Technische Universität München
Fakultät für Chemie
Lehrstuhl für Theoretische Chemie

The Electrostatic Gap: Combining Electrostatic Models with Machine Learning Potentials

Carsten Gerald Staacke

Dissertation

Technische Universität München
TUM School of Natural Sciences

The Electrostatic Gap: Combining Electrostatic Models with Machine Learning Potentials

Carsten Gerald Staacke

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Priv.-Doz. Dr. Aras Kartouzian

Prüfer*innen der Dissertation:

1. Prof. Dr. Karsten Reuter
2. Prof. Dr. Jennifer Rupp
3. Prof. Dr. Volker Deringer

Die Dissertation wurde am 31.08.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 12.10.2022 angenommen.

Concordia res parvae crescunt, discordia maximae dilabuntur.

Preface

This dissertation is publication-based, meaning its scientific content is published in a series of related, but independent articles, all of which have undergone the scientific peer-review process in international scientific journals. The first chapters therefore mainly serve as an introduction to methods and relevant literature. Summaries for each article are then provided in chapter 5. The presented work has been carried out at the Chair of Theoretical Chemistry of the Technical University of Munich (TUM) between March 2019 and December 2020, under the supervision of Prof Dr. Karsten Reuter and has been completed between January 2021 and May 2022 at the Fritz Haber Institute of the Max Planck Society in Berlin. A research stay between January 2022 and Mai 2022 hosted by Prof Dr. Gábor Csányi at the University of Cambridge complemented this work.

Munich, August 2022

Abstract

Growths in the economy and population, a modern lifestyle, and a fully digitalized and connected world increase the global energetic demands each year. Currently, fossil fuels make up 80% of the global energy consumption, the combustion of which being the main driving force for the disastrous effects of climate change. Controlling and reducing CO_2 emissions are therefore key challenges of modern society. Renewable energy sources such as wind and solar panel would be ideal solutions to this problem, however both are a.) not necessarily predictable and b.) not evenly distributed geographicaly. To enable an even energy distribution, we require efficient energy storage. In the past, the combustion of coal and oil has been so successful as that is what carbon-based chemicals are: extraordinarily efficient forms of energy storage.

Many applications, such as laptops, mobile phones, and electric vehicles, utilize lithium-ion batteries as their primary energy storage. While lithium-ion batteries using liquid electrolytes entered the market in 1991, all-solid-state lithium-ion batteries (ASS-LIB), although investigated for decades, are still not widely applied. They promise several advantages in comparison to liquid electrolyte batteries: minimizing fire hazards, longer cycle lifetimes, more comprehensive temperature ranges, and enhanced energy density by potential usage of Li metal anodes. In particular, solid electrolytes of the Li_2S-P_2S_5 (LPS) material class have gained substantial attention due to their favorable properties. First, they possess high RT conductivities of up to 10^{-2} S/cm for crystalline LPS components, which ranks them among the most conductive solid electrolytes. Secondly, they are composed of the earth-abundant elements sulfur and phosphorous enabling applications at large scales. However, this material class' design of potent SSE is hampered by the poor understanding of structure-property relations. This manifests in massive deviations in reported Li-ion conductivity in different experimental setups and from theory and experiment.

Simulations based on Density Functional Theory (DFT) or classical force fields (FF) have enabled material comprehension e.g. new insights into material properties for decades. Insights at the atomistic level are irreplaceable for a mechanistic understanding of chemical processes. Unfortunately, due to high computational costs, DFT methods are limited to small systems while providing a highly accurate and complete description. At a much reduced computational cost, classical FFs allow to account for such effects. Yet, here the problem is an often reduced accuracy in the description of the potential energy surface (PES). To this end, emerging Machine Learning (ML) methods have shown to be increasingly able to bridge this gap, with good first-principles accuracy at a much reduced computational cost. However, the basic assumption of locality, implying the neglect of long-range interactions, is problematic in many cases.

To this end, the central topic of this thesis is threefold. First, we intended to systematically identify systems where this locality assumption does not hold. We especially tried to understand when and why the locality assumption holds for polar and ionic systems and when it fails. Second, as we realized that local ML models accurately predict isotropic bulk material properties, we developed a near-universal Gaussian Approximation Potential (GAP) model for the crystalline and amorphous compounds in Li_2S-P_2S_5. We then used the GAP model to systematically investigate the effect of the local anion composition in glassy Li_2S-P_2S_5 compounds.

At the same time we realized that a short-range model can accurately describe isotropic systems, we understood that we need an accurate description of non-local interactions for non-isotropic systems. To this end, we developed the kernel-based charge equilibration scheme called kQEq. The novel kQEq schemes enable the prediction of partial charges based on local environments by including the ability to predict non-local charge transfer.

Zusammenfassung

Wirtschafts- und Bevölkerungswachstum, ein moderner Lebensstil, in seiner Gesamtheit eine vollständig digitalisierte und vernetzte Welt, erhöhen den Ausstoß von Kohlenstoffdioxid jedes Jahr. Die Verbrennung fossiler Energieträger in Industriesektoren und im Individualverkehr sind dabei der Hauptgrund für katastrophale Auswirkungen sowohl auf den Planeten Erde als auch auf uns als Menschheit. Daher ist die Reduktion von CO_2-Emissionen die größte Herausforderung, vor der die Welt heute steht. Bereits heute stammen 20% des weltweiten Energieverbrauchs aus erneuerbaren Energien. Bei der Energiegewinnung aus nachwachsenden Rohstoffen stehen wir vor dem Problem, dass Wind und Sonne a.) nicht vorhersagbar und b.) nicht gleichmäßig über die Erde verteilt sind. Um eine gleichmäßige Energieverteilung zu ermöglichen, benötigen wir daher effiziente Energiespeicher. In der Vergangenheit war die Verbrennung von Kohle und Öl so allgegenwärtig, da kohlenstoffbasierte Chemikalien genau das sind: ein außerordentlich effizienter Energiespeicher.

Lithium-Ionen-Batterien werden häufig als Energiespeicher der Wahl eingesetzt. Laptops, Mobiltelefone und Elektroautos wären ohne Lithium-Ionen-Akkus nicht realisierbar gewesen. Während Lithium-Ionen-Batterien mit flüssigen Elektrolyten 1991 auf den Markt kamen, haben All-Solid-State-Lithium-Ionen-Batterien trotz jahrzehntelanger Erforschung noch immer keine breite Anwendung gefunden. Sie versprechen mehrere Vorteile im Vergleich zu Flüssigelektrolytbatterien: höhere Leistungsdichte, Minimierung von Sicherheits- und Brandgefahren, längere Zyklenlebensdauer, umfassendere Temperaturbereiche und Erhöhung der Energiedichte durch die potenzielle Verwendung von Li-Metall-Anoden. Festkörperelektrolyte der Materialklasse Li_2S-P_2S_5 haben hier aufgrund ihrer bemerkenswerten Eigenschaften große Aufmerksamkeit erlangt. Erstens besitzen sie hohe Leitfähigkeiten von bis zu 10^{-2} S/cm, womit sie zu den leitfähigsten SSEs zählen. Zweitens bestehen sie aus den auf der Erde reichlich vorkommenden Elementen Schwefel und Phosphor und ermöglichen daher Anwendungen in großem Maßstab. Allerdings wird das Design potenter SSEs dieser Materialklasse durch das schlechte Verständnis der Beziehung von Struktur und Materialeigenschaft behindert. Dies zeigt sich z.B. in großen Abweichungen von Li-Ionen Leitfähigkeit aus Theorie und Experiment.

Simulationen basierend auf der Dichtefunktionaltheorie (DFT) oder klassischen Kraftfeldern (force fields, FF) beschleunigen die Materialanalyse seit Jahrzehnten. Simulation und Modellierung auf atomarer Ebene sind für das mechanistische Verständnis chemischer Prozesse unersetzlich. DFT-Methoden sind auf kleine Systeme beschränkt, liefern aber trotz hoher Rechenkosten eine sehr genaue und vollständige Beschreibung. Obwohl der Rechenaufwand bei der Verwendung eines klassischen FF gering ist, liefern diese hingegen oft nur eine vereinfachte Beschreibung eines Materials. Hier haben sich neue Methoden des maschinellen Lernens (ML) als zunehmend in der Lage erwiesen, diese Lücke zu schließen, indem sie die Genauigkeit der Dichtefunktionaltheorie bei stark reduzierten Rechenkosten ermöglichen.

Die Grundannahme der Lokalität, die die Vernachlässigung langreichweitiger Wechselwirkungen impliziert, ist in vielen Fällen problematisch. Zu diesem Zweck ist das zentrale Thema dieser Arbeit in drei Fragestellungen aufgeteilt. Zunächst war es notwendig Systeme zu identifizieren, bei denen diese Lokalitätsannahme nicht ausreicht. Der Fokus lag insbesondere darauf, wann und warum die Lokalitätsannahme für polare und ionische Systeme wie SSEs gilt und wann sie versagt. Im zweiten Schritt haben wir, als wir erkannten, dass eine genaue Beschreibung isotroper Materialien durch ein lokales ML-Modell erreicht werden kann, ein nahezu universelles Gaussian Approximation Potential (GAP) für die kristallinen und amorphen LPS Materialien entwickelt.

Wir haben das GAP-Model verwendet, um systematisch den Einfluss der lokalen Anionenzusammensetzung auf die Li-Ionen-Leitfähigkeit in glasartigen Li_2S-P_2S_5-Verbindungen zu untersuchen. Genauso wie wir erkannten, dass isotrope Systeme durch ein lokales Modell nahezu exakt beschrieben werden können, haben wir verstanden, dass wir eine genaue Beschreibung nicht-lokaler Wechselwirkungen für nicht-isotrope Systeme benötigen. Hierfür haben wir das Kernel-basierte Ladungsmodel namens kQEq entwickelt. kQEq ermöglicht die Vorhersage von Partialladungen basierend auf lokalen atomaren Umgebungen.

Contents

Nomenclature

As this thesis combines the nomenclature of different fields I will use the following notation throughout this thesis. It is closely related to the original publications i.e. the publication where the underlying theory has been introduced for the first time.

χ^E	Electronegativity
$\chi^{E,0}$	Electronegativity of the isolated atom
χ^E_{CENT}	Environment dependent electronegativity used in the CENT approach
χ^d	Descriptor vector
J_{ii}	Hardness of atom i
J^0_{ii}	Hardness of an isolated atom i
r_{ij}	Distance between atoms i and j
r^A	Atomic radius
q_i	Charge of atom i
Q_{tot}	Total Charge
$\rho_{i,\alpha}$	Element specific neighbour density in SOAP
w	Weighting in the element-specific neighbor densities
$\mathbf{p}$	Power spectra vector
p	Elements of the power spectra vector $\mathbf{p}$
c	Regression weight
R_n	Radial basis function
Y_{lm}	Spherical Harmonics
f_{cut}	cutoff function
k^d	Kernel function using descriptor vectors d
$\mathbf{K}_{NN}$	Kernel matrix of dimension N x N
λ	Regularization parameter
Σ_{NN}	Diagonal matrix containing the regularization parameters λ
N	Number of input parameters
N_t	Number of training points
U_E	Electrostatic potential energy
k_C	Coulomb constant
$E(+1)$	Energy equivalent to removing an electron
E_{tot}	Total energy
$\epsilon^d(\chi^d)$	local energy corresponding to a descriptor d
$\delta^{(d)}$	Scaling parameter of descriptor $d \in$ [2b,3b,SOAP]

List of Abbreviations

ASS-LIB	All-solid-state Lithium ion battery
CENT	Charge equilibration via neural network technique
DFT	d Density Functional Theory
EA	Electron Affinity
EN	Electronegativity
FF	Force Field
GAP	Gaussian Approximation Potential
GPR	Gaussian Process Regression
IP	Ionisation potential
kQEq	Kernel QEq
KRR	Kernel ridge regression
LPS	LiPS material class
ML	Machine Learning
NN	Neural Network
PES	Potential energy surface
QEq	Charge equilibration scheme by Rappe and Goddard
RT	Room temperature
SSE	Solid-state electrolyte

1 Introduction

"Our dangerous reliance on carbon-based fuels is at the core of all these problems- economics, environmental, national security." [1]

In his Nobel prize lecture in 2007, Al Gore pinpointed the need for an immediate carbon-free economy. In times of inner European aggression, Al Gore seems right. More than ever.

Thirteen years later, clean energy technologies' worldwide market lies at approximately 130 billion USD only.[2] Nevertheless, there are signs of improvement. Whilst the global economy suffered from the impact of COVID-19,[3] renewable energies expanded by their fastest rate in two decades.[2] Record sales of electric vehicles are set daily.[4] Policy changes, such as a solar panel mandate for new non-residential buildings in Germany [5, 6], indicate that the energy economy of future generations will be unrecognizable to our current one.[7–9] Although sustainable energy in our economy is without alternative, it is unclear what the supply with solely regenerative energy will look like.

Batteries are by far the largest part of the clean energy technology market, a market expected to increase to 1.2 trillion USD by 2050.[2] The current majority of lithium-ion batteries (LIBs) utilize liquid electrolytes, which are in competition with All-solid-state (ASS)-LIBs.[8, 10–12] In theory, these are advantageous over liquid-electrolye LIBs, as they minimize safety and fire hazards, have longer cycle lifetimes and more comprehensive temperature ranges, and enhance energy density via the potential use of Li metal anodes.[11, 12]

A promising subgroup of ASS-LIBs is the Li_2S-P_2S_5 (LPS) material class.[13] First, they possess high conductivities of up to 10^{-2} S/cm, even higher than the well known electrolytes $Li_{10}GeP_2S_{12}$ or $Li_{1.3}Al_{0.3}Ti_{1.7}(PO_4)_3$.[14, 15] Secondly, they consist of earth-abundant elements sulfur and phosphorous.[16] A critical, yet often neglected factor, as abundance is required for sustainable large-scale global implementation. [8]

When first commercial energy storage devices entered the market, sustainable energy research targeted specific energy technologies and related materials.[17] While there is an ongoing hunt for the jack-of-all-trades material,[18] design of potent SSE is typically hampered by the poor understanding of structure-property relations.[19] Hence, efforts such as the *e*-conversion cluster or the BIG-MAP project target disorder and materials interfaces that underlie these changes of material functions.[20] Examples of successful material improvements by a mechanistic understanding range from atomic-scale complexions [21, 22], via interface amorphization processes that improve stability and capacity of batteries [23–25], to nanoscale disorder in solar cell materials.[26, 27]

Likewise, using LPS on a large scale is hindered by a poor understanding of Li-ion conductivity mechanisms, amorphization and degradation processes.[13] From a modeling perspective, investigating these observations realistically at the atomistic level strains the capabilities of state-of-the-art theoretical approaches.[28] On one hand, the system sizes and simulation time scales required are prohibitive for first-principles methods such as density functional theory (DFT).[29, 30] That simply means that the sheer number of calculations, requires simulation approaches multiple orders of magnitudes faster than standard DFT. On the other hand, parameterizations for

empirical potentials are often not available, and these potentials may ultimately lack the desired predictive accuracy. Fortunately, modern machine learning (ML) potentials are increasingly able to bridge this gap, promising first-principles accuracy at a much reduced computational cost.[28] In the past ML potentials have increasingly been a standard tool for atomistic simulations. The most prominent examples are Neural Networks (NN) [31–36] and Gaussian Approximation Potentials (GAP)[37–40], which enable simulation with nearly DFT accuracy while accelerating simulations almost reaching the speed of classical empirical potentials in some cases.[28, 30] As indicated by their name, empirical potentials gain their speed from a simple fixed analytical form.[41] Flexibility and transferability are dominated by choice of the functional form of the potential energy surface (PES).[42] Opposed to that, ML potentials do not have a fixed functional form, but instead Machine Learning is used to learn an approximated PES as accurately as needed. [30]

When first introduced, ML mehods such as NN used a fixed structure. The NN was then optimized for a certain number of degrees of freedom, i.e., number of atoms.[32] These cannot be used to predict energies for a different system size/different number of atoms, since the optimized weights are valid only for a fixed number of input nodes of the NN. In order to overcome this scaling limitation, the locality assumption was introduced: The total energy E of the system can be defined as a sum of atomic contributions ϵ_i and hence the energy associated with a given atom depends on its immediate environment but not on atoms outside a given cutoff radius.[32, 43, 44]

However, this locality approximation implies that long-range contributions arising, e.g., from electrostatic interactions, are neglected beyond a certain cutoff.[37] Although the need to include long-range electrostatics in a similar fashion as in empirical potentials appears straightforward at first glance, the success of short-range ML potentials for the modeling of certain properties of ionic and polar materials appears to say something different.[45–49] Therefore, the first task was to identify a system and simulation tasks where the locality assumptions hold in some cases but fail in others. To study the role of long-range effects, GAP were constructed with and without an electrostatic correction term. As we studied the role of anisotropy, we realised that current electrostatic and charge equilibration models lack the required flexiblity. Hence, we extended the classical charge equilibration model QEq by an environment-dependent electronegativity and showed that our kernel QEq model (kQEq) can be used to generate accurate and highly data-efficient models for molecular dipole moments.

Chapter 2 will introduce the general concepts of SSEs specifically, materials of the Li_2S-P_2S_5 class, and discuss current challenges. Chapter 3 focuses on the two parts characteristic for ML potentials, namely the atomic descriptor and the concept of regression. I discuss the GAP framework and give a tutorial like introduction to the regression methods used in this work. The final chapter 4 focuses on the derivation of the charge equilibration and long-range interaction and the discussion of previously published machine learning charge prediction schemes.

2 Li_2S-P_2S_5 Solid-State Electrolytes and their challenges

2.1 Solid-State Electrolytes

Lithium-ion batteries (LIBs) are currently the most efficient electrochemical energy storage technology in terms of energy and power densities, reliability, and rechargability. While LIBs with liquid electrolytes reached commercialisation at the beginning of the 90s, all-solid-state Lithium-ion batteries (ASS-LIBs) are still not widely used. ASS-LIBs promise several advantages, regarding reduced flammability and operation safety in electrified mobility, longer lifetimes, and higher energy density. [11, 12]

A common drawback of liquid electrolytes are side reactions in which soluble products are generated at one electrode and consumed or further reacted at the other electrode, often called chemical cross-talk.[50] A well known example is the Li-S shuttle effect.[51, 52] Soluble electrode components can diffuse, being responsible for the progressive leakage of active material from the cathode.[53] In solid-state electrolytes (SSEs) only lithium ion transfer is exhibited as the SSE acts as a functional separators with only resulting in minor self-discharge. The negligible self-discharge is typically attributed to a low residual electronic conductivity.[21] In liquid electrolytes, most compounds (Li ions and most anions) are mobile. The mobility of all chemical species can cause concentration gradients of the conducting salts. This salt gradient leads to bulk polarization, limiting the cell current. As in SSBs only Li ions are mobile, higher current densities can be reached and hence lower charging times are enabled. [11, 14] A desired feature of SSEs is the use of lithium-metal anodes. The hope was that due to mechanical rigidity of SSEs dendrite formation caused by electrodeposition of lithium can be prevented.

SSEs can be divided into three groups. Organic solid polymers, inorganic solids, and solid like dispersion of nanoparticles in liquids.[54] The last group is often referred to as semi-solid electrolytes and form a group on their own. Although solid polymer electrolytes appear to be the preferred choice as they can compensate for volume changes of electrodes by elastic and plastic deformation, they lack the required high ionic conductivities for battery operation.[55] Batteries using inorganic solids — either crystalline, glass or glass-ceramic in nature — have demonstrated to improve battery performance at high currents.[56] This is in contrast to a common misconception that SSEs are inherently poor ionic conductors at ambient temperature.[56] In fact a number of ternary and quaternary sulfides and thiophosphates (Li_2S-P_2S_5 and $Li_{10}GeP_2S_{12}$) have been reported to exhibit room temperature conductivities equal or even higher than typical liquid electrolytes.[14, 57, 58]

So why is it that ASSBs are not fully applicable already? The major drawback of many inorganic SSEs is their low thermodynamic stability.[59] Just like liquid electrolytes, SEs are easily oxidized at intermediate potentials and reduced at low potentials.[60] Protecting interfaces are therefore required to stabilize the electrode/electrolyte contact.[61] While oxides often experience mechanical failure through cracking, thiophosphates such as the herein studied (Li_2S-P_2S_5) are

ductile and easily form dense cathode composites.[62]

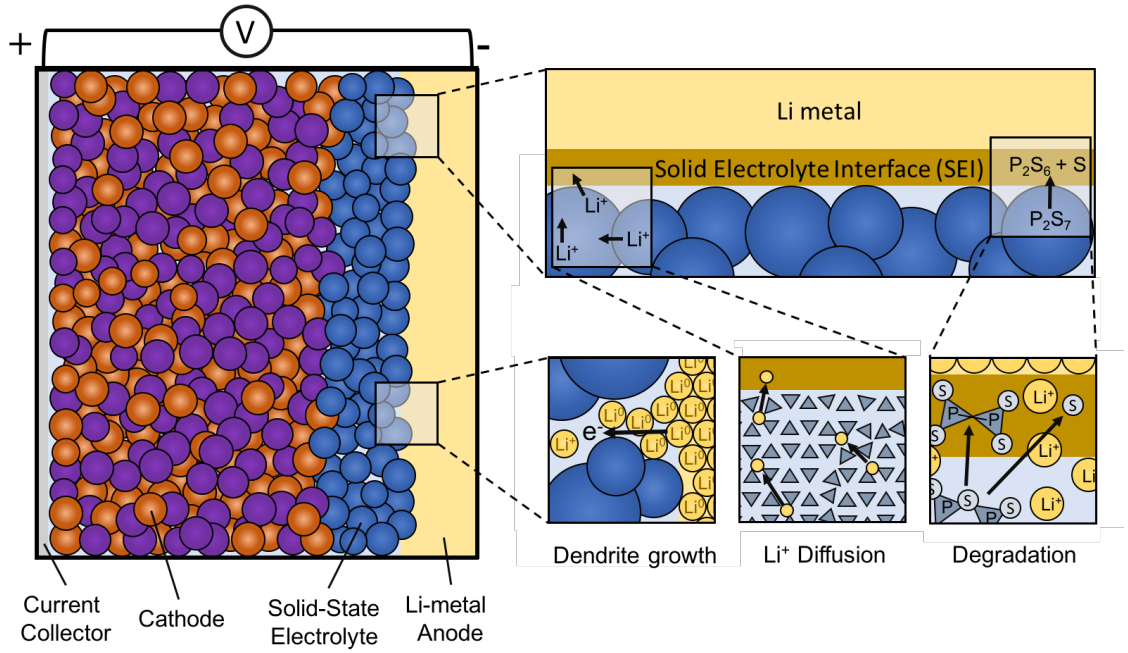


Fig. 2.1: *(Left): Schematic architecture of a solid-state battery adapted from Zeier and Janek.[11] Cathode and anode are separated by the solid-state electrolyte that allows for lithium-ion diffusion. Porous cathodes typically made of layered transition metals are the largest component of a battery. Anodes and cathodes both serve as the active storage component of the battery, but known cathode materials have a lower Li density. The cathode is typically coated on thin aluminium foils (current collector). In a SSB the classically used porous graphite anode could be replaced by Lithium metal. (Right top): Major challenges in modeling solid-state batteries: The use of a lithium metal anode can significantly increase the cell energy density resulting in a dramatically increased cell energy density. However, resistive solid electrolyte interfaces (SEI) may form between the lithium anode and the SSE (Right top). Inhomogeneous lithium metal deposition can form dendrites. Dendrite formation represents a simulation task that requires an accurate description of charge-transfer plus the ability to model SEI. Modeling lithium ion diffusion within a solid and over a solid-solid interface strains the capability of modern simulation techniques. Near ab inito accuracy for long and large simulation cells are required to determine a full picture of lithium ion motion. When modeling amorphization and the formation of interfaces, it is crucial to describe reactivity in order to predict degradation processes that happen during charging and discharging of a battery.*

In Fig. 2.1 I give a schematic architecture of a solid-state battery as proposed by Zeier and Janek and including current challenges in modeling solid-state batteries.[11] Already during synthesis but especially during operation, SSE form resistive solid electrolyte interfaces (SEI) between the lithium anode and the SSE. Although material interfaces can be engineered so they form a protective layer, realistic models are challenging in both cases.[21] Secondly, inhomogeneous lithium metal deposition in the anode as well as in the SSE itself can form dendrites.[63] For modelling dendrite formation we require an accurate description of charge-transfer plus the ability to model disorder in both electrodes and the SSE. The pure size of realistic simulation setups as well as the simulation time in order to realistically model lithium ion diffusion within a solid and over a solid-solid interface strains the capability of modern simulation techniques. Near *ab initio*

accuracy, in long and large simulation cells are required to determine a full picture of lithium ion motion. Finally, when modeling disorder, defects, amorphization and the formation of interfaces, it is crucial that the underlying model is reactive and hence can predict degradation processes that happen during charging and discharging of a battery.

The key to further advance in the ASS-LIB field is to fully understand material disorder, interfacial properties and eventually feature a lithium metal electrode and outperform conventional lithium ion batteries.[11] In this light, it is evident that the development of Li ion batteries doesn't hinge on the development of appropriate solid bulk electrolyte materials, but a systematic understanding of material properties. In the following chapter I will summarize the structural and ionic conduction properties of the Li_2S-P_2S_5 material class, highlighting the challenges for an atomistic understanding of these materials.

2.2 The Li_2S-P_2S_5 Solid-State Electrolytes

Sulfur based electrolytes gained attention as SSEs due to an inherent high ionic conductivity (>10^{-3} Scm^{-1}), [64] and their good contact with electrode materials due to their mechanical soft nature. [65] The thiophosphate electrolytes in the "simple" Li_2S-P_2S_5 two component system (LPS family) are particularly interesting as they possess high conductivities without the addition of a transition metal (Si, Ge, Sn).[58] Accordingly, a variety of crystalline and amorphous materials in the LPS material class were reported over the past two decades.[13] In literature two nomenclatures for the LPS material class are found. Either the material is characterized by its chemical formula or by its mass percentage of Li_2S, i.e. Li_3PS_4 or 75 mol% Li_2S.

Although crystalline and amorphous materials are - in principle - of infinite structural diversity, in LPS five anionic species are commonly observed. As illustrated in Fig. 2.4, these species are characterized by central phosphorus atoms, each bonded to either four sulfur atoms or a neighbouring phosphorous atom and three sulfurs.

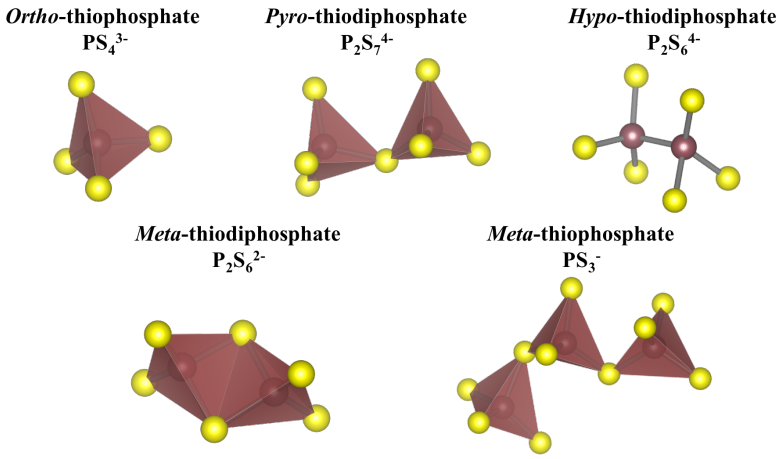


Fig. 2.2: *P-S microchemistry/ anionic species formed within Li_2S - P_2S_5.[13] Phosphorous is displayed in red, sulfur in yellow.*

These features were identified and characterized in detail for the first time by Dietrich *et al.* [58] in 2017 and can be summarized as follows:

- *Ortho*-thiophosphate moieties $\mathbf{PS_4^{3-}}$ (tetrahedra) are dominant for high Li_2S quantities (>75 mol%).
- *Pyro*-thiophosphate moieties $\mathbf{P_2S_7^{4-}}$ are formed by two corner sharing PS_4 tetrahedra. They are typically observed for Li_2S quantities <75 mol%.
- *Hypo*-thiodiphosphate moieties $\mathbf{P_2S_6^{4-}}$ are composed by two PS_3^- units with a direct P-P bond. It should be noted that phosphorus in $P_2S_6^{4-}$ has a formal oxidation state of +IV, whereas its formal charge in the rest of the LPS anions is +V.

- *Meta*-thiodiphosphate moieties $\mathbf{P_2S_6^{2-}}$ and *Meta*-thiophosphate $\mathbf{(PS_3^-)}_n$ are both observed in the 60 mol% Li_2S crystalline composition. *Meta*-thiodiphosphate has two edge sharing PS_4^{3-} units, *Meta*-thiophosphate moieties are described as the polymeric corner-sharing chains of PS_3^-.

Crystalline Phases of Li_3PS_4, $Li_4P_2S_6$, and $Li_7P_3S_{11}$

In the ternary Li-P-S phase diagram, Li_3PS_4, $Li_4P_2S_6$, and $Li_7P_3S_{11}$ are probably the most commonly studied compounds.[13]

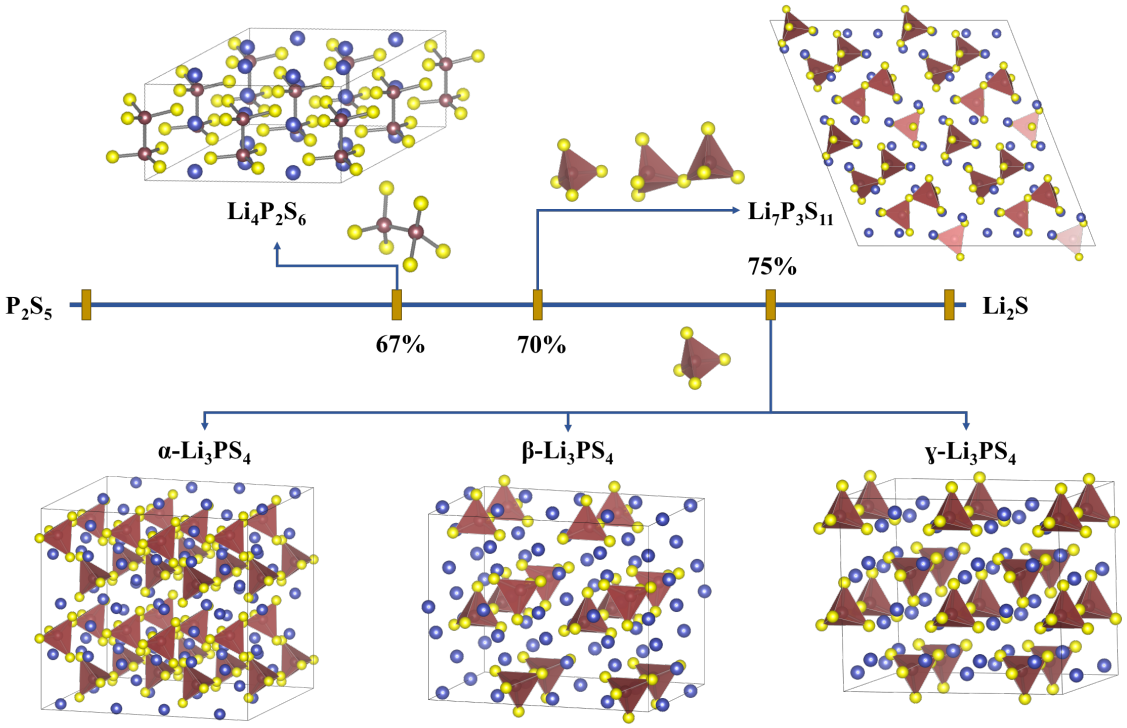


Fig. 2.3: *Crystal structures of the Li_2S–P_2S_5 composition line. The structures are grouped by their local P–S motifs. Note that $Li_4P_2S_6$ does not exactly lie on the Li_2S–P_2S_5 composition line, but is the crystallization product of glassy $Li_4P_2S_7$. (Li: blue; S: yellow; P: red)*

Li_3PS_4: Li_3PS_4 is the crystalline compound of Li_2S content of 75 mol% and can be found in three different phases: α, β, and γ -Li_3PS_4. All of them solely contain the simplest PS_4^{3-} anion. As seen in Fig. 2.4 conceptually, these three phases can be distinguished by their different arrangement of PS_4^{3-} anions. These are either all pointing in the same direction (γ), are arranged in a zig-zag fashion in one (α) or two directions (β) in space.[66] The room temperature stable phase is γ -Li_3PS_4, exhibiting a phase transition at 600K (γ to β) and at 800K (β to α).[67] As operating temperatures of ASSBs are usually smaller than 400K the α phase is less relevant for battery applications as it is not stable at RT. Apart from that, due to small stoichiometric changes, β-Li_3PS_4 has been shown to occur at RT. [68]

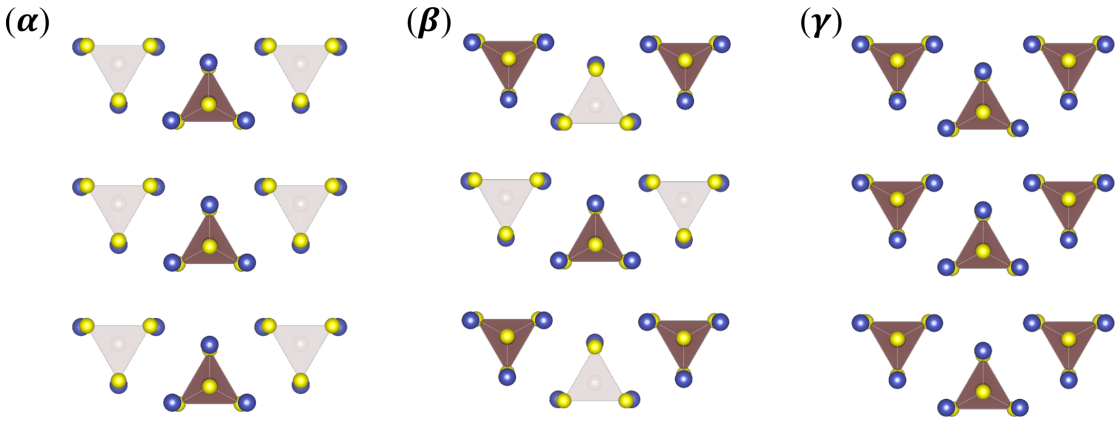


Fig. 2.4: *Arrangements of PS_4 tetrahedra in the α, β, and γ in Li_3PS_4.[69] (Li: blue; S: yellow; P: red)*

A fourth crystalline Li_3PS_4 called δ, was predicted by Iikubo *et al.* using an evolutionary algorithm under high-pressure of 5GPa.[70] To the best of my knowledge, δ-Li_3PS_4 has not been experimentally observed. Low RT-conductivities of 2.6 x 10^{-7} and 9.0 x 10^{-7} S cm^{-1} are experimentally reported for both relevant crystals (γ to β).[67]

$Li_7P_3S_{11}$: Crystalline LiP_3S_{11} is obtained for 70 mol% Li_2S. It is an extremely important member and well studied crystal of the LPS family due to its very high ionic conductivity (up to 1.7 x 10^{-2} Scm^{-1} at RT).[71] It has a triclinic P-1 space group, composed by a 1:1 ratio of PS_4^{3-} and $P_2S_7^{4-}$. Lithium ions are exclusively tetrahedrally coordinated.[72] Obtaining fully crystalline $Li_7P_3S_{11}$ is challenging due to its narrow stability window.

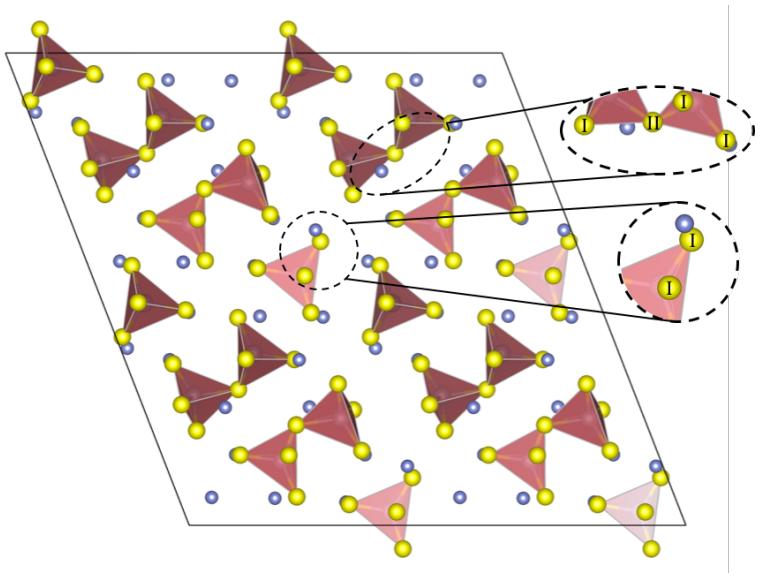


Fig. 2.5: *Crystallographic structure of $Li_7P_3S_{11}$. It features terminal (I) as well bridging (II) sulfur atoms. (Li: blue; S: yellow; P: red)*

As seen in Fig. 2.5 $Li_7P_3S_{11}$ features two distinct sulfur species, namely in bridging and terminal

positions. Our evaluations using Hirshfeld population analysis [73] indicate that these species correspond to different charge states. It is generally obtained as a glass-ceramic. [74] While Li-ion conductivity in LPS is usually determined via a diffusion of defects (Li^+ vacancies), $Li_7P_3S_{11}$ exhibits a more collective Li^+ motion yielding superior conductivity compared to the other of its crystalline counterparts.[75, 76]

$Li_4P_2S_6$ Synthesizing $Li_4P_2S_6$ crystals is unique as its composition is not exactly in the stoichiometric line between Li_2S and P_2S_5. Li_2S contents of 67 mol% yields $Li_4P_2S_7$ glass. The corresponding reaction is:

$$2Li_2S + P_2S_5 \rightleftharpoons Li_4P_2S_7 \rightleftharpoons Li_4P_2S_6 + S$$

The crystal structure of $Li_4P_2S_6$ was initially reported as being P6 3/mcm when first synthesized in 1982,[77] recent studies predicted a stacking of $P_2S_6^{4-}$ yielding P-31 m.[78] Quenching the Li_2S contents of 67 mol% melt yields glassy $Li_4P_2S_7$. Subsequent annealing leads to the formation of crystalline $Li_4P_2S_6$ and sulphur.[77, 78] It is important to emphasize, that the local structures of glassy $Li_4P_2S_7$ and crystalline $Li_4P_2S_6$ differ significantly. The dominant anion of the $Li_4P_2S_7$ glass is $P_2S_7^{4-}$, while as shown in Fig. 2.6 in crystalline $Li_4P_2S_6$ the material solely contains $P_2S_6^{4-}$.[58, 78] $Li_4P_2S_6$ is a commonly observed byproduct of other members of the LPS material class.[78]

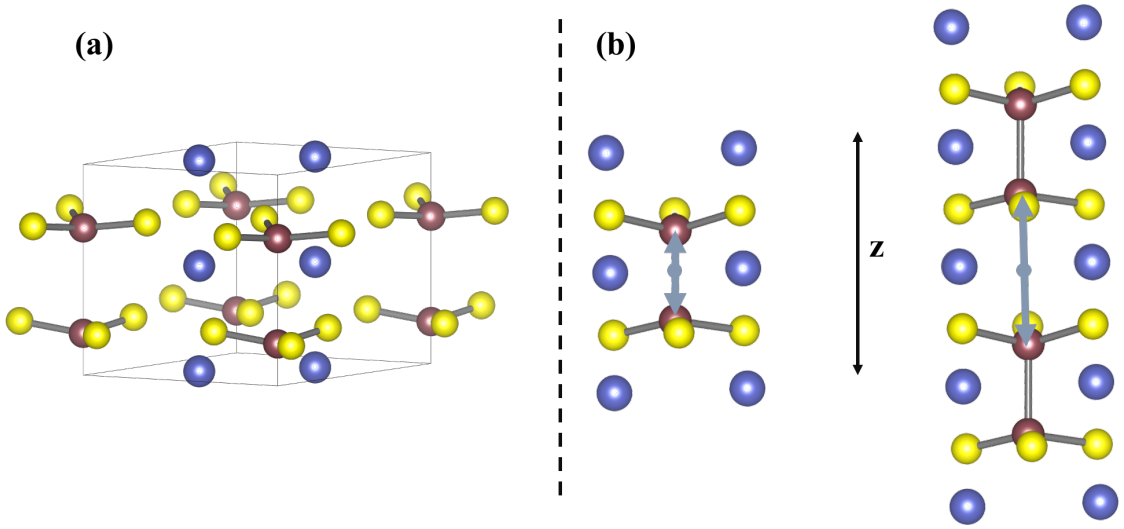


Fig. 2.6: *(a) rystallographic structure of $Li_4P_2S_6$ with a D_3d symmetry. The P–P bonds are along the crystallographic z axis. (b) Ball-and-stick diagram of the P_2S_6 ion units. The blue arrows indicate the two possible placements of the P.[79] The origin of the unit cell can either be at the center of the P-P bond or located between P_2S_6. (Li: blue; S: yellow; P: red)*

Especially, the metastable $Li_7P_3S_{11}$ crystal degrades at high temperatures above 1000K leading to the formation of $Li_4P_2S_6$.[80] Indeed, these reactions can be attributed to the evaporation of $Li_4P_2S_7$ and the high stability of the $Li_4P_2S_6$ phase. The ionic conductivity of $Li_4P_2S_6$ is very low (10^{-6} Scm^{-1}) but can be enhanced when synthesized as a glass ceramic structure.[58] In the glass-ceramic microstructure, the amorphous part is mainly composed of PS_4^{3-} units.[58]

Glasses and Thiophosphate Microchemistry

While LPS crystals have fixed stoichiometries and P-S anion matrices, high structural variability is found in LPS glasses.[72, 78] The plasticity arising from the amorphous nature of the glasses makes them more robust to mechanical stress, exerted by volume changes during cell cycling.[11]

Amorphous structures of $x$$Li_2S$-(100-$x$)$P_2S_5$ have been synthesized and systematically studied in the range $60 < x < 80$.[78] Similar to crystalline LPS, the choice of x influences the anion composition. Except for *Meta*-thiodiphosphate ($P_2S_6^{2-}$) all thiophosphate anions were experimentally verified via ^{31}P magic angle spinning nuclear magnetic resonance (MAS-NMR) and Raman measurements.[78, 81, 82] In general though, compositions are dominated by PS_4^{3-} and $P_2S_7^{4-}$ with the corresponding species ratio highly depending on the choice of x. At low Li_2S contents ($60 \leq x \leq 70$) $P_2S_7^{4-}$ is the dominating anion, while higher Li_2S content favours PS_4^{3-} anions.

The $P_2S_6^{4-}$ anion content is affected by the method of synthesis. A ball milling approach yields up to 10 % $P_2S_6^{4-}$ anion independent of x.[78, 83] In contrast to that a microwave assisted synthesis route minimizes the formation of $P_2S_6^{4-}$ anion at RT.[78, 81, 82] Thiophosphate-chains (PS_3^-) have only been demonstrated for low Li_2S contents (x=60).[58, 78]

Essential differences of crystalline and glassy LPS, in addition to anion content, are Li^+ conduction mechanisms. For glassy Li_3PS_4 , the so-called paddlewheel effect has been found to increase the RT conductivity by several orders of magnitude.[84] The paddlewheel effect describes quasi continuous PS_4^{3-} re-orientations during Li^+ migration. Li ion conductivities are between 10^{-5} and 10^{-4} Scm^{-1}. They are less conductive than the $Li_7P_3S_{11}$ crystal, but significantly higher than the crystalline Li_3PS_4 phases.[78–80, 85, 86] An interesting experimentally observed trend in LPS glasses is an increased Li-ion conductivity with increasing Li_2S content.[58]

2.3 Challenges from a Modeling Perspective

Realistic modelling of battery materials is an ongoing journey, challenging time and size scales of all available methods we currently have in our toolbox. Ranging from polaron assisted Li diffusion[87], studies on nano-scale complexions using empricial potentials[21] to kinetic Monte Carlo investigations of Lithium intercalated in graphite[88] our group always tried to push state of the art approaches to its limits. Developments in the field of operando spectroscopy enabled detailed structural insight into batteries during operation. A key finding of operando experiments, not just in battery research but also catalysis and other fields, is the awareness that a static picture of a material is not enough and sometimes even wrong.[89] To establish multiscale relations between atomistic features and a material's macroscopic behaviour and electrochemical properties we need to describe atomic interactions as accurate as possible on a size and time scale that is statistically robust in an ensemble like fashion. So far, the methods of choice are based on DFT for small/short but very accurate simulations and empirical potentials for large/long simulations that are limited in their accuracy. Empirical potentials gain their speed and robustness from fixed functional forms which comes at the cost of lacking reactivity or describing chemical variations in LPS material inaccurately. In that respect my personal experience was always that accuracy in a numerical sense often seemed to be taken too seriously and accuracy in a sense of consistency with experimental observations to be underrepresented.

The approach taken in this thesis, is to replace fixed functional forms of empirical potentials by flexible functional forms of machine learning interatomic potentials in order to combine the accuracy of first principle methods with the speed of empirical potentials. In materials modelling

we are typically interested in the total energy E_{tot} of a system. We can define E_{tot} as a sum of local contributions E_{local} and long-range electrostatic E_{ES} yielding the following energy expression

$$E_{tot} = E_{local} + E_{ES} \tag{2.1}$$

The following two chapters will evolve around this expression of the total energy. While chapter 2 focuses on the local energy contribution E_{local} retrieved from the GAP framework, chapter 3 derives the E_{ES} expression in the context of the charge equilibration scheme QEq. Hence, the underlying structure of each chapter is always the same: Starting from the energy expression of either the local or electrostatic term each section introduces key steps how these energy expressions are obtained.

3 Machine-learning interatomic potentials in materials science

3.1 Machine Learning Potentials: A general Overview

In principle, the most accurate way to obtain total energies and atomic forces of a system is by solving the Schrodinger equation (SE).Unfortunately, an analytic solution is only feasible for very-simple systems such as the hydrogen atom. For larger chemical structures, the SE is typically solved approximately. However, even with approximations, an accurate numerical solution of the SE is a computationally extraordinary demanding task.

In the past, simple empirical functions are commonly used to model the relevant interactions. From these FF, energies and forces can be obtained with much reduced costs. However, while offering a qualitatively reasonable description of chemical interactions, the accuracy of the underlying FF and hence the quality of the simulations can be very limited.[90] ML methods could bridge this gap between accuracy of *ab initio* methods and efficiency of classical FFs. When using ML methods the user aims to train an algorithm to learn the functional relationship between inputs (chemical descriptors) and outputs (properties) from patterns or structure in the training data.

As we already introduced, in order to create a general ML potential that can be employed for systems of varying size and composition, just as with many empirical potentials (e.g. EAM, Tersoff) a locality assumption is typically made.[32, 38, 91, 92] The system's total energy is thus approximated purely as a sum of local (atomic) contributions:

$$E_{tot} \approx E_{local} = \sum_{i}^{N} \epsilon(Z_i, \chi_i) f_{cut} \tag{3.1}$$

where the sum runs over the N atoms in the system and each atom i contributes with an energy ϵ that only depends on its atomic number Z_i and its local chemical environment, represented by the descriptor χ_i. [32, 38, 93–95] This implies that electrostatic contributions outside a cutoff, in the above function defined by a cutoff function f_{cut}, are negligible.

Three components are needed to generate an ML potential for a given material/material class:

- A database of reference structures and associated quantum-mechanical data.
- A way to represent the atomic structure such that these can be used by the ML algorithm
- The regression or "learning" algorithm itself.

While I discussed the target materials in the previous chapter, I will discuss tasks two and three in the following chapter.

3.2 The Gaussian Approximation Potential Framework

In this work I used the Gaussian approximation potential (GAP) framework.[38, 39] The GAP software is implemented in the QUIP code.[96] As well as in other ML frameworks, the total energy of an atomistic system is a sum of atomic ("local") energies, from training data that consist of the system's cartesian coordinates, total energies and their derivatives. In GAP, the two components for modeling are the representation of atomic environments typically using n-body (with n=2,3) descriptors as well as the many-body descriptor smooth overlap of atomic positions (SOAP), and the regression task which is in this framework a Gaussian Process Regression (GPR).

The commonly used energy expression in GAP is

$$\begin{aligned} E_{total} = E_{local} = (\delta^{(2b)})^2 \sum_{i \in pairs}^{N_t} \epsilon^{(2b)}(\chi_i^{(2b)}) \\ + (\delta^{(3b)})^2 \sum_{j \in triplets}^{N_t} \epsilon^{(3b)}(\chi_j^{(3b)}) \\ + (\delta^{(SOAP)})^2 \sum_{a \in atoms}^{N_t} \epsilon^{(SOAP)}(\chi_a^{(SOAP)}) \end{aligned} \tag{3.2}$$

"2b", "3b", and "SOAP" denote two-, three-, and many-body interactions each containing a scaling parameter $\delta^{(d)}$. The scaling parameter defines the energy contributions of a given interaction. To the best of my knowledge, Volker Deringer and Gabor Csanyi first introduced that energy expression for a GAP on amorphous Carbon in 2017.[37] From that energy expression, I will now introduce the ingredients of the GAP framework in a step-by-step manner starting with descriptors/representations of local atomic environments followed by the GPR framework.

3.3 Descriptors of local atomic environments

The set of descriptors $\chi^d = \{\chi_1^d, \chi_2^d, ..., \chi_N^d\}$ with $d \in \{2b, 3b, MB\}$ encode the local environment of every atom i.[97] One possible terminology defines a descriptor by being a mapping of an atomic configuration i, typically a molecule or a solid defined by the cartesian coordinates and chemical identity of its N atoms, into a suitable representation for the regression task.[30] The mapping associates i with points in feature space, which are then used to construct a machine-learning model to regress (fit) a structure-property relation.

One can define four desirable properties/requirements for a structural descriptor: First, the descriptor should obey fundamental physical symmetries. Second, the descriptor should be smooth i.e. continuous changes of a structure should yield a smooth change in the associated descriptor. Third, it should be complete hence inequivalent structures should yield distinguishable descriptors. Finally, in order to be able to ensure transferability to systems of varying molecular size, the descriptor should be additive e.g. structures should be decomposed in a sum of local environments.[98, 99]

The need to remove the dependency of the Cartesian coordinates on the origin and orientation of the reference system, is a key in chemical simulations. Already in classical FF different sets of internal coordinates (bonds, angles, and torsions) have been proposed, based on chemical intuition, as invariant descriptors of molecular geometry. In fact classical FF have been extremely effective

in the modeling of biological systems. When ensuring fundamental physical symmetries this means that the descriptor $\chi^{(d)}$ has to be invariant [1] under symmetry operations and permutation of equivalent atoms in addition to translation and rotation of structures.

In practice, in addition to SOAP, a multitude of descriptors $\chi^{(d)}$ are available and different authors have their favorite descriptor.[32, 38, 93, 94]

3.3.1 Kernel Function

Before we can introduce n-Body Descriptors, we need to highlight the role of kernel functions. In GAP, the local energy corresponding to each type of descriptor $d \in \{2b,3b,MB\}$ can be given by a linear combination of kernel functions

$$\epsilon_i^{(d)}(\chi_i^{(d)}) = \sum_{t=1}^{N_t} w_t^{(d)} K^{(d)}(\chi_i^{(d)}, \chi_t^{(d)}) \tag{3.3}$$

where t denotes one of N_t training configurations $\chi_t^{(d)}$. Each training configuration has a weighting coefficient w_t. The weighting coefficient is attained during fitting. The covariance kernel K quantifies how similar the input configuration $\chi^{(d)}$ is to the training configuration $\chi_t^{(d)}$. In practice, one sparsifies the representation and only allows the sum to range over a number of "representative points" drawn from the full training database ($N_t \ll N_{full}$).

So what is now a kernel? In machine learning, a 'kernel' is usually used to refer to the kernel trick, a method of using a linear classifier to solve a non-linear problem. The kernel transform linearly inseparable data to linearly separable ones. As seen in Figure 3.1 the kernel function K is thus applied on each data point to map the original non-linear observations into a higher-dimensional space in which they become separable. In this work we thus use the following nomenclature: Cartesian coordinates are transformed by basis functions of the underlying descriptor, yielding a descriptor vector. Each element of that descriptor vector is referred to as a feature.

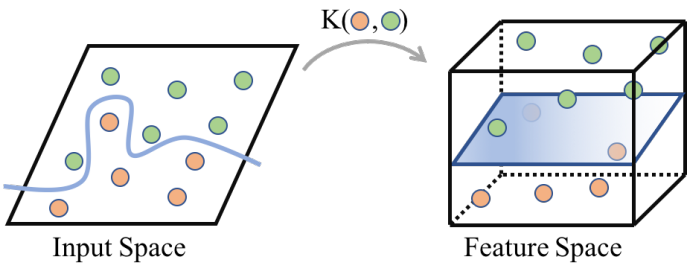


Fig. 3.1: *Graphical representation of the kernel trick: In 2-d, non-separable data (input space) become linearly separable data in 3-d (feature space) after applying the kernel transformation K.*

[1] It is important to note, that $\chi^{(d)}$ should be invariant to the things that the target property is invariant to. For example, energy is a translation and rotationally invariant property, but a molecular dipole isn't.

In the GAP-framework, for 2b and 3b contributions, a squared exponential kernel is

$$K^{(d)}(\chi^{(d)}, \chi_t^{(d)}) = exp\left(- \sum_{\varsigma} \frac{|\chi_\varsigma^{(d)} - \chi_{t,\varsigma}^{(d)}|^2}{2\sigma_\varsigma^2} \right) \tag{3.4}$$

with ς being an index running over the elements of the descriptor vector χ and σ being the width of the exponential kernel. The elements of the descriptor vector are often referred to as features. A polynomial kernel is used in order to compare many-body (SOAP) environments

$$K^{(SOAP)}(\chi^{(SOAP)}, \chi_t^{(SOAP)}) = \left(\frac{\chi^{(SOAP)} \cdot \chi_t^{(SOAP)}}{\sqrt{\chi^{(SOAP)} \cdot \chi_t^{(SOAP)} \chi^{(SOAP)} \cdot \chi_t^{(SOAP)}}} \right)^\zeta \tag{3.5}$$

3.3.2 n-Body Descriptors

In the case of two-body contributions we use the above introduced local energy expression

$$\epsilon_i^{(2b)}(\chi_i^{(2b)}) = \sum_{t=1}^{N_t} w_t^{(d)} K^{(2b)}(\chi_i^{(2b)}, \chi_t^{(2b)}) \tag{3.6}$$

using a squared exponential kernel yielding

$$\epsilon_i^{(2b)}(\chi_i^{(2b)}) = \sum_{t=1}^{N_t} w_t^{(d)} exp\left(- \sum_{\varsigma} \frac{|\chi_\varsigma^{(d)} - \chi_{t,\varsigma}^{(d)}|^2}{2\sigma_\varsigma^2} \right) \tag{3.7}$$

In this case, we simply use distances r_{ij} between atoms i and j, hence the descriptor has one single scalar component and the local energy expression for a two body contribution simplifies to

$$\epsilon_i^{(2b)}(\chi_i^{(2b)}) = \sum_{t=1}^{N_t} \sum_{j=1}^{N_t} w_t^{(d)} exp\left(- \frac{|r_{ij} - r_{jk}|^2}{2\sigma_{2b}^2} \right) \tag{3.8}$$

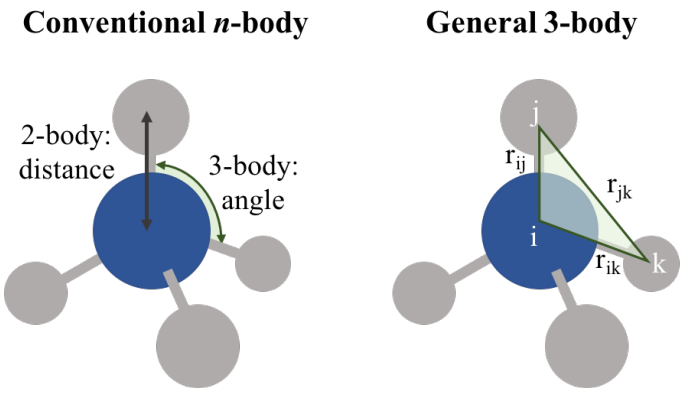


Fig. 3.2: Graphical representation of n-body (here n=2,3) descriptors for a simple methane molecule. 2-body terms encode the chemical environments via atomic distances. 3-body terms add angular information between atoms. Instead of describing 3-body terms via the angle between three atoms, a general 3-body defines angular information by three atomic distances.

The simplest way to guarantee rotational, translational and permutational invariance is to represent an atomic structure via a set of atomic pairwise distances $\mathbf{r}_{ij}$ accounting for all atom pairs i and j.

The two-body contribution $\mathrm{E}^{(2b)}$ can be obtained by comparing all eligible distances $r_{ij} = |r_i - r_j|$ in a molecule or crystal via a smooth and regular kernel function $k^{2b} : \mathbb{R}\mathrm{x}\mathbb{R} \rightarrow \mathbb{R}$

$$k^{(2b)}(r_{ij}, r_{ik}) = exp\left(- \frac{|r_{ij} - r_{jk}|^2}{2\sigma_{2b}^2} \right) \tag{3.9}$$

defining the similarity of distances via a Gaussian function of width σ_{2b}. The width σ_{2b} controls the smoothness of the kernel. The energy $E_i^{(2b)}(\mathbf{r}_{ij})$ of atom i is given by summing local energies, each defined by its kernel values $k^{(2b)}(r_{ij})$

$$E_i^{(2b)}(\mathbf{r}_{ij}) = \sum_j^{N_t} \epsilon_{ij} = \sum_j^{N_t} w_j k^{(2b)}(r_{ij}, r_{ik}) \tag{3.10}$$

with w being the regression weights. For a detailed discussion of the regression weights see sec. 3.4.

Three-body descriptors add angular information, i.e. energy contributions arising from the relative position of three atoms i, j, k. Equally to the 2-body descriptor, we can define a three-body kernel $k^{3b} : \mathbb{R}^2\mathrm{x}\mathbb{R}^2 \rightarrow \mathbb{R}$ defining the energy as

$$E_i^{(2b,3b)}(\mathbf{r}_{ijk}) = \sum_j^{N_t} w_j k^{(2b)}(r_{ij}, r_{ik}) + \sum_j^{N_t} w_{jk} k^{(3b)}(r_{ik}, r_{ij}, r_{jk}) \tag{3.11}$$

In the 3-body descriptor of GAP, distances r_{ij}, r_{ik}, and r_{jk} are not directly used, but a different form is used to enforce symmetry over permutation of the neighbor atoms j and k.

$$\mathbf{q}^{(3b)} = \begin{pmatrix} r_{ij} + r_{ik} \\ (r_{ij} - r_{ik})^2 \\ r_{jk} \end{pmatrix} \tag{3.12}$$

But as shown in Fig. 3.2, already a 3-body descriptor can be defined in multiple ways. Either by a set of three atomic distances (or a combination of these) or by an angle between three atoms. While it can be crucial to include higher body order terms to achieve a higher accuracy, when defining a many-body representation, the possibilities to define the descriptor space seems nearly infinite. [97] In the next section I will introduce the widely applied Smooth Overlap of Atomic Positions (SOAP) descriptor as my many-body descriptor of choice.

3.3.3 Smooth Overlap of Atomic Positions Descriptor

In the past, SOAP has been one of the most widely applied many-body representations.[39] In SOAP, neighboring atoms are represented by overlapping Gaussian functions yielding the neighbour density. While the neighbor density is by construction already invariant to permutation and translation, rotational invariance still has to be introduced. This is achieved by expanding the

neighbor density in the basis of orthogonal radial distribution functions and spherical harmonics. We again start by the energy expression

$$\epsilon_i^{(d)}(\chi_i^{(d)}) = \sum_{t=1}^{N_t} w_t^{(d)} K^{(d)}(\chi_i^{(d)}, \chi_t^{(d)}) \tag{3.13}$$

using a polynomial kernel

$$\epsilon_i^{(d)}(\chi_i^{(d)}) = \sum_{t=1}^{N_t} w_t^{(d)} \Big(\frac{\chi^{(SOAP)} \cdot \chi_t^{(SOAP)}}{\sqrt{\chi^{(SOAP)} \cdot \chi_t^{(SOAP)} \chi^{(SOAP)} \cdot \chi_t^{(SOAP)}}} \Big)^{\zeta} \tag{3.14}$$

In a first step the atomic structure is transformed into atomic density fields ρ for each species α. We can define a set of element-specific neighbor densities $\rho_{i,\alpha}$ for each central atom i

$$\rho_{i,\alpha}(\mathbf{r}) = \sum_j f_{cut} \cdot (\mathbf{r}_{ij}) exp\Big(- \frac{|\mathbf{r} - \mathbf{r}_{ij}|^2}{2\sigma_\alpha^2} \Big) \tag{3.15}$$

with σ_α being an element-specific descriptor width and f_{cut} a cutoff function. As discussed the neighbour density in eq. 3.15 is already invariant to permutations between equivalent atoms and translation. Rotational invariance can then be introduced by expanding the neighbor density in a set of orthonormal radial basis functions g_n and spherical harmonics Y_{lm}

$$\rho_{i,\alpha}(\mathbf{r}) = \sum_{nlm} w_{nlm}^{i,\alpha} g_n(r) Y_{lm}(\hat{\mathbf{r}}) \tag{3.16}$$

In this definition of the neighbour density $\rho_{i,\alpha}$ $\mathbf{r}$ is the vector containing the cartesian coordinates of atom i, r in the radial basis functions $g_n(r)$ is the magnitude and $\hat{\mathbf{r}}$) in the spherical harmonics $Y_{lm}(\hat{\mathbf{r}})$) is the direction. The coefficients $w_{nlm}^{i,\alpha}$ can be obtained by projecting the density onto the basis functions via

$$w_{nlm}^{i,\alpha} = \iiint dV g_n(r) Y_{lm}(\theta, \phi) \rho_{i,\alpha}(\mathbf{r}). \tag{3.17}$$

One typically does not use the entire powerspectrum, but the elements of a finite truncation of the power spectrum (up to $n \leq n_{max}$ and $l \leq l_{max}$). Hence, these hyperparameters n_{max}, l_{max} have to be chosen according to the investigated system. The now rotationally invariant output is the partial power spectra vector $\mathbf{p}$ with the individual elements:

$$p_{nn'l}^{i\alpha\alpha'} = \sqrt{\frac{8\pi^2}{2l+1}} \sum_m w_{nlm}^{i,\alpha} \cdot w_{n'lm}^{i,\alpha'} \tag{3.18}$$

In an alternative definition of SOAP by Ceriotti and coworkers[97], the spherical harmonics can be defined by angular terms θ and ϕ as

$$\rho_{i,\alpha}(\mathbf{r}) = \sum_{nlm} w_{nlm}^{i,\alpha} g_n(r) Y_{lm}(\theta, \phi) \tag{3.19}$$

with the coefficients $w_{nlm}^{i,\alpha}$

$$w_{nlm}^{i,\alpha} = \iiint dV g_n(r) Y_{lm}(\theta,\phi)\rho_{i,\alpha}(\mathbf{r}). \tag{3.20}$$

and the elements of the powerspectrum

$$p_{nn'l}^{i\alpha\alpha'} = \sqrt{\frac{1}{2l+1}} \sum_m w_{nlm}^{i,\alpha} \cdot w_{n'lm}^{i,\alpha'} \tag{3.21}$$

The kernel function for SOAP is a dot product of the power spectrum elements. When then using a polynomial kernel (raising the dot product to a small integer ζ) we get the final polynomial kernel K as defined above

$$K(\chi_i, \chi_t) = \left(\frac{\chi_i \cdot \chi_t}{\sqrt{\chi_i \cdot \chi_t \chi_i \cdot \chi_t}}\right)^{\zeta} \tag{3.22}$$

In addition to the derivation of SOAP, Fig. 3.3 gives a graphical depiction of the above SOAP derivation. This figure focuses on the symmetry (permutation, translation, rotation) that is introduced in the individual steps.

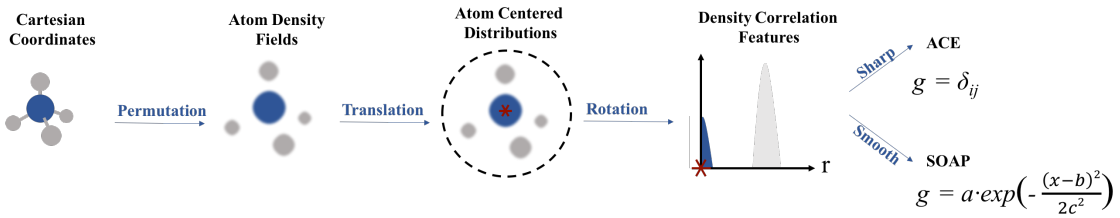


Fig. 3.3: *Schematic summary of the steps in a symmetrized field construction. In contrast to cartesian coordinates, the atom density field is permutationally invariant. By summing over the continuous translation group we yield an atom centered distribution. Using orthonormal radial functions and spherical harmonics yields a discrete set of coefficients that transform as spherical harmonics. The atomic density functions can be either finite-width Gaussians, which leads to representations to SOAP features, or Dirac δ distributions, which recovers the third body order term of the atomic cluster expansion.[93]*

The energy expression can now be expressed in terms of kernel functions

$$\begin{aligned} E_{total} = E_{local} = (\delta^{(2b)})^2 \sum_i^{N_{full}} \sum_t^{N_t} K^{(2b)}(\chi_i^{(2b)}, \chi_t^{(2b)}) \\ + (\delta^{(3b)})^2 \sum_j^{N_{full}} \sum_t^{N_t} K^{(3b)}(\chi_j^{(3b)}, \chi_t^{(3b)}) \\ + (\delta^{(SOAP)})^2 \sum_a^{N_{full}} \sum_t^{N_t} K^{(SOAP)}(\chi_a^{(SOAP)}, \chi_t^{(SOAP)}) \end{aligned} \tag{3.23}$$

The coefficients w are determined during the fitting process. In the next section we turn to gaussian process regression (GPR).

3.4 Regression Models

The third component determining the use of ML in computational chemistry in addition to the database and a set of suitable descriptors is the regression of atomic properties. In that sense regression defines the functional dependence of a given quantity on the local structural environment. Regression models define y_i as a function of x_i and c_i, with c_i being the respective regression weights. The penalty term λ, often referred to as the regularization, is introduced in some cases

$$y = f(x, c) + \lambda \quad . \tag{3.24}$$

The aim is to find the function $f(x, c)$ that most closely fits the given data. In order to carry out regression tasks, the form of function f must be specified. First, I will introduce a generalized formalism of the underlying GPR approach including a simple tutorial-style example. The last section will discuss the formalism of a sparse regression approach and the role of sparsification for the prediction of atomic properties.

3.4.1 Gaussian Process Regression

In a recent review on "Gaussian Process Regression for Materials and Molecules" Deringer and coworkers defined two equivalent approaches deriving the GPR framework.[100] Both approaches highlight different aspects of the fitting process. While from my point of view the weight-space approach highlights the similarity of Kernel-Ridge Regression (KRR) and GPR based on the choice of regression weights, the function-space view discusses the fact that the estimator of the local energies only depends explicitly on the kernel function, and not on the basis functions. In this section I will discuss the weight-view derivation of GPR but will highlight key information taken from the function-space derivation.

In the weight-space view of GPR, a function $y(x)$ can be approximated by a function $f(x, c)$, defined as a linear combination of N data points in the training set (usually atoms)

$$f(x, c) = \sum_{n}^{N} c_n k(x, x_n) = \mathbf{c}^{\mathrm{T}} \mathbf{K}_{NN}, \tag{3.25}$$

with c being the regression weights, k the kernel function, and x_n the input data. I use the matrix notation for that regression problem $\mathbf{c}^{\mathrm{T}}\mathbf{K}_{\mathrm{NN}}$ in order to be consistent with Fig. 3.4. The fitting of the GPR model to the data is done by finding the coefficients c that minimize the loss function

$$\mathcal{L} = \sum_{i} (y_i - f(x_i, c_i))^2 + \lambda^2 \sum_{n,n'}^{N} c_n k(x_n, x_{n'}) c_{n'} = ||(\mathbf{y}_{\mathrm{ref}} - \mathbf{y})||^2 + \Sigma \mathbf{c}^T \mathbf{K}_{NN} \mathbf{c}. \tag{3.26}$$

Σ is a diagonal matrix of size N containing all values of λ. By setting $\frac{d\mathcal{L}}{dc} = 0$ to minimize $\mathcal{L}$ and solve for $\mathbf{c}$, to obtain:

$$\mathbf{c} = \left(\mathbf{K} - \Sigma\right)^{-1} \mathbf{y}_{\mathrm{ref}}. \tag{3.27}$$

In the GPR framework the Gaussian kernel, sometimes also called square exponential kernel,

$$k(x, x_n) = exp\left(\frac{(x - x_n)^2}{2\sigma^2_{\text{length}}}\right) \tag{3.28}$$

with σ_{length} being the spatial length scale/width of the Gaussian is used.

Let's now evaluate an example function: y: $\mathbb{R} \rightarrow \mathbb{R}$. These data points are reproduced from Ref. [101]. The exact function form of $y(x)$ is unknown so we collect a total of N observations $\mathbf{y}_{\text{ref}}$. In Fig. 3.4 the increasing similarity of the predicted function $f(x, c)$ and the true function $y(x)$ with an increasing number of data points is depicted. While this in principle shows the relevance of additional data for the accuracy of the prediction, additional data come with an increasing computation cost. In the last section of this chapter I discuss this and address the role of sparsification for such problems.

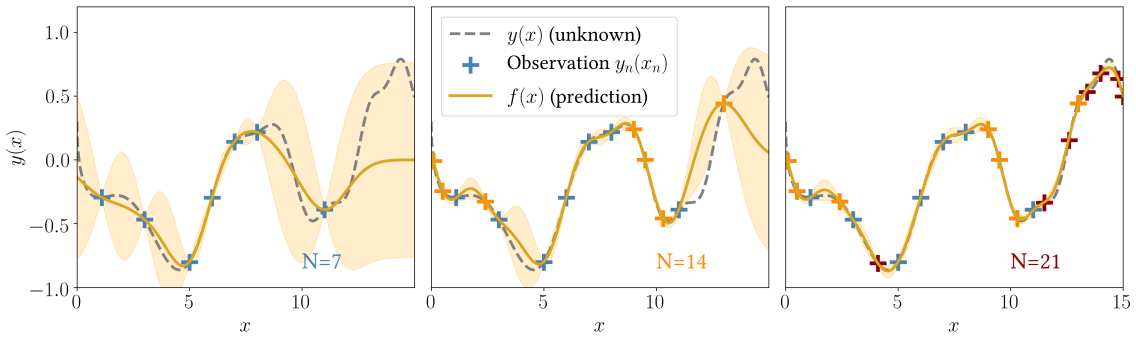


Fig. 3.4: *GPR prediction of a function $f(x, c)$ (solid yellow line) of an unknown one-dimensional function $y(x)$ (dashed black line). The orange shaded area is the standard deviation. GPR provides an uncertainty information that is not available with a kernel ridge. Using an increasing number of observations N of data points* $\mathbf{y}$ *the model hyperparameters are fixed.*

When updating the Gaussian process (hence adding a new observation n based on its predicted uncertainty), the updated Gaussian process is constrained to the possible functions that fit the observations N. Hence, the mean of $f(x, c)$ intercepts all observations N. Additionally, it is clear that the standard deviation is higher away from the observations, which reflects our lack of knowledge about these part of the function.

The kernel width σ_{length} in Gaussian kernel functions together with the regularization λ are both crucial for the smoothness and the accuracy of the predicted function. The role of σ_{length} and λ is depicted in Fig. 3.5. The choice is related to the often used terms 'underfitting' and 'overfitting' being the left and right panel in Fig. 3.5.

When decreasing regularization and/or kernel width ($\lambda \rightarrow 0$, $\sigma \rightarrow 0$) we obtain an improved alignment of training data N and the predicted function f. As illustrated in the right panel of Fig. 3.5 an increased precision on the prediction of the training data N when decreasing λ and or σ diminishes the predictive accuracy for input data unequal to the training data. It is therefore crucial to determine the accuracy/ predictive power on a set of independent data points.

3.4.2 Sparse Gaussian Process Regression

In practice when training an interatomic potential, full GPR is not applicable for large training sets, because the computational costs of training scale with $\mathcal{O}(N^3)$ i.e. cubical with the training

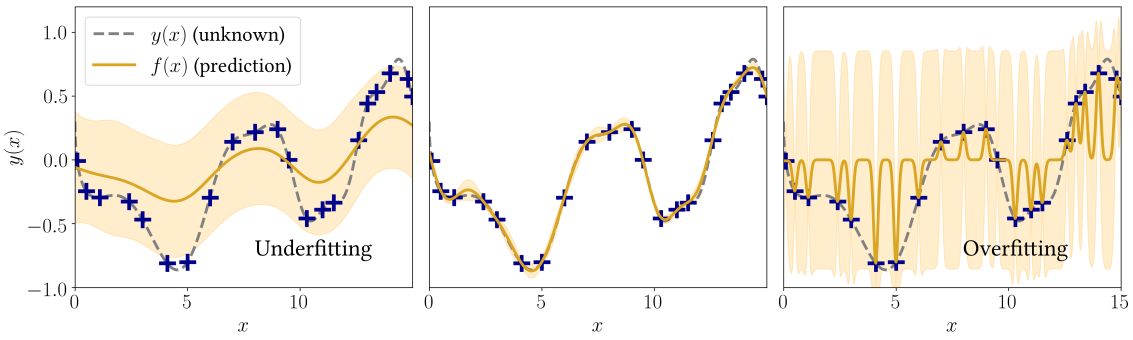


Fig. 3.5: *Effect of regularization λ and kernel width σ on the GPR prediction (solid yellow line). Left panel: For small penalties on the prediction (large λ, σ) $f(x, c)$ is centered around the average of the observations. Although the predicted function is smooth, it can not accurately predict the original function and has a high uncertainty. Right panel: In contrast, small λ, σ overfit the function y(x). Although high accuracy is achieved for the training data, high uncertainties are yielded for all other data points. Center panel: An optimal performance is acchieved when balancing smoothness and accuracy with appropriate λ and σ.*

set size.[100] In GAP, instead of using all data points N a smaller set of representative points M(i.e. $M \ll N$) defining the 'sparse' representation is used. We first recall the loss function

$$\mathcal{L} = \sum_{n}^{N} (y_n - f(x_n, c_n))^2 + R \tag{3.29}$$

where the relative importance of individual data points being controlled by R. Opposite to the general case discussed above in sparse GPR the regularization term R is now only depending on this representative set of M instead of N data points

$$R = \Sigma \sum_{m,m'}^{M} c_m k(x_m, x_{m'}) c_{m'}. \tag{3.30}$$

Adding eq. 3.29 to eq. 3.30 and rewriting the loss function in matrix form yields:

$$\mathcal{L} = \left(\mathbf{y}_{\text{ref}} - \mathbf{K}_{NM}\mathbf{c}\right)^T \Sigma^{-1} \left(\mathbf{y}_{\text{ref}} - \mathbf{K}_{NM}\mathbf{c}\right) + \mathbf{c}^T \mathbf{K}_{MM}\mathbf{c}. \tag{3.31}$$

The matrix elements are defined as $\mathbf{K}_{NM} = k(\mathrm{x}_n, \mathrm{x}_m)$ where N indicates the number of data points in the data set and M indicates the number of representative points, respectively. Minimizing $\mathcal{L}$ we obtain:

$$-\mathbf{K}_{MN}\Sigma^{-1}\mathbf{y} + \mathbf{K}_{MN}\Sigma^{-1}\mathbf{K}_{NM}\mathbf{c} + \mathbf{K}_{MM}\mathbf{c} = 0 \quad . \tag{3.32}$$

When solving for $\mathbf{c}$ we yield the following expression:

$$\mathbf{c} = \left(\mathbf{K}_{MM} + \mathbf{K}_{MN}\Sigma^{-1}\mathbf{K}_{NM}\right)^{-1} \mathbf{K}_{MN}\Sigma^{-1}\mathbf{y}_{\text{ref}} \tag{3.33}$$

Although, in the first glance eq. 3.27 appears to be 'simpler' compared to eq. 3.33, Fig. 3.7 clarifies that the coefficient vector is shorter. Fig 3.7 also sorts out the misconception of sparsification,

that data points are left out. In sparse GPR, the full data vector $\mathbf{y}$ is used in training, yet now M ('sparse') locations are chosen to evaluate (unkown) input data. The coefficient vector is therefore of length M in the case of sparse GPR, while full GPR yields a coefficent vector of lenght N. Evaluating $f(x_{\text{new}}, c)=\tilde{y}$ for a new configuration $\mathbf{x}_{\text{new}}$ is done using eq. 3.25

$$f(x_{\text{new}}, \mathbf{c}) = \mathbf{c}^T \mathbf{k}(x_{\text{new}}) \tag{3.34}$$

A key benefit of sparse GPR becomes apparent. While the cost of prediction in full GPR scales with N, sparse GP is now independent of N. The notation of $\mathbf{k}(\mathbf{x}_{\text{new}})$ is used for the vector of kernel values at $\mathbf{x}_{\text{new}}$ and the set of representative points

$$\mathbf{k}(\mathbf{x})_m = k(x, x_m) \tag{3.35}$$

But what does that now mean in the case of the GAP framework? All GAP models are sparse kernel models, i.e. the basis functions for the linear expansion of the atomic energy do not directly correspond to the set of input data N. In the GAP framework individual atomic environments are chosen as the elements of the representative set M, and the corresponding kernel basis functions are used to expand the atomic energy. The crucial factor, beside the actual number of sparse points (which can be seen as convergence parameter), is the choice of representative environments for training. While for a two body-descriptor a homogeneously spaced grid is suitable due to low dimensionality, the CUR algorithm has been found to be a good algorithm to provide a decent set of representative data points within the SOAP-GAP model.[100]

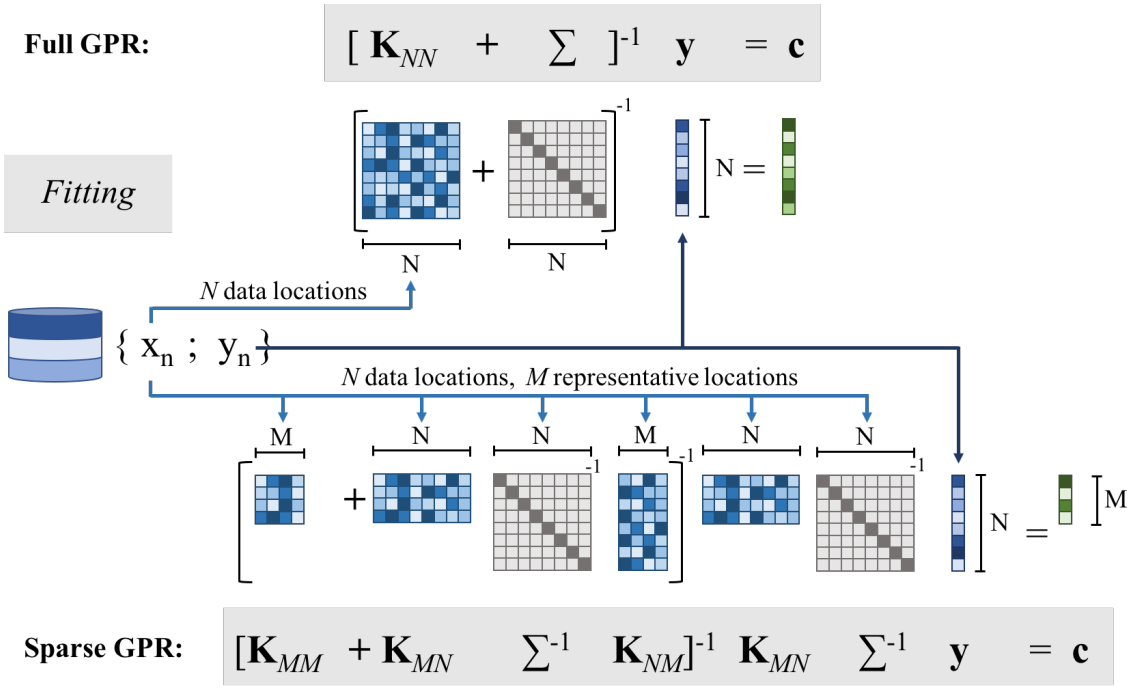


Fig. 3.6: *Visualization of the full GPR (top) eq. 3.27 and sparse GPR (bottom) eq. 3.33. The training database consists of N entries. In full GPR all N entries are used to obain the coefficient vector* c *of length N. In sparse GPR, still all entries y are used, but M representative ("sparse") locations are chosen to obtain a coefficient vector c of length M. The figure is adapted from a review by Deringer and coworkers.[100]*

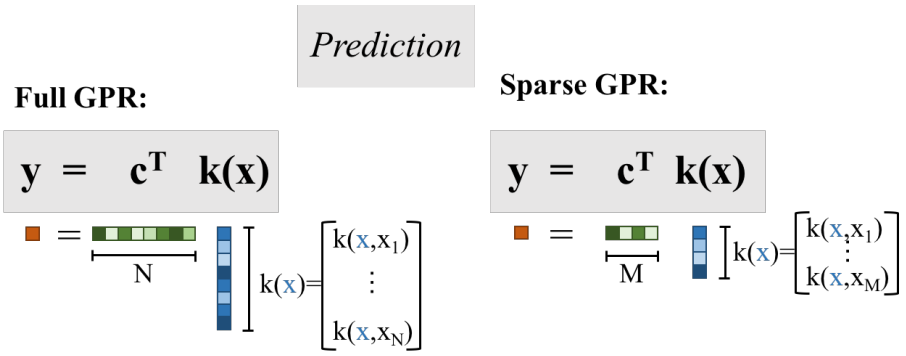


Fig. 3.7: *The coefficient vector* c *(green) is computed, and can be used to make a prediction at a new location $f(x_{new})$ as given in eq. 3.34. The cost scales with the number of data locations, N. In sparse GPR, the full data vector* y *is used as well, but now M representative ('sparse') locations are chosen, with $M \ll N$. The coefficient vector is therefore of length M, and the cost of prediction is now independent of N. The figure is adapted from a review by Deringer and coworkers.[100]*

4 Long-Range Electrostatics and Non-local Charge Transfer

The locality assumption kick-started the fitting of ML potentials and their wide applicability spanning the range from molelucar to condensed systems.[37, 89] Even in water, where one would expect long-range electrostatics to be crucial for an accurate description, extraordinary accuracy of the structural and dynamic properties of bulk liquid water and different ice phases was achieved using a short-range potential.[102] This can be rationalized by the fact that these properties/materials are presumably highly isotropic, so that long-range interactions average out. The importance of long-range effects depends on the material and property of interest and thus demands a systematic analysis.

In this last chapter I will give a historical outline on long-range interactions. Starting with Coulomb's law from 18th century, to Sanderson's concept of electronegativity equalization introduced in the early 50s, and the first Charge Equilibration (QEq) scheme by Rappe and Goddard in 1991, section 4.1 will introduce the basic formalism of charge equilibration.

Section 4.2 covers machine-learning based charge equilibration approaches and recent developments in the field of machine learning based charge assignments.

4.1 First steps towards QEq

Coulomb's law, when first introduced in 1785 enabled the quantification of the force between two stationary, electrically charged particles. Charged particles' attraction or repulsion is directly proportional to the product of charges (q_i, q_j) and inversely proportional to the square of the distance between them:

$$F = k^c \frac{q_i q_j}{(r_i - r_j)^2} \quad . \tag{4.1}$$

Here, k^c is Coulomb's constant.

From Coulomb's law we can derive the electrostatic potential energy $(U_E(q_i))$ of one point charge q_i at position r_i in the presence of another point charge q_j as:

$$U_E(q_i) = k^c \frac{q_i q_j}{|r_i - r_j|} \quad , \tag{4.2}$$

as well as the electrostatic potential energy $(U_E(q_i))$ of one point charge q_i at position r_i in the presence of $N-1$ other point charges q_j

$$U_E(q_i) = k^c q_i \sum_{j=1}^{N-1} \frac{q_j}{|r_i - r_j|} \quad , \tag{4.3}$$

In a molecule or condensed system we can then define the electrostatic potential energy of the entire system of N point charges as:

$$U_E = \frac{1}{2} k^c \sum_i^N q_i \sum_{j \neq i}^N \frac{q_j}{r_{ij}} \quad . \tag{4.4}$$

For small atomic distances r_{ij} e.g. in chemical bonds it is intuitive that the energy description can be beyond coulombic interactions. Linus Pauling defined electronegativity χ by the ability of an atom to attract shared electrons/electron density when forming a chemical bond.[103] He proposed an electronegativity scale which depends on bond energies, as a development of valence bond theory. By that we are able to understand that a covalent bond between two different atoms (i–j) is stronger than the average of the i–i and the j–j bonds. He proposed an electronegativity scale which depends on bond energies, as a development of valence bond theory. It is important to remember that electronegativity cannot be directly measured and strongly correlates with a number of other chemical properties. Two of these properties are the ionization potential (IP) and the electron affinity (EA).[104] The ionization potential is the energy needed to remove the outer valence electron. Electron affinity is the energy difference related to the injection of an extra electron. As such we can define the electronegativity of atom i as:

$$\frac{IP + EA}{2} = \chi_i^E \tag{4.5}$$

When two atoms i and j form a bond Sanderson postulated that the corresponding electronegativities equalize.[105] He first showed that concept based on bond lengths in alkali halide gas molecules, and proved that the same holds true for more than two atoms combining within a molecule. He defined the electronegativity as a stability ratio of the formed bond of atoms i and j

$$\chi^E = \frac{Z}{4.19 r_A^3 \rho_e} \tag{4.6}$$

where Z is the atomic number, r_A the atomic radius of each atom and ρ_e the electronic density of atom i. When comparing reported bond lengths to ionic or non-polar covalent radii, Sanderson identified electronegativities based on bond stability for different molecules. He then averaged χ^E to yield one electronegativity per element. But in principle he already put forth that the electronegativity depends on the formed bond and hence the atomic environment.

Rappe and Goddard extended the idea of the electronegativity equilibration of a bond and developed a framework to estimate the electrostatic potential energy of molecules and periodic systems by its charges. The QEq methods computes partial charges by using atomic coordinates and the two previous defined properties of isolated atoms $\chi_i^{E,0}$, J_{ii}^0. I use the notation $\chi_i^{E,0}$ instead of χ_i when we treat an isolated atom i instead of atom i in a molecule. This is one of the underlying assumptions: atomic ionization potential and electron affinity of isolated atoms are similar to the ones of the same element bonded inside a molecule or a crystal. The concepts involved in the QEq approach manages to combine earlier ideas of Pauling (chemical bond) [103], Mulliken (electron affinity)[104], Margrave (an extended definition of electronegativity)[106], Parr and Pearson (concept of hardness)[107], Mortier (Electronegativity-equalization for the prediction of partial charges)[108], and others. In that respect, I can highly recommend the original QEq

publication of Rappe and Goddard. In their publication they clearly outline the relationship between QEq and some of these earlier ideas and methods.

Above, I already defined electronegativity based on IP and EA. Now let's consider how the energy of an isolated atom changes as a function of charge, e.g. how the energy changes by adding or removing an electron. Starting with the neutral reference, one can write the energy of atom i by a Taylor expansion

$$E_i(q_i) = E_{i0} + q_i\left(\frac{\partial E}{\partial q}\right)_{i0} + \frac{1}{2}q_i^2\left(\frac{\partial^2 E}{\partial q^2}\right)_{i0} + \dots \tag{4.7}$$

When including terms through second order and defining two examples of adding/removing an electron we yield:

$$E_i(+1) = IP = E_{i0} + \left(\frac{\partial E}{\partial q}\right)_{i0} + \frac{1}{2}\left(\frac{\partial^2 E}{\partial q^2}\right)_{i0} \tag{4.8}$$

$$E_i(-1) = EA = E_{i0} - \left(\frac{\partial E}{\partial q}\right)_{i0} + \frac{1}{2}\left(\frac{\partial^2 E}{\partial q^2}\right)_{i0} \tag{4.9}$$

These two cases represent the IP (the energy needed to remove the outer valence electron) and the EA (the energy difference related to the injection of an extra electron). When solving for $\left(\frac{\partial E}{\partial q}\right)_{i0}$ and $\left(\frac{\partial^2 E}{\partial q^2}\right)_{i0}$ we can define

$$\left(\frac{\partial E}{\partial q}\right)_{i0} = \frac{E_i(+1) + E_i(-1)}{2} = \frac{IP + EA}{2} \tag{4.10}$$

$$\left(\frac{\partial^2 E}{\partial q^2}\right)_{i0} = E_i(+1) - E_i(-1) = IP - EA \tag{4.11}$$

In eq. 4.5 we already identified the first term $\left(\frac{\partial E}{\partial q}\right)_i = \chi_i^{E,0}$ being equivalent to the electronegativity. For the second term let's consider a neutral atom with a singly occupied orbital. The orbital is empty for the positive ion and double occupied for the negative ion. The difference between the IP and EA for that orbital is:

$$IP - EA = J_{ii}^0 \tag{4.12}$$

with J_{ii}^0 being the Coulomb repulsion between two electrons in the orbital. This electronic repulsion quantity is the idempotential and is known as atomic hardness. This is an approximation as the optimum shape of the orbital changes upon adding an additional electron.
Using the definitions of electronegativity and hardness we get

$$E_i(q_i) = E_{i0} + \chi_i^0 q_i + \frac{1}{2}J_{ii}^0 q_i^2 \tag{4.13}$$

In order to yield an energy expression for molecules and crystals the sum of atomic values for $\chi_i^{E,0}$ and J_{ii}^0 are extended by pairwise interactions between the atoms:

$$E_{tot}(q_1, \ldots, q_N) = \sum_i^N \left(E_{i0} + \chi_i^0 q_i + \frac{1}{2} J_{ii}^0 q_i^2 \right) + \sum_{i<j}^N q_i q_j J_{ii} \tag{4.14}$$

with J_{ij} being the Coulomb interaction between atom i and j. This can be rewritten as

$$E_{tot}(q_1, \ldots, q_N) = \sum_i^N \left(E_{i0} + \chi_i^{E,0} q_i \right) + \frac{1}{2} \sum_{i,j}^N q_i q_j J_{ij} \tag{4.15}$$

In order to obtain the partial charges we set up a system of partial differential equations of the energy with respect to the system charges as χ_i

$$\chi_i(q_1, \ldots, q_N) = \frac{\partial E}{\partial q_i} = \chi_i^{E,0} + \sum_B^N q_B J_{ij} \tag{4.16}$$

χ_i is a function of the charges on all the atoms of the system. With the constraint on the total charge

$$Q_{tot} = \sum_{i=1}^N q_i \tag{4.17}$$

the minimun energy is found if

$$\chi_1 = \chi_2 = \cdots = \chi_N \tag{4.18}$$

By that we have a set of N partial differential equations for the equilibrium self-consistent charges that are solved once for a given structure.

In order to solve the QEq scheme, we require the Coulomb potential J_{ij}, defined as the potential between the charge centers on atoms i and j, to be separated by a distance R. For distances R where the charge distributions of atoms overlap, the simple Coulomb law (eq. 4.3) is not valid. For $R \rightarrow 0$, the Coulomb interaction $J_{ij} \rightarrow \infty$, whereas it should lead to a finite value related to J_{ii} and J_{jj}. To ensure that $J_{ij}(r)$ is physical meaningful, a shielding is needed. A variety of shielding approaches exist, Rappe and Goddard choose the shielding to be the Coulomb integral between atomic densities. For simplicity Slater-type densities of neighboring atoms are considered instead of atomic densities from *ab initio* calculations on atoms. Hence, QEq still considers neighbouring atoms like isolated atoms that are pushed close.

The electron densities of an atom with valence orbitals *n*s, *n*p, or *n*d can be written as normalized single *nS* Slater densities of the form

$$\rho_{n\zeta}^{Slater} = N_n r^{n-1} e^{-\zeta r} \tag{4.19}$$

where Nn is the normalization constant, n the valence shell and ζ the valence orbital exponent. The valence orbital simply represents the characteristic size of each atom by

$$\zeta_i = \frac{\lambda(2n+1)}{2r_i} \tag{4.20}$$

The scaling factor λ accounts for the difference between an average atom size and the covalent radius r_i. The Coulomb integral for short distances can now be expressed by atomic densities

$$J_{ij}(R) = \iint \rho_i(r_i)\frac{1}{r_{ij}}\rho_j(r_j)dV_i dV_j \tag{4.21}$$

In later work Rappe and coworker's extended the QEq for periodic systems by using the Ewald summation. By using the Ewald summation they ensure the convergence of the Coulomb term in an infinite periodic system. Since the first implementation of the original QEq scheme, a variety of extensions have been proposed in order to improve the quality of the computed charges. Ongari *et al.* compared systematic errors of different classical approaches for gas adsorption predictions in metalorganic frameworks (MOFs). [109] These variations can in general be distinguished by these four parameters:

- Choice of the atomic parameters
- Center and the order of the Taylor expansion of the energy
- Analytic form to compute the pairwise interaction between atoms with respect to it's geometry
- Inclusion of further parameters to characterize each bond type

In the next section, I want to discuss a few selected machine learning approaches for charge prediction and charge equilibration.

4.2 Machine Learning Charge Prediction Schemes

Including long-range electrostatic interactions in ML potentials as an electrostatic baseline is a great challenge. We not only require to take interactions beyond the cutoff radius into account, but also to include physically meaningful energy terms. Huge efforts have been made by various groups in the past, ranging from NN approaches like PhysNet[31] and HIPNN[110] predicting partial charges, electrostatic multipole coefficients for organic molecules from kernel-ridge regression [111, 112], partial charge prediction by random forest regression in drug like molecules [113], and many more.[114–116]

In this section I want to focus on two ideas that - from my perspective - are the key developments in the last decade that influenced the development of our kQEq model. Namely Behler's third generation NN in 2011[33] and Goedecker's charge equilibration neural network technique (CENT) in 2015[117]. Behler and coworkers were the first group to use a baseline neural network predicting environment-dependent charges from ab initio atomic charges. By that they could predict long-range interactions but lack the ability to predict non-local charge transfer. A first important step towards non local charge-transfer in ML potentials has then been done by Goedecker and co-workers. By predicting environment dependent electronegativities and determine charges by the QEq scheme they were able to include long-range charge transfer in a qualitatively correct way. In all subsections, I focus on the main idea of how to incorporate long-range electrostatics. Details of the underling NN frameworks can be found in the corresponding publications.

4.2.1 Third generation NN potentials

The key component for a successful application of such an electrostatic baseline is the accurate prediction of atomic charges. While for some cases, including our first attempt to use an electrostatic baseline[73], an element-wise fitting of charges can be sufficient, the work of Behler and others have shown the possibility to predict environment-dependent atomic charges.[33, 118]

In this approach a baseline NN (accounting for the long-range electrostatic) in addition to a NN potential (covering all remaining short-range interactions within the cutoff) is trained. This idea of using an electrostatic baseline is often referred to Δ-learning.

$$E_{total} = E_{short} + E_{elec} \tag{4.22}$$

In Δ-learning a double counting of electrostatic energy contributions is avoided by a simple subtraction of the electrostatic energy E_{elec} from the total energy E_{total}. As shown in Fig. 4.1, atomic charge NNs are trained using reference atomic charges obtained from electronic structure calculations and atomic positions.

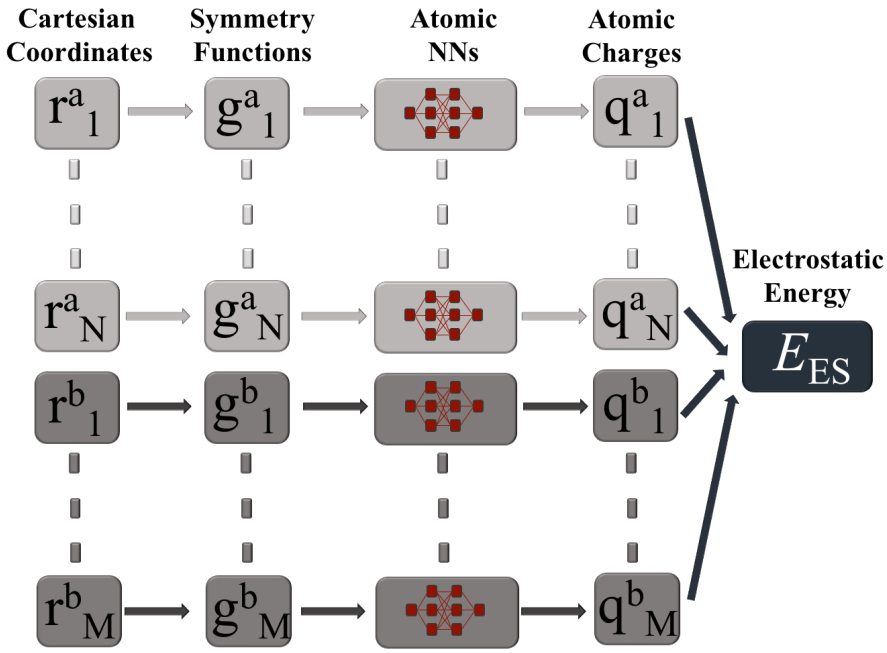


Fig. 4.1: *Schematic structure of a third generation NN potential by Behler. A set of atomic NNs (shown in red) is used to construct environment-dependent atomic charges. These predicted atomic charges can be used to calculate the long-range electrostatic energy. The total energy of the system is then given by the sum of the short-range E_{short} and the electrostatic energy E_{elec}.The figure is adapted from a review by Behler.[119]*

Using first-principle partial charges is a major drawback in that approach. Atomic partical charges are not physical observables and there is no unique/best choice. Although all of them are mathematically well-defined, benchmark studies have shown that different partitioning schemes yield very different results.[120] In addition, predicting partial charges directly from atomic environments, will not cover non-local charge transfer, e.g. charge redistribution outside the

cutoff radius. As charges are predicted for each atomic environment individually without a constrain on Q_{tot}, the total charge is not conserved and molecules, which are formally charge neutral might be predicted to be charged by these type of NN.

4.2.2 The CENT approach

To account for non-local charge transfer Goedecker and co-workers define the total-energy expression similare to QEq by:

$$E_{tot}(q_1, \ldots, q_N) = \sum_i^N \left(E_i^0 + \chi_i^{CENT} q_i + \frac{1}{2} J_{ii} q_i^2 \right) + \frac{1}{2} \iint \rho_i(r_i) \frac{1}{r_{ij}} \rho_B(r_i) dV_i dV_j \tag{4.23}$$

with E_i^0 being a reference energy, q_i the atomic charges, J_{ii} the atomic hardness, and χ_i^{CENT} the environment-dependent atomic electronegativity of atom i. The charge density ρ in CENT is a superposition of normalized spherically symmetric Gaussian functions of width α_i centered at atomic positions r_i given by

$$\rho_i(\mathrm{r}) = \frac{q_i}{\alpha_i^3 \pi^{\frac{3}{2}}} exp\Big(- \frac{|\mathrm{r} - \mathrm{r}_i|^2}{\alpha_i^2} \Big) \tag{4.24}$$

The QEq approach as outlined above only requires three parameters, namely the electronegativity (χ_i^E), the non-classical contribution to the hardness (J_{ii}^0) and the atomic size (r_i) for each species in the system. As a flipside of this elegant simplicity, the accuracy and transferability of the QEq method is limited, however.

In the CENT approach this limitation is hurdled by allowing the electronegativity χ^{CENT} of an atom to change as a function of its chemical environment.[117] As shown in Fig. 4.2, for an ionic system, the cartesian coordinates (input) are transformed to atom-centered symmetry function vectors (descriptors for the NN). These inputs for the atomic NNs yield the environment-dependent electronegativities χ^{CENT}. Using the charge equilibarion framework similar to QEq, atomic charges q can be used to compute the total energy E_{tot} using eq. 4.23.

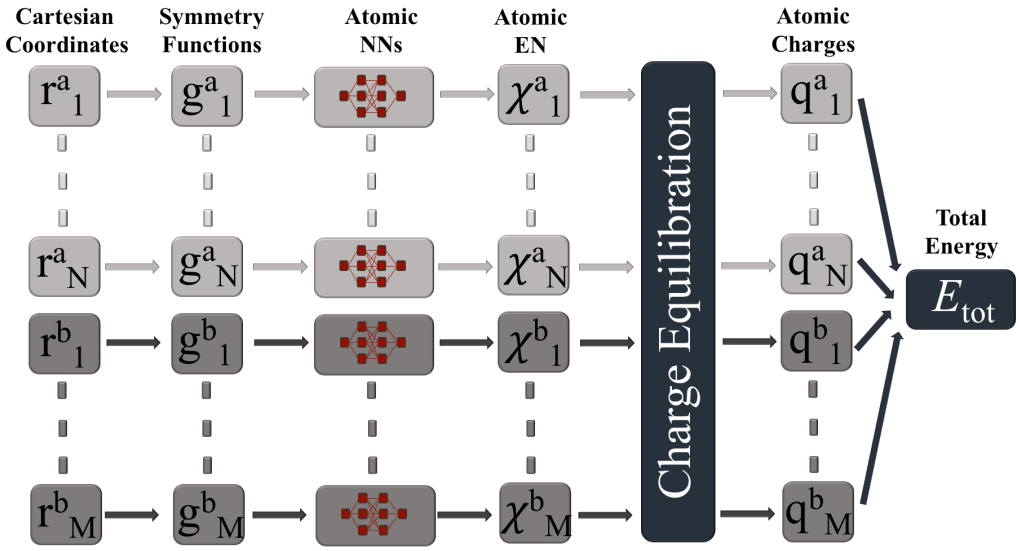


Fig. 4.2: *Schematic structure of the charge equilibration neural network technique (CENT) for a system with two elements (a,b) with N atoms of element a and M elements for element b. Cartesian coordinates are transformed to atom-centered symmetry function vectors. These basis functions are the input for atomic NNs yielding environment-dependent electronegativities χ. Using a charge equilibration framework similar to QEq, atomic charges q can be used to compute the total energy E_{tot}. The figure is adapted from a review by Behler.[119]*

Goedecker, and co-workers applied the NN-based QEq model to ionic crystals.[117] In these ionic systems, the total energy can entirely be represented by the electrostatic contributions. In general there is no need for Q_{tot} to be zero such that the method is also applicable to charged systems. Notably in the CENT approach, Q_{tot} is conserved. For details on the setup on the NN infrastructure see [117].

5 Publications

As this thesis is publication based, in this chapter a summary of my two publications that resulted from my research during my PhD period is given. Each overview includes a brief summary of the genesis and its content and is followed by a more detailed elaboration on my personal contribution. The corresponding full articles together with the respective supporting information can be found in the appendix of this thesis.

5.1 On the role of long-range electrostatics in machine-learned interatomic potentials for complex battery materials

Carsten G. Staacke, Hendrik H. Heenen, Christoph Scheurer, Gábor Csányi, Karsten Reuter, Johannes T. Margraf
ACS Appl. Energy Mater. 2021, 4, 12562-12569
DOI: 10.1021/acsaem.1c02363
Summary: This first project aimed for the question, are long-range interactions needed when using an ML interatomic potential and if so when are they relevant? Modeling complex energy materials such as the herein investigated $Li_7P_3S_{11}$ SSE realistically at the atomistic level strains the capabilities of state-of-the-art theoretical approaches. Fortunately, modern ML potentials promise first-principles accuracy at a much reduced computational cost. However, the local nature of these ML potentials typically means that long-range contributions arising, e.g., from electrostatic interactions are neglected. Clearly, such interactions can be large in polar materials like electrolytes. In this work we investigated the effect that the locality assumption of ML potentials has on lithium mobility and defect formation energies in SSEs. We therefore developed a Δ learning protocol using a simple electrostatic baseline (ES-GAP). Comparing the classical GAP model with the newly developed ES-GAP, we found that neglecting long-range electrostatics is unproblematic for the description of lithium transport in isotropic bulk like environments. In contrast, simulating non-isotropic systems yielded the importance of ES contributions and provided new insights into interphase stability of $Li_7P_3S_{11}$.

Specifically, we studied Frenkel defects in an applied field mimicking the potential drop at a solid/solid interface. In this setup we found that a stabilization of the defects can occur already at moderate fields. This would favor the accumulation of defects towards the interphase, which could influence the kinetic stability of $Li_7P_3S_{11}$/electrode interfaces. Additionally, such stabilizations are anisotropic to crystallographic orientation making grain shape and orientation an additional parameter to be considered in battery engineering and beyond.

The findings of this work laid the foundation for the following two projects. First, now that we understood when ES contributions are relevant, we need to develop a suitable electrostatic model, that has the favourable properties of an ML approach, keeps the reactivity of ML interatomic potentials and can describe non-local charge transfer. Second, if we want to model SSE interphases realistically, we need to develop a near-universal ML potential that can likewise describe crystalline and amorphous LPS compounds.

Individual Contributions

The idea for this project was jointly conceived by Johannes Margraf, Gábor Csányi, Karsten Reuter, Christoph Scheurer and myself. Gábor Csányi introduced me to the GAP model while Johannes Margraf and Christoph Scheurer helped me to understand electrostatic modeling beyond first principles. While we have been able to identify cases of similar predictive power of the GAP and ES-GAP model, Karsten Reuter and Hendrik Heenen suggested interfacial like simulation tasks to strain the capability of a local GAP model. The manuscript was jointly written and edited by all authors.

5.2 Kernel Charge Equilibration: Efficient and Accurate Prediction of Molecular Dipole Moments with a Machine-Learning Enhanced Electron Density Model

Carsten G. Staacke, Simon Wengert, Christian Kunkel, Gábor Csányi, Karsten Reuter, Johannes T. Margraf
Mach. Learn.: Sci. Technol. 3, 2022, 015032
DOI: 10.1088/2632-2153/ac568d

Summary

In the previous project we identified the need for a charge model that can describe non-local charge transfer. Hence, in this project we developed a kernel based extension of the widely used charge equilibrition model (QEq) termed kernel Charge Equilibration (kQEq). In contrast to conventional QEq, a data-driven, environment-dependent description of atomic electronegativities is introduced. For this work we trained kQEq models on molecular dipoles and have been able to show an excellent performance, *en par* with or better than state-of-the-art kernel models, specifically tuned to predicting dipole moments.

The kQEq formalism presented in this work opens the door towards physics-based kernel ML models for predicting atomic charges, to be used in combination with reactive interatomic potentials such as the ES-GAP model. Most importantly, the presented approach was designed quite general and can be extended to other fit targets (*e.g* quadrupole moments and electrostatic potentials) and to more flexible density representations (*e.g* using atom centered dipoles in addition to partial charges).

While this work served as a proof of concept for molecular systems, we envision the extension for more complex finite systems, such as catalytic processes on nano-particles, and periodic systems, such as SSE interfaces.

Individual Contributions

Inspired by the work of Goedecker et al., Johannes Margraf suggested to replace fixed electronegativities by an atomic-environment based term and we quickly developed the first working implementation together with Simon Wengert. Thanks to Christian Kunkel, an online documentation is available. While we first aimed for partial charges as the fitting target, Gábor Csányi suggested to aim for atomic dipoles as a more realistic target. The manuscript was jointly written and edited by all authors.

5.3 Tackling structural complexity in Li_2S-P_2S_5 solid-state electrolyte using Machine Learning Potentials

Carsten G. Staacke*, Tabea Huss*, Johannes T. Margraf, Karsten Reuter, and Christoph Scheurer
* These authors contributed equally to the work.
Nanomaterials, 12, 2950, (2022)
DOI: 10.3390/nano12172950

Summary

For the final project we aimed for a near universal ML potential for the LPS material class that can likewise describe crystalline and amorphous LPS. So, how can we model amorphous LPS realistically at the atomic level and is there a way for a data-efficient description for glassy SSEs? As the lithium thiophosphate (LPS) material class provides promising candidates for solid-state electrolytes (SSEs) in lithium ion batteries due to high lithium ion conductivities, non-critical elements, and low material cost we want to be able to investigate this materials class as a whole. LPS materials are characterized by complex thiophosphate microchemistry and structural disorder influencing the material performance. To overcome the length and time scale restrictions of *ab initio* calculations in simulations of industrially applicable LPS materials, we develop a data-efficient training approach for SSEs with an emphasis on complex microchemistries. Our trained GAP model can likewise describe crystal and glassy materials and di erent P-S connectivities P_mS_n.

As we have been able to model this material class as a whole, we apply the GAP surrogate model to probe lithium ion conductivity and the influence of thiophosphate subunits on the latter. In our work we found that the paddle wheel effect, and hence, a constant reorientation of S in PS_4^{3-}, is happening as long as PS_4^{3-} is present, indepent of glassy or crystalline environments.

The general structure of the training protocol furthermore allows for a variety of extensions, such as dopants, other selection criteria, and including an electrostatic baseline. For future work we are currently aiming for a combination of the kQEq model with our training approach to realistically model internal SSE interfaces.

Individual Contributions

This project has been an ongoing effort of Tabea Huss and myself, that included a research internship and Tabea Huss masterthesis. While I developed the general fitting procedure, Tabea Huss refined the iterative training scheme. Christoph Scheurer and Johannes Margraf guided us when we were stuck in one of the many approaches we tried and helped us to analyse the data we produced. The manuscript was jointly written and edited by all authors.

5.4 Additional Work on Machine Learning Potentials: Machine Learning Surface Complexions of Rutile IrO_2 and RuO_2

IrO_2 Surface Complexions Identified through Machine Learning and Surface Investigations

Jakob Timmermann, Florian Kraushofer, Nikolaus Resch, Peigang Li, Yu Wang, Zhiqiang Mao, Michele Riva, Yonghyuk Lee, Carsten Staacke, Michael Schmid, Christoph Scheurer, Gareth S. Parkinson, Ulrike Diebold and Karsten Reuter
Phys. Rev. Lett. 125, 206101 (2020).
DOI: 10.1103/PhysRevLett.125.206101

Summary:

During an initial *ab initio* study of IrO_2 nanoparticles and surfaces, it quickly became obvious that relevant insights considering the catalytic surface demand a method multiple orders of magnitude faster than DFT. As classical force fields in this case are insufficient due to their lack of reactivity we turned to the Gaussian Approximation Potential (GAP) framework as a reactive, interatomic Machine Learning (ML) potential. Fruitful discussion and testing of hyperparameters with Jakob Timmermann led to a first GAP model. Already the first simulated annealing (SA) simulations based on the initial GAP revealed several unknown GAP minimum structures. In further iterative training and back then visual inspection of Jakob Timmermann revealed a variety of new, low energy surface complexions that have been revealed by surface reconstructions during the SA simulations. Together with the colleages from Prof Ulrike Diebold's group at the Technical University Vienna we have been able to confirm experimentally and theoretically that solely (101)-type surfaces confirm a (1 x 1) surface unit cell size and hence ruling out any reconstruction of higher symmetry.

Data-Efficient Iterative Training of Gaussian Approximation Potentials: Application to Surface Structure Determination of Rutile IrO_2 and RuO_2

Jakob Timmermann, Yonghyuk Lee, Carsten Staacke, Johannes T. Margraf, Christoph Scheurer and Karsten Reuter
J. Chem. Phys. 155, 244107 (2021)
DOI: 10.1063/5.0071249

Summary:

The instant success of the generation-based training protocol had an obvious flaw: the selection process via visual inspection is highly subjective, and can not be automated. Accordingly, we introduced a similarity measure as a systematic, quantifiable selection criterion into our workflow and developed an updated iterative and automated training protocol for the identification of global minimum structures of arbitrary metal oxide surfaces. I then helped with the hyperparameter selection procedure for rutile IrO_2 and RuO_2. Jakob Timmermann and Yonghyuk Lee then took this initial work and developed a systematic yet data-efficient scheme for bootstrapping the initial training set, detailed heuristics on how to test and select appropriate hyperparameters, and a straight-forward approach to determine the similarity threshold. This then fully-automatized, iterative training protocol was successfully applied to IrO_2 and RuO_2 and revealed additional low-energy complexions for both materials.

6 Summary, Conclusions and Outlook

Large scale energy storage is already indispensable in our todays life. In that respect batteries play a crucial role in modern mobility, transport and communication. While we are struggling with the complete picture of liquid electrolytes, we still barely scratch the surface when it comes to solid-state batteries. In the past, *ab initio* methods and empirical potentials have been powerful tools. Unfortunately, in the same way *ab initio* methods are limited given the size and time scales required for a detailed understanding of interfaces, while empirical potentials don't reach the required accuracy or lack flexibility. Here, emerging Machine Learning (ML) methods have shown to be increasingly able to combine the strength of both approaches, with good first-principles accuracy at a much reduced computational cost.

In this thesis, I developed a variety of new approaches and tackled these three challenges:

First, we identified simulation tasks which need an electrostatic baseline in order to achieve a correct description of the material. Using the same training data for crystalline $Li_7P_3S_{11}$, we found significant differences between a short range GAP model and the ES-GAP model when studying isotropic *vs.* non-isotropic systems. Here, for standard isotropic simulation tasks, such as determining Li diffusion barriers and ionic conductivities, both models yield similar results. In contrast, simulations on non-isotropic systems show the importance of ES contributions. More generally, our results confirm that short-ranged ML potentials can be surprisingly accurate for polar and ionic materials in the absence of non-isotropic chemical environments like interfaces or electric fields. In contrast we found important qualitative deviation between our GAP models in non-isotropic systems.

Second, we developed the kernel-based charge equilibration scheme called kQEq. The novel kQEq schemes enable the prediction of partial charges based on local environments by including the ability to predict non-local charge transfer. For a first application, kQEq models trained on molecular dipole moments display excellent performance, *en par* with or better than state-of-the-art dipole prediction schemes. The formalism of kQEq allows for physics-based kernel ML models for predicting atomic charges, to be used in combination with interatomic potentials such as GAP.

Third, we developed a near-universal GAP model for the crystalline and amorphous compounds in Li_2S-P_2S_5. We then used the GAP model to systematically investigate the effect of the local anion composition in glassy Li_2S-P_2S_5 compounds. The general structure of our training protocol allows for a variety of extensions, such additional selection criteria, doping with transition metals, and the future modeling of solid/solid interfaces.

Now in 2022, it seems we finally have all the tools to study solid-solid battery interfaces in an operando like fashion. Nevertheless, open challenges remain. First, we need to define a training procedure to train a kQEq model so that we can fit to energies, higher order moments, and periodic systems. Especially fitting to energies in a Δ-learning fashion is challenging. Second, material interfaces have shown to exhibit distinctively different stoichiometries, structure, and properties than either of the neighboring bulk phases. We therefore need a sensible way to to establish iterative training and exploration protocols that systematically extend the transferability of an electrostatic GAP model to the complex interphases that truly determine the performance of

all solid-state batteries. The speed-up achieved by electrostatic GAPs as compared to direct *ab initio* calculations will then allow extensive searches and sampling that should provide a much clearer picture of the yet missing structure-performance relationships that will ultimately enable a rational design and advancement.

Danksagung

Mein erster Dank gilt meinem Doktorvater Prof. Dr. Karsten Reuter ohne den diese Arbeit nicht möglich gewesen wäre. Er war stets ein ehrlicher Berater und hat mir alle Türen geöffnet, die mich hierhergebracht haben.

Mein weiterer wissenschaftlicher Dank gilt meinem Gruppenleiter Dr. Christoph Scheurer und meinem Mentor Dr. Johannes Margraf. Beide haben wesentlich meine wissenschaftliche Ausbildung vorangetrieben und mir gezeigt wie ein Wissenschaftler in seiner Arbeitsweise, wie auch moralisch handeln sollte.

A special thanks belongs to Prof Gabor Csanyi. He supported me, as if I was one of his students, taught me that we should never stop asking and welcomed me in his group in Cambridge.

Forschung ist nur möglich, wenn man von einer Gruppe von Menschen bei den nicht wissenschaftlichen Themen unterstützt wird. Hier gilt mein Dank dem IT Support und Admin Team: Martin, Simon, Simeon, Xhristoph, Matthias, David und Steffen. Ein ganz besonderer Dank gilt Ruth und Julia, ohne die der administrative Dschungel uns alle mit Sicherheit schon verschlungen hätte.

Ich möchte einigen besonderen Kollegen bedanken: Zuerst meinem Mistery Room Mitstreitern Hanna und Frederic. Zu dritt haben wir uns nicht nur durch das Studium, sondern auch durch unsere ganze Diss durchgeboxt. Ein besonderer Dank gilt meinem Compadre Jakob u.a. dafür, das wir beide jetzt in einer der besten Bars Münchens als Pilsdoktoren bekannt und verewigt sind. Danke an die Studenten die ich über die Zeit betreuen durfte und an denen ich meine Fähigkeiten als Lehrer üben durfte. Hier gilt mein besonderer Dank Tabea, die ich zwei Jahre betreuen durfte.

Unendlicher Dank gilt meinen treuen Freunden die mich über die Jahre begleitet haben. Angefangen bei Nati und Konni meinen ältesten Schulfreunden. Danke an meine Freunde aus der Heckscherstrasse 15. Hier habe ich gelernt was es heißt einen Bund fürs Leben zu schließen. Mein Dank gilt meinen Freunden aus meinen ersten universitären Schritten in der Mathematik Fakultät. Mein vermutlich größter Dank gilt aber meinen Chemikern - Markus, Ala, Paddi, Michi, Sina, Maxi, Chrissi, Jonas und PG9.

Zuletzt möchte ich meiner Familie und Louisa danken. Danke für einfach alles.

Munich, August, 2022

Bibliography

[1] A. Gore, *www.algore.com* (cit. on p. 1).

[2] International Energy Agency, *World energy outlook 2021, https://www.iea.org/reports/world-energy-outlook-2021* (cit. on p. 1).

[3] Organisation for Economic Co-operation and Development, *Oecd regional recovery platform, https://www.oecd.org/regional/recovery-platform.htm* (cit. on p. 1).

[4] International Energy Agency, *Https://www.iea.org/news/global-electric-car-sales-have-continued-their-strong-growth-in-2022-after-breaking-records-last-year* (cit. on p. 1).

[5] Senatsverwaltung für Wirtschaft, Energie und Betriebe Berlin, *Https://www.berlin.de/sen/energie/erneuerbare-energien/solargesetz-berlin/artikel.1053243.php* (cit. on p. 1).

[6] Bayerische Staatskanzlei, *Https://www.gesetze-bayern.de/content/document/bayklimag* (cit. on p. 1).

[7] J. B. Goodenough and M. H. Braga, Dalton Trans. **47**, 645 (2018) (cit. on p. 1).

[8] Z. Yang, J. Zhang, M. C. Kintner-Meyer, X. Lu, D. Choi, J. P. Lemmon, and J. Liu, Chem. Rev. **111**, 3577 (2011) (cit. on p. 1).

[9] C. Bauer, S. Burkhardt, N. P. Dasgupta, L. A.-W. Ellingsen, L. L. Gaines, H. Hao, R. Hischier, L. Hu, Y. Huang, J. Janek, et al., Nat. Sustain. **5**, 176 (2022) (cit. on p. 1).

[10] *Https://battery2030.eu/research/roadmap/* (cit. on p. 1).

[11] J. Janek and W. G. Zeier, Nat. Energy **1**, 1 (2016) (cit. on pp. 1, 3–5, 10).

[12] A. L. Robinson and J. Janek, *Solid-state batteries enter ev fray*, 2014 (cit. on pp. 1, 3).

[13] Ö. U. Kudu, T. Famprikis, B. Fleutot, M.-D. Braida, T. L. Mercier, M. S. Islam, and C. Masquelier, J. Power Sources **407**, 31 (2018) (cit. on pp. 1, 6, 7).

[14] N. Kamaya, K. Homma, Y. Yamakawa, M. Hirayama, R. Kanno, M. Yonemura, T. Kamiyama, Y. Kato, S. Hama, K. Kawamoto, et al., Nat. Mater. **10**, 682 (2011) (cit. on pp. 1, 3).

[15] H. Aono, E. Sugimoto, Y. Sadaoka, N. Imanaka, and G.-y. Adachi, J Electrochem Soc **137**, 1023 (1990) (cit. on p. 1).

[16] F. Mizuno, A. Hayashi, K. Tadanaga, and M. Tatsumisago, Solid State Ion. **177**, 2721 (2006) (cit. on p. 1).

[17] J. G. Kim, B. Son, S. Mukherjee, N. Schuppert, A. Bates, O. Kwon, M. J. Choi, H. Y. Chung, and S. Park, J. Power Sources. **282**, 299 (2015) (cit. on p. 1).

[18] A. D. Sendek, Q. Yang, E. D. Cubuk, K.-A. N. Duerloo, Y. Cui, and E. J. Reed, Energy Environ. Sci. **10**, 306 (2017) (cit. on p. 1).

[19] A. M. Abakumov, S. S. Fedotov, E. V. Antipov, and J.-M. Tarascon, Nat. Comm. **11**, 1 (2020) (cit. on p. 1).

[20] *Https://battery2030.eu/battery2030/projects/big-map/* (cit. on p. 1).

[21] S. Stegmaier, R. Schierholz, I. Povstugar, J. Barthel, S. P. Rittmeyer, S. Yu, S. Wengert, S. Rostami, K. Hans, K. Reuter, R. Eichel, and C. Scheurer, Adv. Energy Mater, 2100707 (2021) (cit. on pp. 1, 3, 4, 10).

[22] H. Türk, F.-P. Schmidt, T. Götsch, F. Girgsdies, A. Hammud, D. Ivanov, I. C. Vinke, L. de Haart, R.-A. Eichel, K. Reuter, et al., Adv. Mater. Interfaces **8**, 2100967 (2021) (cit. on p. 1).

[23] Y. Nagata, K. Nagao, M. Deguchi, A. Sakuda, A. Hayashi, H. Tsukasaki, S. Mori, and M. Tatsumisago, Chem. Mater. **30**, 6998 (2018) (cit. on p. 1).

[24] J. A. Lewis, F. J. Q. Cortes, M. G. Boebinger, J. Tippens, T. S. Marchese, N. Kondekar, X. Liu, M. Chi, and M. T. McDowell, ACS Energy Lett. **4**, 591 (2019) (cit. on p. 1).

[25] W. D. Richards, L. J. Miara, Y. Wang, J. C. Kim, and G. Ceder, Chem. Mater. **28**, 266 (2016) (cit. on p. 1).

[26] P. K. Nayak, G. Garcia-Belmonte, A. Kahn, J. Bisquert, and D. Cahen, Energy Environ. Sci. **5**, 6022 (2012) (cit. on p. 1).

[27] J. J. Choi, X. Yang, Z. M. Norman, S. J. Billinge, and J. S. Owen, Nano Lett. **14**, 127 (2014) (cit. on p. 1).

[28] H. Kulik, T. Hammerschmidt, J. Schmidt, S. Botti, M. A. Marques, M. Boley, M. Scheffler, M. Todorović, P. Rinke, C. Oses, et al., Electronic Structure **1**, 1 (2022) (cit. on pp. 1, 2).

[29] B. Huang and O. A. von Lilienfeld, Chem. Rev. **121**, 10001 (2021) (cit. on p. 1).

[30] M. Ceriotti, C. Clementi, and O. Anatole von Lilienfeld, Chem. Rev. **121**, 9719 (2021) (cit. on pp. 1, 2, 14).

[31] O. T. Unke and M. Meuwly, J. Chem. Theory Comput. **15**, 3678 (2019) (cit. on pp. 2, 29).

[32] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007) (cit. on pp. 2, 13, 15).

[33] N. Artrith, T. Morawietz, and J. Behler, Phy. Rev. B **83**, 153101 (2011) (cit. on pp. 2, 29, 30).

[34] N. Artrith and A. Urban, Comput. Mater. Sci. **114**, 135 (2016) (cit. on p. 2).

[35] N. Artrith, A. Urban, and G. Ceder, Phys. Rev. B **96**, 014112 (2017) (cit. on p. 2).

[36] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, J. Chem. Phys. **148**, 241722 (2018) (cit. on p. 2).

[37] V. L. Deringer and G. Csányi, Phys. Rev. B **95**, 094203 (2017) (cit. on pp. 2, 14, 25).

[38] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010) (cit. on pp. 2, 13–15).

[39] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013) (cit. on pp. 2, 14, 17).

[40] A. P. Bartók and G. Csányi, Int. J. Quantum Chem. **115**, 1051 (2015) (cit. on p. 2).

[41] F. Dommert, K. Wendler, R. Berger, L. Delle Site, and C. Holm, Chem. Phys. Chem. **13**, 1625 (2012) (cit. on p. 2).

[42] S. Riniker, J. Chem. Inf. Model **58**, 565 (2018) (cit. on p. 2).

[43] K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, J. Chem. Phys. **148**, 241727 (2018) (cit. on p. 2).

[44] F. Q. Nazar and C. Ortner, Arch. Ration. Mech. Anal **224**, 817 (2017) (cit. on p. 2).

[45] A. V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016) (cit. on p. 2).

[46] K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. Von Lilienfeld, K.-R. Müller, and A. Tkatchenko, J. Phys. Chem. Lett **6**, 2326 (2015) (cit. on p. 2).

[47] V. L. Deringer, N. Bernstein, G. Csányi, C. Ben Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, and S. R. Elliott, Nature **589**, 59 (2021) (cit. on p. 2).

[48] M. Gastegger, J. Behler, and P. Marquetand, Chem. Sci. **8**, 6924 (2017) (cit. on p. 2).

[49] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, Npj Comput. Mater. **6**, 1 (2020) (cit. on p. 2).

[50] O. C. Harris, S. E. Lee, C. Lees, and M. Tang, J. Phys. Energy **2**, 032002 (2020) (cit. on p. 3).

[51] J. Wang, Y.-S. He, and J. Yang, Adv. Mater. **27**, 569 (2015) (cit. on p. 3).

[52] Y. Jiang, F. Chen, Y. Gao, Y. Wang, S. Wang, Q. Gao, Z. Jiao, B. Zhao, and Z. Chen, J. Power Sources. **342**, 929 (2017) (cit. on p. 3).

[53] Y. Diao, K. Xie, S. Xiong, and X. Hong, J. Power Sources **235**, 181 (2013) (cit. on p. 3).

[54] A. Manthiram, X. Yu, and S. Wang, Nat. Rev. Mater. **2**, 1 (2017) (cit. on p. 3).

[55] A. M. Stephan and K. Nahm, Polymer **47**, 5952 (2006) (cit. on p. 3).

[56] Y. Kato, S. Hori, T. Saito, K. Suzuki, M. Hirayama, A. Mitsui, M. Yonemura, H. Iba, and R. Kanno, Nat. Energy **1**, 1 (2016) (cit. on p. 3).

[57] T. Minami, A. Hayashi, and M. Tatsumisago, Solid State Ion. **177**, 2715 (2006) (cit. on p. 3).

[58] C. Dietrich, D. A. Weber, S. J. Sedlmaier, S. Indris, S. P. Culver, D. Walter, J. Janek, and W. G. Zeier, J. Mater. Chem. A **5**, 18111 (2017) (cit. on pp. 3, 6, 9, 10).

[59] Y. Zhu, X. He, and Y. Mo, ACS Appl. Mater. Interfaces **7**, 23685 (2015) (cit. on p. 3).

[60] K. J. Kim, M. Balaish, M. Wadaguchi, L. Kong, and J. L. Rupp, Adv. Energy Mater. **11**, 2002689 (2021) (cit. on p. 3).

[61] A. C. Luntz, J. Voss, and K. Reuter, J. Phys. Chem. Lett. **6**, 4599 (2015) (cit. on p. 3).

[62] P. Li, Y. Zhao, Y. Shen, and S.-H. Bo, J. Phys. Energy **2**, 022002 (2020) (cit. on p. 4).

[63] F. Han, A. S. Westover, J. Yue, X. Fan, F. Wang, M. Chi, D. N. Leonard, N. J. Dudney, H. Wang, and C. Wang, Nature Energy **4**, 187 (2019) (cit. on p. 4).

[64] S. Chen, D. Xie, G. Liu, J. P. Mwizerwa, Q. Zhang, Y. Zhao, X. Xu, and X. Yao, Energy Storage Mater. **14**, 58 (2018) (cit. on p. 6).

[65] A. Sakuda, A. Hayashi, and M. Tatsumisago, Sci. Rep. **3**, 1 (2013) (cit. on p. 6).

[66] J.-S. Kim, W. D. Jung, S. Choi, J.-W. Son, B.-K. Kim, J.-H. Lee, and H. Kim, J. Phys. Chem. Lett. **9**, 5592 (2018) (cit. on p. 7).

[67] N. H. H. Phuc, M. Totani, K. Morikawa, H. Muto, and A. Matsuda, Solid State Ion. **288**, 240 (2016) (cit. on pp. 7, 8).

[68] M. Murayama, N. Sonoyama, A. Yamada, and R. Kanno, Solid State Ion. **170**, 173 (2004) (cit. on p. 7).

[69] K. Homma, M. Yonemura, T. Kobayashi, M. Nagao, M. Hirayama, and R. Kanno, Solid State Ion. **182**, 53 (2011) (cit. on p. 8).

[70] S. Iikubo, K. Shimoyama, S. Kawano, M. Fujii, K. Yamamoto, M. Matsushita, T. Shinmei, Y. Higo, and H. Ohtani, AIP Advances **8**, 015008 (2018) (cit. on p. 8).

[71] Y. Seino, T. Ota, K. Takada, A. Hayashi, and M. Tatsumisago, Energy Environ. Sci. **7**, 627 (2014) (cit. on p. 8).

[72] H. Yamane, M. Shibata, Y. Shimane, T. Junke, Y. Seino, S. Adams, K. Minami, A. Hayashi, and M. Tatsumisago, Solid State Ion. **178**, 1163 (2007) (cit. on pp. 8, 10).

[73] C. G. Staacke, H. H. Heenen, C. Scheurer, G. Csányi, K. Reuter, and J. T. Margraf, ACS Appl. Energy Mater. **4**, 12562 (2021) (cit. on pp. 9, 30).

[74] S. Wenzel, D. A. Weber, T. Leichtweiss, M. R. Busche, J. Sann, and J. Janek, Solid State Ion. **286**, 24 (2016) (cit. on p. 9).

[75] Y. Wang, W. D. Richards, S. P. Ong, L. J. Miara, J. C. Kim, Y. Mo, and G. Ceder, Nat. Mater. **14**, 1026 (2015) (cit. on p. 9).

[76] I.-H. Chu, H. Nguyen, S. Hy, Y.-C. Lin, Z. Wang, Z. Xu, Z. Deng, Y. S. Meng, and S. P. Ong, ACS Appl. Mater. Interfaces **8**, 7843 (2016) (cit. on p. 9).

[77] R. Mercier, J. Malugani, B. Fahys, J. Douglande, and G. Robert, J. Solid State Chem. **43**, 151 (1982) (cit. on p. 9).

[78] C. Dietrich, M. Sadowski, S. Sicolo, D. A. Weber, S. J. Sedlmaier, K. S. Weldert, S. Indris, K. Albe, J. Janek, and W. G. Zeier, Chem. Mat. **28**, 8764 (2016) (cit. on pp. 9, 10).

[79] Z. D. Hood, C. Kates, M. Kirkham, S. Adhikari, C. Liang, and N. Holzwarth, Solid State Ion. **284**, 61 (2016) (cit. on pp. 9, 10).

[80] Y. Seino, M. Nakagawa, M. Senga, H. Higuchi, K. Takada, and T. Sasaki, J. Mater. Chem. **3**, 2756 (2015) (cit. on pp. 9, 10).

[81] M. B. Preefer, J. H. Grebenkemper, C. E. Wilson, M. Everingham, J. A. Cooley, and R. Seshadri, ACS Appl. Mater. Interfaces **13**, 57567 (2021) (cit. on p. 10).

[82] K. Ohara, A. Mitsui, M. Mori, Y. Onodera, S. Shiotani, Y. Koyama, Y. Orikasa, M. Murakami, K. Shimoda, K. Mori, et al., Sci. Rep. **6**, 1 (2016) (cit. on p. 10).

[83] Y. Aoki, K. Ogawa, T. Nakagawa, Y. Hasegawa, Y. Sakiyama, T. Kojima, and M. Tabuchi, Solid State Ion. **310**, 50 (2017) (cit. on p. 10).

[84] J. G. Smith and D. J. Siegel, Nature Comm. **11**, 1 (2020) (cit. on p. 10).

[85] T. Baba and Y. Kawamura, Front. Energy Res. **4** (2016) 10.3389/fenrg.2016.00022 (cit. on p. 10).

[86] S. Shiotani, K. Ohara, H. Tsukasaki, S. Mori, and R. Kanno, Sci. Rep. **7** (2017) 10.1038/s41598-017-07086-y (cit. on p. 10).

[87] M. Kick, C. Grosu, M. Schuderer, C. Scheurer, and H. Oberhofer, J. Phys. Chem. Lett. **11**, 2535 (2020) (cit. on p. 10).

[88] C. Grosu, S. Döpking, S. Merz, C. Panosetti, P. Jakes, S. Annies, S. Matera, J. Granwehr, and C. Scheurer, in Ecs meeting abstracts, 1 (IOP Publishing, 2020), p. 96 (cit. on p. 10).

[89] J. Timmermann, F. Kraushofer, N. Resch, P. Li, Y. Wang, Z. Mao, M. Riva, Y. Lee, C. Staacke, M. Schmid, et al., Phys. Rev. Lett. **125**, 206101 (2020) (cit. on pp. 10, 25).

[90] F. Vitalini, A. S. Mey, F. Noé, and B. G. Keller, J. Chem. Phys. **142**, 611 (2015) (cit. on p. 13).

[91] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984) (cit. on p. 13).

[92] J. Tersoff, Phys. Rev. B **37**, 6991 (1988) (cit. on p. 13).

[93] R. Drautz, Phys. Rev. B **99**, 014104 (2019) (cit. on pp. 13, 15, 19).

[94] F. A. Faber, A. S. Christensen, B. Huang, and O. A. von Lilienfeld, J. Chem. Phys. **148**, 241717 (2018) (cit. on pp. 13, 15).

[95] A. S. Christensen, L. A. Bratholm, F. A. Faber, and O. Anatole von Lilienfeld, J. Chem. Phys. **152**, 044107 (2020) (cit. on p. 13).

[96] G. Csányi, S. Winfield, J. R. Kermode, A. De Vita, A. Comisso, N. Bernstein, and M. C. Payne, IoP Comput. Phys. Newsletter, 1 (2007) (cit. on p. 14).

[97] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Chem. Rev. **121**, 9759 (2021) (cit. on pp. 14, 17, 18).

[98] S. N. Pozdnyakov, M. J. Willatt, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, Phys. Rev. Lett **125**, 166001 (2020) (cit. on p. 14).

[99] B. Onat, C. Ortner, and J. R. Kermode, J. Chem. Phys. **153**, 144106 (2020) (cit. on p. 14).

[100] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Chem. Rev. **121**, 10073 (2021) (cit. on pp. 20, 22–24).

[101] J. C. Timmermann, "Iridium oxide as catalyst in water electrolysis: identification of novel surface structures via machine learning," PhD thesis (Technische Universität München, 2022) (cit. on p. 21).

[102] B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, Proc. Natl. Acad. Sci. **116**, 1110 (2019) (cit. on p. 25).

[103] L. Pauling, *The nature of the chemical bond, and the structure of molecules and crystals - an introduction to modern structural chemistry*, 3rd ed. (Cornell University Press, New York, 1960) (cit. on p. 26).

[104] R. S. Mulliken, J. Chem. Phys. **2**, 782 (1934) (cit. on p. 26).

[105] R. Sanderson, Science **114**, 670 (1951) (cit. on p. 26).

[106] R. P. Iczkowski and J. L. Margrave, J. Am. Chem. Soc. **83**, 3547 (1961) (cit. on p. 26).

[107] R. G. Parr and R. G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983) (cit. on p. 26).

[108] W. J. Mortier, S. K. Ghosh, and S. Shankar, J. Am. Chem. Soc. **108**, 4315 (1986) (cit. on p. 26).

[109] D. Ongari, P. G. Boyd, O. Kadioglu, A. K. Mace, S. Keskin, and B. Smit, J. Chem. Theory Comput. **15**, 382 (2018) (cit. on p. 29).

[110] B. Nebgen, N. Lubbers, J. S. Smith, A. E. Sifain, A. Lokhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, J. Chem. Theory Comput. **14**, 4687 (2018) (cit. on p. 29).

[111] T. Bereau, D. Andrienko, and O. A. Von Lilienfeld, J. Chem. Theory Comput. **11**, 3225 (2015) (cit. on p. 29).

[112] T. Bereau, R. A. DiStasio Jr, A. Tkatchenko, and O. A. Von Lilienfeld, J. Chem. Phys. **148**, 241706 (2018) (cit. on p. 29).

[113] P. Bleiziffer, K. Schaller, and S. Riniker, J. Chem. Inf. Model. **58**, 579 (2018) (cit. on p. 29).

[114] B. K. Rai and G. A. Bakken, J. Comp. Chem. **34**, 1661 (2013) (cit. on p. 29).

[115] V. V. Korolev, A. Mitrofanov, E. I. Marchenko, N. N. Eremin, V. Tkachenko, and S. N. Kalmykov, Chem. Mater. **32**, 7822 (2020) (cit. on p. 29).

[116] R. Zubatyuk, J. S. Smith, B. T. Nebgen, S. Tretiak, and O. Isayev, Nat. Comm. **12**, 1 (2021) (cit. on p. 29).

[117] S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, Phys. Rev. B **92**, 045131 (2015) (cit. on pp. 29, 31, 32).

[118] T. Morawietz, V. Sharma, and J. Behler, J. Chem. Phys. **136**, 064103 (2012) (cit. on p. 30).

[119] J. Behler, Chem. Rev. **121**, 10037 (2021) (cit. on pp. 30, 32).

[120] A. E. Sifain, N. Lubbers, B. T. Nebgen, J. S. Smith, A. Y. Lokhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, J. Phys. Chem. Lett. **9**, 4495 (2018) (cit. on p. 30).

ACS APPLIED ENERGY MATERIALS

www.acsaem.org Article

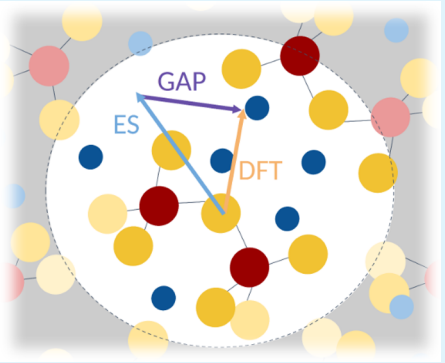
On the Role of Long-Range Electrostatics in Machine-Learned Interatomic Potentials for Complex Battery Materials

Carsten G. Staacke, Hendrik H. Heenen, Christoph Scheurer, Gábor Csányi, Karsten Reuter, and Johannes T. Margraf*

Cite This: *ACS Appl. Energy Mater.* 2021, 4, 12562−12569 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Modeling complex energy materials such as solid-state electrolytes (SSEs) realistically at the atomistic level strains the capabilities of state-of-the-art theoretical approaches. On one hand, the system sizes and simulation time scales required are prohibitive for first-principles methods such as the density functional theory. On the other hand, parameterizations for empirical potentials are often not available, and these potentials may ultimately lack the desired predictive accuracy. Fortunately, modern machine learning (ML) potentials are increasingly able to bridge this gap, promising first-principles accuracy at a much reduced computational cost. However, the local nature of these ML potentials typically means that long-range contributions arising, for example, from electrostatic interactions are neglected. Clearly, such interactions can be large in polar materials such as electrolytes, however. Herein, we investigate the effect that the locality assumption of ML potentials has on lithium mobility and defect formation energies in the SSE $Li_7P_3S_{11}$. We find that neglecting long-range electrostatics is unproblematic for the description of lithium transport in the isotropic bulk. In contrast, (field-dependent) defect formation energies are only adequately captured by a hybrid potential combining ML and a physical model of electrostatic interactions. Broader implications for ML-based modeling of energy materials are discussed.



KEYWORDS: *machine learning, electrostatics, battery, solid-state electrolyte, locality*

1. INTRODUCTION

The development of new analytical approximation frameworks is currently leading to an unparalleled surge of machine learning (ML) approaches in all areas of chemistry and materials science.[1−5] Here, ML is typically considered a universal approach for learning (fitting) a complex relationship $y = f(x)$ without explicitly knowing the physical (analytic) form of f.[6,7] In the context of interatomic potentials, this means establishing the relationship between a system's atomistic structure and its total energy $E = f(\{\boldsymbol{Z}, \boldsymbol{R}\})$, where $\boldsymbol{Z}$ are the atomic numbers and $\boldsymbol{R}$ are the position vectors of the constituting atoms. The expectation here is that flexible ML potentials can overcome long-standing limitations of empirical potentials that use simple fixed functional forms.[7]

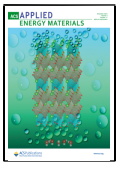
Such limitations are especially acute when covalent bonds are formed or broken, when atoms vary their hybridization or charge state, and generally when large changes in chemical environments occur. All these aspects apply prominently to the simulation of *operando* energy conversion systems in general and battery materials in particular.[3,8−15] With the structural and compositional complexity of contemporary battery materials severely limiting direct first-principles-based simulations, there is thus considerable hope that ML potentials trained with first-principles data will enable simulations at unprecedented length and time scales and a predictive quality matching that of electronic structure methods.[16−23]

To achieve size extensivity and create a general ML potential that can be employed for systems of varying size and composition, just as with many empirical potentials (*e.g.*, embedded atom models and Tersoff potentials), a locality assumption is typically made.[6,7,24] The system's total energy is thus approximated as a sum of local (atomic) contributions

$$E = \sum_{i}^{N} \epsilon(Z_i, \chi_i) \tag{1}$$

where the sum runs over the N atoms in the system and each atom i contributes with an energy ϵ that only depends on its atomic number Z_i and its local chemical environment χ_i (which is itself a function of $\{\boldsymbol{Z}, \boldsymbol{R}\}$). This local environment is suitably encoded into a representation that (similar to the total energy)

Received: August 5, 2021
Accepted: October 19, 2021
Published: November 2, 2021



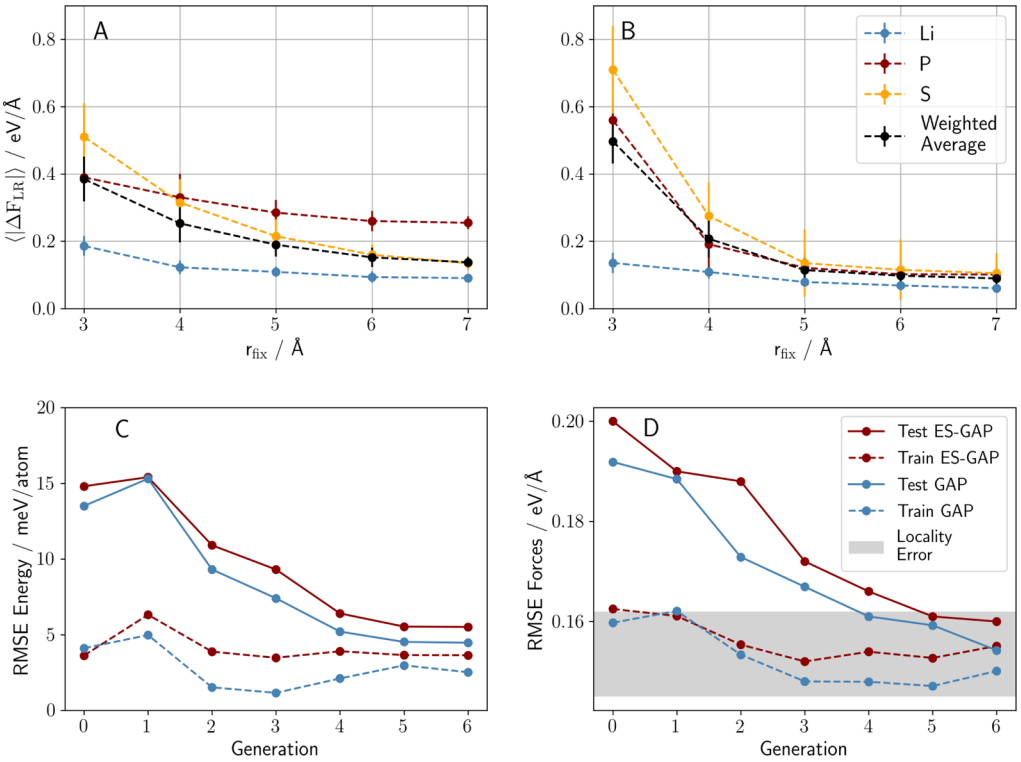


Figure 1. (A) Force locality in crystalline $Li_7P_3S_{11}$. (B) Force locality in crystalline $Li_7P_3S_{11}$ after subtracting the ES baseline model. Convergence of energies (C) and forces (D) throughout the generations for both models (GAP: blue and ES-GAP: red). Solid lines correspond to test errors and dashed lines to training errors. The gray shading corresponds to the expected force accuracy according to the locality test.

obeys general symmetries such as the invariance to translation, rotation, and permutation of atoms of the same element.[6,7,25−27] Importantly, to allow for efficient training and scalability to large simulation cells, these representations are almost always short-ranged, that is, they only describe the environment within a few Ångstroms around each atom.

In recent years, a large variety of ML potentials using kernel- or neural-network-based regression has been developed for molecular and condensed systems, heralding the great potential of this new data-driven approach.[6,7,28] Notably, such short-ranged ML potentials have also been applied to polar systems such as water.[29−32] Though one would naively expect long-range electrostatic interactions to play a significant role here, these potentials provide a remarkably accurate description of the structural and dynamic properties of bulk liquid water and different ice phases.[33] This is presumably because these systems are highly isotropic so that long-range interactions average out.[34] Consequently, short-ranged ML potentials are now commonly applied to study polar and even ionic systems.[30,35−37]

However, the importance of long-range effects will clearly depend on the material and property of interest and thus demands more systematic scrutiny.[38] For example, ionic diffusion in electrolytes may lead to the transient local accumulation and depletion of charges, which can break the isotropy of the electrostatic environment. Even more critically, grain-boundaries, interfaces, and defects may lead to a permanent localized polarization of materials. Finally, the effect of applied electric fields, for example, in batteries, can obviously only be studied if an electrostatic description is part of the model.

Notably, several groups have recently proposed ML models that explicitly include long-range electrostatics.[39−45] These range from simple point-charge models to polarizable models and full self-consistent approximations of the charge density. While these approaches offer a route to overcome the locality constraints of current ML potentials, they also lead to an increased computational cost, both in terms of training and evaluation. In particular, they break the favorable linear scaling of the computational cost with the system size. This makes it crucial to understand when such explicit treatments of electrostatics are necessary and when a local ML potential can be used instead.

The goal of this paper is to analyze the effect of electrostatic contributions on ML-based simulations of battery materials, using the Li mobility and interface stability of the crystalline phase of the solid-state electrolyte (SSE) $Li_7P_3S_{11}$ as an example.[46,47] $Li_7P_3S_{11}$ exhibits an exceptionally high Li-ion conductivity and has been suggested as a promising candidate for all-solid-state lithium batteries.[46,47] Large-scale simulations of this material, possible, for example, through Gaussian approximation potentials (GAPs), can provide critical insights into solving open challenges in connection with this material.

We develop two GAP models where one is strictly short-ranged (GAP), while the other includes a simple electrostatic baseline (ES-GAP). We find that the inclusion of long-range ES interactions only benefits the description of nonisotropic chemical environments, while diffusion properties in the homogeneous bulk material are well-captured by both potentials. In contrast, the stability of Frenkel defects in the presence of electric fields (which influence the material's stability at the inhomogeneous electrode/electrolyte interface) can only be captured with a model that includes long-range interactions.

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

2. METHODS

2.1. Computational Details. Reference density functional theory (DFT) calculations are performed with the Perdew−Burke−Ernzerhof functional, default "light" integration grids, and a "tier 1" basis set of numerical atomic orbitals, as implemented in FHI-aims.[48,49] The Brillouin zone is sampled with a 2 × 2 × 2 *k*-grid. Initial training configurations are generated with *ab initio* molecular dynamics (AIMD) using the Γ-point approximation for the k-grid. GAP-based MD and nudged-elastic-band (NEB) simulations are performed using the LAMMPS[50] code and the corresponding interface to QUIP.[51] Pairwise electrostatic interactions in the ES-GAP model are included *via* a fixed-charge model. To avoid the divergence of point-charge Coulomb interaction at short distances, atomic charge densities are modeled by s-type Slater orbitals as, for example, in the QEq charge equilibration model.[52,53] Further details on the ES-GAP are noted in Section A of the Supporting Information. For training set construction and data analysis, the atomic simulation environment, SciPy, and scikit-learn are used.[54−56]

2.2. GAP Training. To train the GAP models, a simple iterative procedure is used. Briefly, an initial model is trained on a set of 80 crystalline $Li_7P_3S_{11}$ configurations, taken from a short DFT-based MD simulation and Monte Carlo-sampled Li-ion distributions on crystal and interstitial sites. This potential is then used to generate new configurations *via* MD simulations at 800 K, which are added to the training set. This procedure is repeated for several iterations (termed "generations") until the force and energy errors on new configurations no longer improve. The thus-obtained models provide an increasingly accurate description of high-temperature crystalline $Li_7P_3S_{11}$. GAP and ES-GAP models are trained on identical configurations, and the ES-GAP model was used to generate new configurations in the iterative procedure. Further details can be found in the Supporting Information Section C.

3. RESULTS

3.1. GAP and ES-GAP Models. As discussed above, a hallmark of many-body ML potentials is the assumption that the total energy can be described as a sum of local atomic contributions, which corresponds to a complete neglect of long-range interactions. The locality of these interactions is thus often tacitly assumed when ML potentials are generated. However, it can also be quantified more rigorously by analyzing the force ΔF_{LR} induced on a reference atom by perturbations of other atoms in the distance beyond a given cutoff radius.[57] This locality test is performed with the reference method before fitting an ML potential and can thus be considered the material property. Figure 1A shows the mean induced force and its standard deviation for $Li_7P_3S_{11}$ (see the Supporting Information for details).

As can be seen, the induced forces are quite large (between 0.1 and 0.5 eV/Å) and decay slowly with the cutoff. This is particularly evident for the phosphorous atoms, which bear the largest formal charge in this system (+5). Importantly, these forces are by construction long-ranged since they exclusively originate from perturbations beyond the given radius. They therefore cannot be described by an ML potential with the corresponding cutoff. In this sense, the mean induced force $\langle|\Delta F_{LR}|\rangle$ provides a lower bound for the residual force errors that an ML potential with a given cutoff can achieve.

In Figure 1B, the same locality test is performed after subtracting the fixed-charge ES baseline model from the DFT forces (see the Supporting Information for details). For cutoffs larger than 5 Å, this significantly lowers $\langle|\Delta F_{LR}|\rangle$, most prominently for phosphorous. Counterintuitively, the induced forces on phosphorous and sulfur are actually increased at shorter distances. On one hand, this is because the charges of this model were parameterized to minimize the locality error at 6 Å and are thus not ideal for shorter cutoffs. On the other hand, a fixed-charge ES model is generally inaccurate for short-range interactions, where polarization, charge transfer, and induction effects become important. To capture such effects with a baseline model would require the use of more complex polarizable models. Nonetheless, cutoffs of 6 Å are commonly used in state-of-the-art ML potentials so that a fixed charge ES baseline can be used here, though residual errors remain and are discussed below.

The convergence of the iterative training procedure can be seen in Figure 1C,D. This shows that the energy and force errors of both models show no further improvement between the fifth and sixth generations. Importantly, the final force errors fall into the expected range estimated from the locality test (see Figure 1B). This indicates the convergence of the training process, meaning that the remaining error will not be significantly reduced by further training but is instead related to the locality of the model and/or potential inadequacies of the representation.

Interestingly, the root mean squared errors (RMSEs) for predicted energies and forces are actually slightly lower for the short-range GAP model. By analyzing the errors of the individual elements separately (see the Supporting Information), we find that the ES-GAP displays somewhat higher errors for sulfur but lower errors for lithium and especially phosphorous. Since sulfur is the most abundant element in $Li_7P_3S_{11}$, this leads to the better average performance of the short-range GAP.

This points to a disadvantage of the fixed-charge approach used here: $Li_7P_3S_{11}$ features two distinct sulfur species, namely, in bridging and terminal positions. Hirshfeld population analysis (see the Supporting Information) indicates that these species correspond to different charge states, while they are treated equivalently by the ES baseline. This introduces an error, which needs to be compensated by the GAP potential. In principle, this could be mitigated by assigning different charges to these sulfur species. However, such atom typing would run counter to one of the main advantages of ML potentials relative to classical force-fields, namely, the fact that they can break and form bonds. A more satisfying solution would be the use of floating-charge models, and this will be explored in future work.[42,58]

From a different perspective, the differences in force errors observed for different elements also illustrate a weakness in average error metrics such as the RMSE (or the least-squares loss function minimized by the GAP) for multispecies systems: if the stoichiometry of a material is not balanced, more abundant species are implicitly weighted more strongly by the metric. In the present case, sulfur has the largest weight, although lithium is arguably more important. Nevertheless, the accuracy of both models is actually quite satisfying overall, considering the magnitudes of the force components in the training and test sets, which range up to ca. 10 eV/Å. Note, however, that in principle, it would be possible to use different weightings in the loss function for forces on different elements. In GAP models, this can be achieved by specifying individual regularization parameters for each force.[59]

3.2. Lithium-Ion Mobility. Both potentials introduced in the previous section are trained on the same data and have approximately the same force error, though small differences can be seen when looking at the description of individual elements. Do these differences affect the prediction of the

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

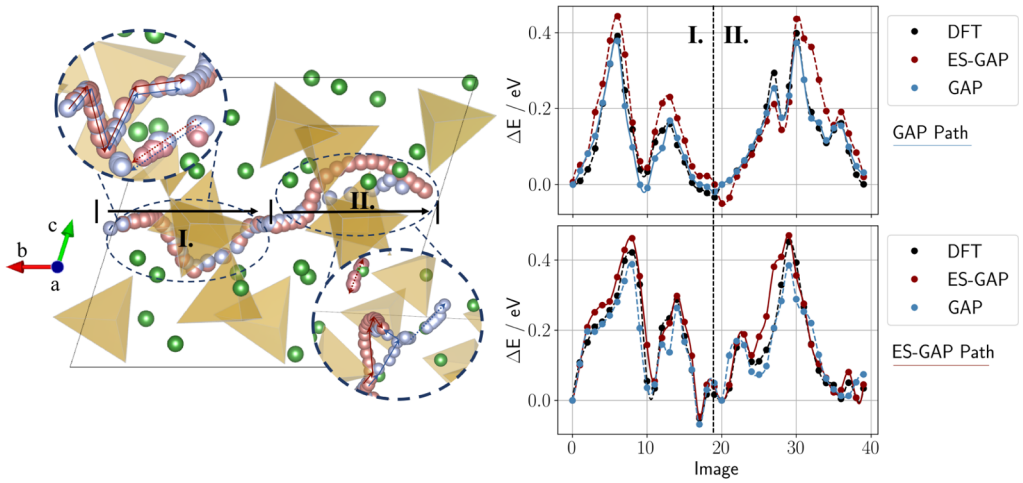


Figure 2. Left: Illustration of minimum-energy paths for lithium diffusion through $Li_7P_3S_{11}$ obtained with short-range GAP (light blue) and ES-GAP (red) interatomic potentials. Li ions are shown as green spheres and thiophosphates as orange tetrahedra. Solid and dashed arrows highlight the main migration path and displacements of neighboring lithium atoms, respectively. The labels I and II indicate two consecutive NEB calculations. Initial and final positions are obtained by analyzing hopping events from MD trajectories (see Supporting Information Section G). Right: Energies for DFT, GAP, and ES-GAP potentials on the minimum-energy paths calculated with the short-range GAP (top) and the ES-GAP (bottom). Solid lines indicate the potential with which the path was obtained and dashed lines indicate single-point calculations.

observables relevant to battery performance? As the first case in point, we investigate the Li-ion mobility in the (isotropic) bulk material. To this end, we consider lithium diffusion barriers obtained *via* the NEB method (see Figure 2) and Li-ion mobilities obtained from MD simulations.

NEB calculations allow the investigation of minimum-energy paths of individual Li hops between two equilibrium positions (an initial state and a final state).[60] Choosing these states is actually nontrivial since $Li_7P_3S_{11}$ contains a large variety of possible Li interstitial positions, reflecting the highly dynamic nature of the Li sublattice. This was previously demonstrated by Chang *et al.*, who reported a number of Li configurations with nearly the same ground-state energy as that of the equilibrium crystal structure.[61,62] It is therefore important to focus on Li hopping events that actually contribute to conductivity and not just dynamic rearrangements of Li positions.

To obtain these relevant pathways, we therefore analyzed the training MD trajectories to isolate individual Li hopping events. In this manner, a variety of diffusion pathways were obtained. Further information on all pathways is given in the Supporting Information (Section G). In the following, the lowest-barrier pathway is discussed in more detail. Here, a Li ion diffuses in a channel formed by the PS_4^- and $P_2S_7^-$ anion complexes (positions are tabulated in the Supporting Information Section G), along the *b* lattice vector, as shown in Figure 2. It can be seen that the $1 \times 2 \times 1$ supercell is traversed with two consecutive Li-ion hops (obtained from two NEB calculations Figure 2 I,II). These NEB calculations were performed based on both the GAP and ES-GAP potentials, leading to slightly different minimum-energy paths.

Nevertheless, the optimized lowest-energy paths display similar characteristics of a correlated ion migration where lithium ions diffuse, while neighboring ions are slightly displaced from the diffusion path (highlighted in the figure with solid and dashed arrows for the migrating and displaced atoms, respectively). For the first hop, an almost identical path is found, while the second hop yields a slightly different path when optimized with the two potentials (I and II in Figure 2, respectively). The deviating paths should not be understood as different mechanisms favored by the respective potentials, however, but merely as two feasible paths found by the NEB. We confirm this assumption by evaluating the energies along the short-range GAP path with the ES-GAP (and vice-versa). This analysis reveals almost identical barrier heights for a given potential on both its own NEB path and the one from the other potential. It is further noteworthy that the ES-GAP consistently predicts somewhat higher barriers than the short-range GAP. For both paths, reference DFT single-point values tend to lie between the GAP and ES-GAP values. In other words, the GAP somewhat underestimates the barriers, while the ES-GAP overestimates them to a similar degree. The deviation in energies is also not perfectly uniform along the paths so that the agreement with the DFT can be excellent for both potentials, at different points of the potential energy surface.

Despite the overall similar performance of the models, we thus find small systematic differences between the predictions of the two models. Interestingly, the short-range GAP is actually slightly more accurate in predicting the lowest barrier heights. This indicates that the static-charge model used for the ES-GAP does not faithfully reflect the electrostatics of the full DFT calculation, and the ES baseline represents an over-correction: it correctly increases the barriers but by too much. As a sidenote, we emphasize that both models are mainly trained on high-temperature MD data, while the NEB corresponds to the minimum energy path at 0 K. Presumably, an even higher accuracy for NEB calculations could be achieved for both potentials by training on the corresponding data.

Next, we turn to the Li-ion conductivity σ at finite temperature, predicted from MD simulations *via* the Nernst−Einstein equation (see the Supporting Information for details). Here, we shift the focus from a microscopic property (the Li migration barrier height) to a macroscopic observable (Li-ion conductivity). In principle, the two are closely related since the barrier height determines the rate of the Li transport in the transition state theory. However, MD

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

simulations sample a multitude of different diffusion events. These are often dominated by the lowest barrier mechanism but may also be influenced by higher barrier pathways (*e.g.*, because they are entropically favored). In this sense, the NEB and MD simulations provide complementary information.

The corresponding conductivities within the two potentials (at 800 K) are shown in Figure 3 as a function of the MD

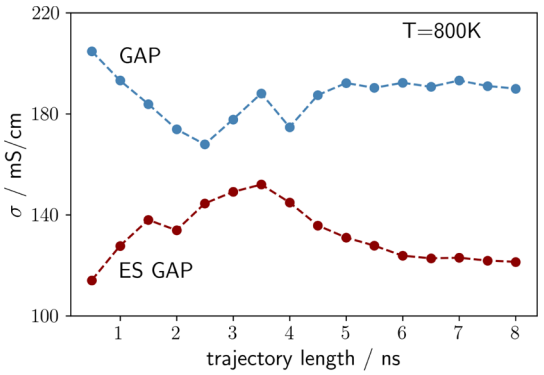


Figure 3. Convergence of the Li-ion conductivity σ at 800 K during an MD simulation. The red curve corresponds to the ES-GAP model and blue to the GAP model.

trajectory length. After sufficient sampling, the conductivities converge to 190 mS/cm for the GAP and 120 mS/cm for the ES-GAP. This trend in conductivities perfectly reflects the NEB barrier differences discussed above (*i.e.*, slightly higher mobility and lower barriers for the short-range GAP model).

Figure 3 also highlights the benefit of using ML potentials for battery research more generally: to obtain fully converged conductivities from these simulations, MD trajectory lengths far beyond the tractability of typical AIMD simulations are necessary. Hence, AIMD simulations of Li-ion conductivities should generally not be considered converged and yield, at best, a qualitative indicator of relative performance between closely related materials. Even at longer timescales in the low-nanosecond range (<5 ns), conductivity differences predicted by both models vary from 10 to 80%.

While we do find some differences between the predicted Li mobilities of the ES-GAP and GAP, these should be put into perspective. Ion mobilities of potential SSE materials can vary by several orders of magnitude, and the small differences in observed barrier heights are certainly within the margin of the DFT error. From a practical perspective, there is thus no significant difference between the two models.[63] At this point, one could conclude from a practitioner's point of view that a short-ranged ML potential would be fully sufficient to treat a complex battery material such as $Li_7P_3S_{11}$.

More broadly speaking, the observed insensitivity of energetic and dynamic properties of bulk $Li_7P_3S_{11}$ to long-range electrostatics is also in line with the previous reports on other polar liquids and solids.[29−32] For dynamic properties, electrostatic screening by mobile ions is clearly a significant cause of this behavior. Nonetheless, it is notable that the ES-GAP and GAP also perform very similarly in static calculations, as illustrated by the energy−volume curve of $Li_7P_3S_{11}$ (see the Supporting Information). This is a consequence of the fact that the crystalline system studied herein is highly isotropic. As a consequence, long-range interactions contribute to a large but nearly constant Madelung potential. These contributions cancel each other out to a large extent when considering forces and relative energies.

3.3. Applied Fields and Defects. Having established the similar behavior of the GAP and ES-GAP for bulk Li mobilities, we now turn to their description of an anisotropic environment. Specifically, we consider a model system that mimics the effect of the potential drop in the interphase region at the solid/solid interface between an electrode and an SSE. This potential variation has been previously computed using a 1-D continuum model[64] for a Li anode and graphite (C) cathode setup (see Figure 4). In this model, the potential drop

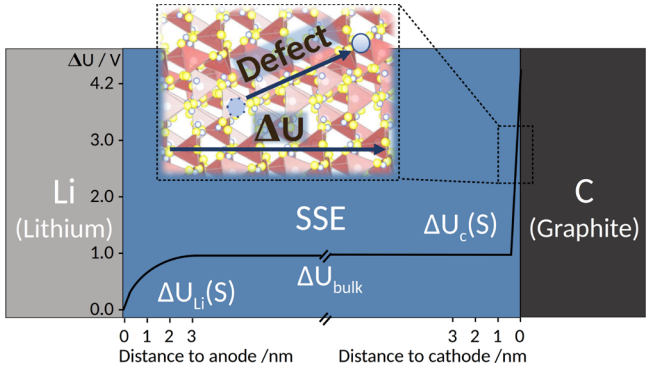


Figure 4. Schematic of the spatially dependent potential variation ΔU in a prototypical Li‖SSE‖C solid-state Li-ion battery.[64] Since the grain orientation varies throughout the SSE, the field-dependent defect stability is studied by applying electric fields along different crystallographic axes of the simulation cell.

reaches several nanometers into the bulk region before it is completely screened by displaced ions. To model the effect of this potential drop on the SSE near the interface, we apply electric fields to a 2 × 3 × 2 $Li_7P_3S_{11}$ supercell and investigate how the field strength and direction affect the stability of a Frenkel defect (*i.e.*, a Li vacancy/interstitial pair).

Defects have been argued to play an important role in the kinetics of the decomposition processes at SSE interfaces.[65] This kinetic stability is of high relevance to the applicability of SSEs.[66−68] Here, we consider a Frenkel defect in particular as it allows keeping the simulation cell overall charge neutral and lies within the phase space covered by the training set of the GAP potentials. Clearly, the effect of an electric potential drop can only be captured by the ES-GAP model, which contains charges that are able to respond to the applied field. In contrast, the short-range GAP model can only model the "zero-field" scenario. We also compare the "zero-field" defect stability in both models with that in the DFT.

We construct the Frenkel defect by shifting a Li atom into an interstitial position and relaxing the resulting structure using the two GAP models (see Figure 5 and Supporting Information Section I). For comparison, single-point DFT calculations are also performed on both the GAP and ES-GAP geometries. Without an applied field, this leads to a predicted defect formation energy of ca. 0.8 eV with the ES-GAP and DFT and ca. 0.6 eV with the short-range GAP model. Since the interstitial Li ion has a net positive charge and the ion vacancy a net negative charge, the defect forms a dipole. The observed 0.2 eV deviation of the defect formation energies between the GAP and DFT can thus be attributed to the absence of long-range dipole−dipole interactions in the former. Notably, the formation energies of neutral Li vacancies, where such dipoles

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

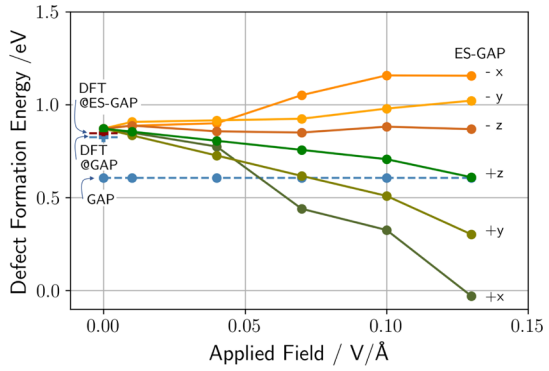


Figure 5. Frenkel defect formation energies against the applied field. The electric field is applied along the different crystallographic axes (*x*, *y*, *z*). Solid lines (green and orange) correspond to ES-GAP calculations with different applied field directions. Defect positions are given in the Supporting Information I.

are absent, reportedly do not show such a strong dependence on long-range electrostatics.[16]

Nonetheless, all models predict defect formation to be highly unfavorable. When applying an electric field, this picture changes. We find an anisotropic response to the field where both the destabilization and stabilization can occur (Figure 5). This anisotropy reflects the fields' orientation relative to the defect dipole. Structural relaxation effects then lead to an unsymmetric stabilization/destabilization of the defect in either field direction.

While this is a rather simple model system, it already yields insights into the stability of $Li_7P_3S_{11}$ at the SSE/electrode interface. As shown in Figure 5, field strengths typically occurring at this interface (which can reach up to 0.3 V/Å[64]) are sufficient to make the formation of this defect energetically favorable. Consequently, one would expect an accumulation of such defects toward the interface. As recently suggested,[66] the kinetic processes in electrolyte decomposition can be related to the delithiation of the SSE. Intermediate to this delithiation process are local concentration gradients by Frenkel defects. We can therefore hypothesize from our findings that local fields play a crucial role in the evaluation of interphase stabilities. Further, we find that the defect stabilization is anisotropic to the crystallographic orientation. This finding might explain the previous observations that the SSE/anode interface stability of $Li_7P_3S_{11}$ was dependent on the crystallographic orientation of the latter.[69]

From a methodological perspective, this example shows that the explicit inclusion of electrostatic interactions will be indispensable for the computational study of battery materials under operating conditions. Indeed, even the contact between two different materials will cause a potential drop across the interface, albeit at a smaller length scale.[13] The good performance of the short-range GAP model in the previous section is thus not because long-range electrostatic interactions are small but because they are reasonably isotropic in a periodic calculation. Breaking this symmetry with an interface or by applying an electric field clearly shows the importance of electrostatics which, by design, cannot be incorporated into a model with a short-range cutoff.

We note that the response of the ES-GAP model to the applied field relies on the ionic partial charges of the baseline ES model. Hence, we exploited the corresponding *a priori* knowledge. Ideally, these charges could instead be determined in the training procedure, including model structures of the full interface, which, however, is beyond the scope of this conceptual study.

4. CONCLUSIONS

In this paper, we have systematically explored the influence of explicitly including electrostatic interactions in ML potentials for battery materials. Using the same ML approach and training data, we find significant differences between a short-range GAP model and the ES-GAP model that uses an ES baseline when studying isotropic *versus* nonisotropic systems. In standard isotropic simulation tasks, such as determining Li diffusion barriers and ionic conductivities, both models yield similar results. In contrast, simulations on nonisotropic systems show the importance of ES contributions and provide new insights into the interphase stability of $Li_7P_3S_{11}$.

Specifically, we studied Frenkel defects in an applied field, mimicking the potential drop at a solid/solid interface. In this setup, we found that the stabilization of the defects can occur already at moderate fields. This would favor the accumulation of defects toward the interphase, which could influence the kinetic stability of $Li_7P_3S_{11}$/electrode interfaces. Additionally, such stabilizations are anisotropic to the crystallographic orientation, making the grain shape and orientation an additional parameter to be considered in battery engineering and beyond.[70]

More generally, our results confirm that short-ranged ML potentials can be surprisingly accurate for polar and ionic materials in the absence of nonisotropic chemical environments such as interfaces or electric fields. In contrast, we found important qualitative deviation between our GAP models in nonisotropic systems. The further development of ML potentials with an explicit description of electrostatics therefore represents an important research goal on the way to the computational study of battery materials in *operando* conditions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsaem.1c02363.

GAP model details, locality test and locality-based charge fit, interstitial configurations via Voronoi tessellation, ionic mobility, species-resolved model convergence, charge analysis, NEB calculations and interstitial Li position, defect positions, and energy−volume curve (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Johannes T. Margraf − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany; Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany;* orcid.org/0000-0002-0862-5289; Email: margraf@fhi.mpg.de

Authors

Carsten G. Staacke − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany; Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

Hendrik H. Heenen − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany;* orcid.org/0000-0003-0696-8445

Christoph Scheurer − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany; Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

Gábor Csányi − *Engineering Laboratory, University of Cambridge, Cambridge CB2 1PZ, U.K.*

Karsten Reuter − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany; Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsaem.1c02363

Funding

Open access funded by Max Planck Society.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the funding from the German Research Foundation (DFG) through its Cluster of Excellence e-conversion EXC 2089/1.

■ REFERENCES

(1) von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. Exploring chemical compound space with quantum-based machine learning. *Nat. Rev. Chem.* **2020**, *4*, 347−358.

(2) Cheng, B.; Griffiths, R.-R.; Wengert, S.; Kunkel, C.; Stenczel, T.; Zhu, B.; Deringer, V. L.; Bernstein, N.; Margraf, J. T.; Reuter, K.; Csanyi, G. Mapping Materials and Molecules. *Acc. Chem. Res.* **2020**, *53*, 1981−1991.

(3) Deringer, V. L. Modelling and Understanding Battery Materials with Machine-Learning-Driven Atomistic Simulations. *J. Phys. Energy* **2020**, *2*, 041003.

(4) Stocker, S.; Csányi, G.; Reuter, K.; Margraf, J. T. Machine Learning in Chemical Reaction Space. *Nat. Commun.* **2020**, *11*, 5505.

(5) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71*, 361.

(6) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(7) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(8) Artrith, N.; Urban, A.; Ceder, G. Constructing First-Principles Phase Diagrams of Amorphous LixSi Using Machine-Learning-Assisted Sampling with an Evolutionary Algorithm. *J. Chem. Phys.* **2018**, *148*, 241711.

(9) Urban, A.; Seo, D.-H.; Ceder, G. Computational Understanding of Li-Ion Batteries. *npj Comput. Mater.* **2016**, *2*, 16002.

(10) Stegmaier, S.; Schierholz, R.; Povstugar, I.; Barthel, J.; Rittmeyer, S. P.; Yu, S.; Wengert, S.; Rostami, S.; Kungl, H.; Reuter, K.; Eichel, R. A.; Scheurer, C. Nano-Scale Complexions Facilitate Li Dendrite-Free Operation in LATP Solid-State Electrolyte. *Adv. Energy Mater.* **2021**, *11*, 2100707.

(11) Heenen, H. H.; Voss, J.; Scheurer, C.; Reuter, K.; Luntz, A. C. Multi-ion Conduction in Li3OCl Glass Electrolytes. *J. Phys. Chem. Lett.* **2019**, *10*, 2264−2269.

(12) Heenen, H. H.; Scheurer, C.; Reuter, K. Implications of occupational disorder on ion mobility in Li4Ti5O12 battery materials. *Nano Lett.* **2017**, *17*, 3884−3888.

(13) Stegmaier, S.; Voss, J.; Reuter, K.; Luntz, A. C. Li+ defects in a solid-state Li ion battery: theoretical insights with a Li3OCl electrolyte. *Chem. Mater.* **2017**, *29*, 4330−4340.

(14) Bruix, A.; Margraf, J. T.; Andersen, M.; Reuter, K. First-principles-based multiscale modelling of heterogeneous catalysis. *Nat. Catal.* **2019**, *2*, 659−670.

(15) Timmermann, J.; Kraushofer, F.; Resch, N.; Li, P.; Wang, Y.; Mao, Z.; Riva, M.; Lee, Y.; Staacke, C.; Schmid, M.; Scheurer, C.; Parkinson, G. S.; Diebold, U.; Reuter, K. IrO2 Surface Complexions Identified through Machine Learning and Surface Investigations. *Phys. Rev. Lett.* **2020**, *125*, 206101.

(16) Li, W.; Ando, Y.; Minamitani, E.; Watanabe, S. Study of Li Atom Diffusion in Amorphous Li3PO4 with Neural Network Potential. *J. Chem. Phys.* **2017**, *147*, 214106.

(17) Onat, B.; Cubuk, E. D.; Malone, B. D.; Kaxiras, E. Implanted Neural Network Potentials: Application to Li-Si Alloys. *Phys. Rev. B* **2018**, *97*, 094106.

(18) Lacivita, V.; Artrith, N.; Ceder, G. Structural and Compositional Factors That Control the Li-Ion Conductivity in LiPON Electrolytes. *Chem. Mater.* **2018**, *30*, 7077−7090.

(19) Deng, Z.; Chen, C.; Li, X.-G.; Ong, S. P. An Electrostatic Spectral Neighbor Analysis Potential for Lithium Nitride. *npj Comput. Mater.* **2019**, *5*, 75.

(20) Wang, C.; Aoyagi, K.; Wisesa, P.; Mueller, T. Lithium Ion Conduction in Cathode Coating Materials from On-the-Fly Machine Learning. *Chem. Mater.* **2020**, *32*, 3741−3752.

(21) Jalem, R.; Kanamori, K.; Takeuchi, I.; Nakayama, M.; Yamasaki, H.; Saito, T. Bayesian-Driven First-Principles Calculations for Accelerating Exploration of Fast Ion Conductors for Rechargeable Battery Application. *Sci. Rep.* **2018**, *8*, 5845.

(22) Houchins, G.; Viswanathan, V. An Accurate Machine-Learning Calculator for Optimization of Li-Ion Battery Cathodes. *J. Chem. Phys.* **2020**, *153*, 054124.

(23) Fujikake, S.; Deringer, V. L.; Lee, T. H.; Krynski, M.; Elliott, S. R.; Csányi, G. Gaussian Approximation Potential Modeling of Lithium Intercalation in Carbon Nanostructures. *J. Chem. Phys.* **2018**, *148*, 241714.

(24) Jung, H.; Stocker, S.; Kunkel, C.; Oberhofer, H.; Han, B.; Reuter, K.; Margraf, J. T. Size-Extensive Molecular Machine Learning with Global Representations. *ChemSystemsChem* **2020**, *2*, No. e1900052.

(25) Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **2019**, *99*, 014104.

(26) Faber, F. A.; Christensen, A. S.; Huang, B.; von Lilienfeld, O. A. Alchemical and Structural Distribution Based Representation for Universal Quantum Machine Learning. *J. Chem. Phys.* **2018**, *148*, 241717.

(27) Christensen, A. S.; Bratholm, L. A.; Faber, F. A.; Anatole von Lilienfeld, O. FCHL Revisited: Faster and More Accurate Quantum Machine Learning. *J. Chem. Phys.* **2020**, *152*, 044107.

(28) Shapeev, A. V. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **2016**, *14*, 1153−1173.

(29) Monserrat, B.; Brandenburg, J. G.; Engel, E. A.; Cheng, B. Extracting ice phases from liquid water: why a machine-learning water model generalizes so well. **2020**, arXiv preprint arXiv:2006.13316.

(30) Cooper, A. M.; Kästner, J.; Urban, A.; Artrith, N. Efficient Training of ANN Potentials by Including Atomic Forces via Taylor Expansion and Application to Water and a Transition-Metal Oxide. *npj Comput. Mater.* **2020**, *6*, 54.

(31) Schran, C.; Behler, J.; Marx, D. Automated Fitting of Neural Network Potentials at Coupled Cluster Accuracy: Protonated Water Clusters as Testing Ground. *J. Chem. Theory Comput.* **2020**, *16*, 88−99.

(32) Schran, C.; Brieuc, F.; Marx, D. Transferability of Machine Learning Potentials: Protonated Water Neural Network Potential Applied to the Protonated Water Hexamer. *J. Chem. Phys.* **2021**, *154*, 051101.

(33) Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab initio thermodynamics of liquid and solid water. *PNAS* **2019**, *116*, 1110−1115.
(34) Cox, S. J. Dielectric response with short-ranged electrostatics. *PNAS* **2020**, *117*, 19746−19752.
(35) Vandermause, J.; Torrisi, S. B.; Batzner, S.; Xie, Y.; Sun, L.; Kolpak, A. M.; Kozinsky, B. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Comput. Mater.* **2020**, *6*, 20.
(36) Jinnouchi, R.; Lahnsteiner, J.; Karsai, F.; Kresse, G.; Bokdam, M. Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference. *Phys. Rev. Lett.* **2019**, *122*, 225701.
(37) Tovey, S.; Narayanan Krishnamoorthy, A.; Sivaraman, G.; Guo, J.; Benmore, C.; Heuer, A.; Holm, C. DFT Accurate Interatomic Potential for Molten NaCl from Machine Learning. *J. Phys. Chem. C* **2020**, *124*, 25760−25768.
(38) Yue, S.; Muniz, M. C.; Calegari Andrade, M. F.; Zhang, L.; Car, R.; Panagiotopoulos, A. Z. When do short-range atomistic machine-learning models fall short? *J. Chem. Phys.* **2021**, *154*, 034111.
(39) Faraji, S.; Ghasemi, S. A.; Rostami, S.; Rasoulkhani, R.; Schaefer, B.; Goedecker, S.; Amsler, M. High Accuracy and Transferability of a Neural Network Potential through Charge Equilibration for Calcium Fluoride. *Phys. Rev. B* **2017**, *95*, 104105.
(40) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic Potentials for Ionic Systems with Density Functional Accuracy Based on Charge Densities Obtained by a Neural Network. *Phys. Rev. B* **2015**, *92*, 045131.
(41) Grisafi, A.; Fabrizio, A.; Meyer, B.; Wilkins, D. M.; Corminboeuf, C.; Ceriotti, M. Transferable Machine-Learning Model of the Electron Density. *ACS Cent. Sci.* **2019**, *5*, 57−64.
(42) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer. *Nat. Commun.* **2021**, *12*, 398.
(43) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678−3693.
(44) Xie, X.; Persson, K. A.; Small, D. W. Incorporating Electronic Information into Machine Learning Potential Energy Surfaces via Approaching the Ground-State Electronic Energy as a Function of Atom-Based Electronic Populations. *J. Chem. Theory Comput.* **2020**, *16*, 4256−4270.
(45) Yao, K.; Herr, J. E.; Toth, D. W.; McKintyre, R.; Parkhill, J. The TensorMol-0.1 Model Chemistry: A Neural Network Augmented with Long-Range Physics. *Chem. Sci.* **2018**, *9*, 2261−2269.
(46) Busche, M. R.; Weber, D. A.; Schneider, Y.; Dietrich, C.; Wenzel, S.; Leichtweiss, T.; Schröder, D.; Zhang, W.; Weigand, H.; Walter, D.; Sedlmaier, S. J.; Houtarde, D.; Nazar, L. F.; Janek, J. In Situ Monitoring of Fast Li-Ion Conductor Li7P3S11 Crystallization Inside a Hot-Press Setup. *Chem. Mater.* **2016**, *28*, 6152−6165.
(47) Dietrich, C.; Weber, D. A.; Sedlmaier, S. J.; Indris, S.; Culver, S. P.; Walter, D.; Janek, J.; Zeier, W. G. Lithium ion conductivity in Li2S-P2S5 glasses - building units and local structure evolution during the crystallization of superionic conductors Li3PS4, Li7P3S11 and Li4P2S7. *J. Mater. Chem. A* **2017**, *5*, 18111−18119.
(48) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(49) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175−2196.
(50) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(51) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 184115.
(52) Rappe, A. K.; Goddard, W. A., III Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.
(53) Streitz, F. H.; Mintmire, J. W. Electrostatic potentials for metal-oxide surfaces and interfaces. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 11996.
(54) Hjorth Larsen, A.; et al. The atomic simulation environment-a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.
(55) Virtanen, P.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.
(56) Pedregosa, F.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(57) Deringer, V. L.; Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **2017**, *95*, 094203.
(58) Unke, O. T.; Meuwly, M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678−3693.
(59) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.
(60) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(61) Chang, D.; Oh, K.; Kim, S. J.; Kang, K. Super-ionic conduction in solid-state Li7P3S11-type sulfide electrolytes. *Chem. Mater.* **2018**, *30*, 8764−8770.
(62) Yamane, H.; Shibata, M.; Shimane, Y.; Junke, T.; Seino, Y.; Adams, S.; Minami, K.; Hayashi, A.; Tatsumisago, M. Crystal structure of a superionic conductor, Li7P3S11. *Solid State Ionics* **2007**, *178*, 1163−1167.
(63) Döpking, S.; Plaisance, C. P.; Strobusch, D.; Reuter, K.; Scheurer, C.; Matera, S. Addressing global uncertainty and sensitivity in first-principles based microkinetic models by an adaptive sparse grid approach. *J. Chem. Phys.* **2018**, *148*, 034102.
(64) Luntz, A. C.; Voss, J.; Reuter, K. Interfacial challenges in solid-state Li ion batteries. *J. Phys. Chem. Lett.* **2015**, *6*, 4599−4604.
(65) Hong, L.; Yang, K.; Tang, M. A mechanism of defect-enhanced phase transformation kinetics in lithium iron phosphate olivine. *npj Comput. Mater.* **2019**, *5*, 118.
(66) Schwietert, T. K.; Arszelewska, V. A.; Wang, C.; Yu, C.; Vasileiadis, A.; de Klerk, N. J. J.; Hageman, J.; Hupfer, T.; Kerkamm, I.; Xu, Y.; van der Maas, E.; Kelder, E. M.; Ganapathy, S.; Wagemaker, M.; Wagemaker, M. Clarifying the relationship between redox activity and electrochemical stability in solid electrolytes. *Nat. Mater.* **2020**, *19*, 428−435.
(67) Zhu, Y.; He, X.; Mo, Y. Origin of outstanding stability in the lithium solid electrolyte materials: insights from thermodynamic analyses based on first-principles calculations. *ACS Appl. Mater. Interfaces* **2015**, *7*, 23685−23693.
(68) Xiao, Y.; Wang, Y.; Bo, S.-H.; Kim, J. C.; Miara, L. J.; Ceder, G. Understanding interface stability in solid-state batteries. *Nat. Rev. Mater.* **2020**, *5*, 105−126.
(69) Mangani, L. R.; Villevieille, C. Mechanical vs. chemical stability of sulphide-based solid-state batteries. Which one is the biggest challenge to tackle? Overview of solid-state batteries and hybrid solid state batteries. *J. Mater. Chem. A* **2020**, *8*, 10150−10167.
(70) Opalka, D.; Scheurer, C.; Reuter, K. Ab initio thermodynamics insight into the structural evolution of working IrO2 catalysts in proton-exchange membrane electrolyzers. *ACS Catal.* **2019**, *9*, 4944−4950.

https://doi.org/10.1021/acsaem.1c02363
ACS Appl. Energy Mater. 2021, 4, 12562−12569

Kernel charge equilibration: efficient and accurate prediction of molecular dipole moments with a machine-learning enhanced electron density model

IOP Publishing Mach. Learn.: Sci. Technol. 3 (2022) 015032 https://doi.org/10.1088/2632-2153/ac568d

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
24 November 2021

REVISED
14 February 2022

ACCEPTED FOR PUBLICATION
18 February 2022

PUBLISHED
11 March 2022

PAPER

Kernel charge equilibration: efficient and accurate prediction of molecular dipole moments with a machine-learning enhanced electron density model

Carsten G Staacke[1], Simon Wengert[1], Christian Kunkel[1], Gábor Csányi[2], Karsten Reuter[1] and Johannes T Margraf[1,*]

[1] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany
[2] Engineering Laboratory, University of Cambridge, Cambridge, CB2 1PZ, United Kingdom
[*] Author to whom any correspondence should be addressed.

E-mail: margraf@fhi.mpg.de

Keywords: kernel, charge equilibration, partial charges, dipole moment

Supplementary material for this article is available online

Abstract

State-of-the-art machine learning (ML) interatomic potentials use local representations of atomic environments to ensure linear scaling and size-extensivity. This implies a neglect of long-range interactions, most prominently related to electrostatics. To overcome this limitation, we herein present a ML framework for predicting charge distributions and their interactions termed kernel charge equilibration (kQEq). This model is based on classical charge equilibration (QEq) models expanded with an environment-dependent electronegativity. In contrast to previously reported neural network models with a similar concept, kQEq takes advantage of the linearity of both QEq and Kernel Ridge Regression to obtain a closed-form linear algebra expression for training the models. Furthermore, we avoid the ambiguity of charge partitioning schemes by using dipole moments as reference data. As a first application, we show that kQEq can be used to generate accurate and highly data-efficient models for molecular dipole moments.

1. Introduction

Kernel and neural network (NN) based machine-learning (ML) methods have in recent years become established as an essential addition to the toolbox of computational chemistry [1–4]. In particular, ML-based interatomic potentials have had great success in providing energies and forces with quantum mechanical accuracy at a fraction of the cost of first-principles calculations [1, 5–10]. To achieve size-extensivity and a linear computational scaling with system size, these ML potentials typically rely on a local representation of atomic environments and consequently assume that the energy can be decomposed into local atomic contributions [9, 11]. This simple idea has led to a strong focus of chemical ML research on developing sophisticated representations of local atomic environments and, relatedly, NN architectures that directly embed atoms in their neighborhood [11–14, 14–18].

At the same time, it is clear that the assumption of locality does not hold for all systems to the same extent [19]. Indeed, strongly polar or ionic systems display very long-ranged Coulomb interactions. Even for a fairly unpolar (e.g. organic) system, the locality of the energy does not necessarily imply that other electronic properties are similarly local. In particular, electronic properties such as molecular orbital energies or dipole moments can break locality assumptions [20]. Consequently, such properties tend to be more challenging to predict with purely data-driven ML methods [16, 20]. Beyond this methodological challenge, dipole moments are an intrinsically interesting target as they govern the asymptotic decay of interactions of neutral molecules and their absorption cross-sections in vibrational spectroscopy.

A promising route to overcome the limitations of local ML models is to include known physical interactions explicitly [19, 21, 22]. For example, a description of long-range electrostatics can be obtained by

learning atom-centered charge distributions (e.g. atomic charges, dipoles or partitioned electron densities) [19, 23–28]. A prominent recent example of this is the MuML dipole model of Veit *et al*, which uses atomic charges and atom-centered dipoles to predict molecular dipole moments [20]. This idea takes advantage of the fact that the charge distributions around atoms can be predicted with reasonable accuracy from local environments, even if their interactions are long-ranged. While this solves some of the issues of local interatomic potentials, there are also significant downsides: firstly, charge conservation of the overall system is generally not ensured, and secondly non-local charge transfer (e.g. through conjugated π-systems) is not captured [29].

These issues can be fundamentally addressed by switching the target of the ML model: instead of predicting the charge distribution directly, one can predict a charge-dependent energy expression. The charge distribution is then obtained by minimizing this energy expression under the constraint that the charge is conserved. This idea is closely related to classical charge equilibration approaches like QEq [30]. In this manner, charge conservation is rigorously ensured, the description of non-local charge transfer is enabled and a simple route to analytical derivatives is provided through a Hellmann-Feynman-like approach. The advantage of this approach, compared to directly predicting the charge density, can perhaps be understood in analogy to the choice of initial guess in density functional theory (DFT) calculations: while it is common practice to construct the initial guess from a superposition of atomic densities, it has been found that the superposition of atomic potentials yields a significantly improved starting point [31].

So far only few examples of ML-based charge equilibration models have been reported, however. Most notably, Goedecker and co-workers applied a NN-based QEq model to ionic crystals [32, 33]. The corresponding models were trained to predict the energies and forces of reference DFT calculations, using the corresponding partial charges merely as an intermediate quantity. More recently, Behler, Goedecker and co-workers combined this approach with local NN potentials for the description of organic molecules and MgO surfaces [29, 34]. Here, the charge equilibration models were trained on partial charges from reference DFT calculations. Finally, Xie, Persson and Small applied a more flexible charge-dependent NN to describe lithium hydride nanoparticles, using training data from constrained DFT calculations [35].

Herein, we present a new kernel-based approach to charge equilibration termed kernel charge equilibration (kQEq). These models are directly trained on molecular dipole moments and thus avoid the ambiguity associated with choosing population analysis or projection approaches required in other methods. A closed-form linear algebra expression for training kQEq models is derived and their accuracy is benchmarked on the prediction of molecular dipole moments. Finally, limitations and possible extensions are discussed.

2. Theory

2.1. Charge equilibration

Different conventional (i.e. non-ML) charge equilibration and electronegativity equalization methods have been proposed in the literature [30, 36–45]. In the derivation of the charge equilibration approach we largely follow the formalism of Goedecker and coworkers [29, 32–34], which is in turn based on the QEq method of Rappé and Goddard [30, 46]. In this context, QEq can be understood as a kind of semi-empirical, orbital-free DFT, where the electron density $\rho(\mathbf{r})$ is expanded as:

$$\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \delta\rho(\mathbf{r}), \tag{1}$$

where, $\rho_0(\mathbf{r})$ is a reference density (here the superposition of isolated atom densities) and $\delta\rho(\mathbf{r})$ is a fluctuation term, which describes charge transfer and polarization in the interacting system. We expand $\delta\rho$ into a linear combination of normalized 1 *s* Gaussians centered at the atomic positions $\mathbf{r}_\mathrm{A}$ and of width α_A

$$\delta\rho(\mathbf{r}) \approx \sum_{A=1}^{N} -q_A \frac{1}{\pi^{\frac{3}{2}}\alpha_A^3} \exp\left(\frac{|\mathbf{r}-\mathbf{r_A}|^2}{\alpha_A^2}\right), \tag{2}$$

where N is the number of atoms and q_A are the expansion coefficients. Note that we use the negative of the expansion coefficients q_A here, so that these can directly be interpreted as atomic partial charges. With this approximation, the electron density is completely defined via the charges q_A and it remains to find their optimal values.

To this end, a simple form of the charge-dependent electrostatic energy is assumed:

$$E[\rho] = E_0 + \underbrace{\sum_{A=1}^{N}\left(\chi_A q_A + \tfrac{1}{2}\eta_A q_A^2\right)}_{\mathrm{Site-Energy}} + \underbrace{\tfrac{1}{2}\iint \frac{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'}_{\mathrm{Coulomb-Integral}}. \tag{3}$$

Here, E_0 is a charge-independent reference energy, which we set to zero throughout. The second term (labeled 'Site-Energy') is the well-known second-order Taylor expansion of the atomic energy with respect to the charge, with the atomic electronegativity χ_A and the hardness η_A [47]. It provides the energetic contribution incurred by adding or removing electron density from a given atom. The third term (labeled 'Coulomb-Integral') is the classical Coulomb potential of the fluctuation density $\delta\rho$. This integral can be computed analytically, using the definition of $\delta\rho$ (see equation (2)):

$$\iint \frac{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' = \sum_{A=1}^{N}\left(\frac{2\gamma_{AA}}{\sqrt{\pi}}q_A^2 + \sum_{B=1}^{N} q_A q_B \frac{\mathrm{erf}(\gamma_{AB}r_{AB})}{r_{AB}}\right), \tag{4}$$

with $\gamma_{AB} = \sqrt{(\alpha_A^2+\alpha_B^2)}^{-1}$.

This allows rewriting equation (3) to:

$$E[\rho] = E_0 + \sum_{A=1}^{N}\left[\chi_A q_A + \tfrac{1}{2}\left(\eta_A + \frac{2\gamma_{AA}}{\sqrt{\pi}}\right)q_A^2\right] + \tfrac{1}{2}\sum_{A,B}^{N} q_A q_B \frac{\mathrm{erf}(\gamma_{AB}r_{AB})}{r_{AB}}, \tag{5}$$

which makes it explicit that $E[\rho]$ only depends on the charges q_A. We may therefore equivalently use the notation $E(q_1,\ldots,q_N)$. Note that this equation has the familiar form of the original QEq formulation, with the slight difference that the hardness parameter in QEq implicitly includes the electrostatic idempotential $\frac{2\gamma_{AA}}{\sqrt{\pi}}$, whereas here η_A only describes the non-classical (e.g. exchange-correlation) contributions to the hardness.

Given the definitions of ρ and $E[\rho]$, we now search for the density that minimizes the energy functional under the constraint that the total number of electrons is conserved. From the definition of $\delta\rho$, it can be seen that this is equivalent to the constraint that $\sum_A q_A = Q_{\mathrm{tot}}$, with the total system charge Q_{tot}. This can be achieved by defining the Lagrangian

$$\mathcal{L}(q_1,\ldots,q_N,\lambda) = E(q_1,\ldots,q_N) + \lambda\left(\sum_{A=1}^{N} q_A - Q_{\mathrm{tot}}\right). \tag{6}$$

The constrained minimization of the charges can then be performed by setting up a linear system of equations so that:

$$\frac{\partial\mathcal{L}(q_1,\ldots,q_N,\lambda)}{\partial q_A} = \sum_{B=1}^{N} H_{A,B}q_B + \lambda + \chi_A = 0 \tag{7}$$

$$\frac{\partial\mathcal{L}(q_1,\ldots,q_N,\lambda)}{\partial \lambda} = \sum_{A=1}^{N} q_A - Q_{\mathrm{tot}} = 0 \tag{8}$$

with the elements of the hardness matrix $\mathbf{H}$ defined as:

$$H_{A,B} = \begin{cases} \eta_A + \frac{2\gamma_{AA}}{\sqrt{\pi}} & \text{for } A = B, \\ \frac{\mathrm{erf}(\gamma_{AB}r_{AB})}{r_{AB}} & \text{otherwise.} \end{cases} \tag{9}$$

In matrix notation, this linear system can be formulated as:

$$\underbrace{\begin{pmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,N} & 1 \\ H_{2,1} & H_{2,2} & \cdots & H_{2,N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ H_{N,1} & H_{N,2} & \cdots & H_{N,N} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}}_{\bar{\mathbf{H}}} \cdot \underbrace{\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \\ \lambda \end{pmatrix}}_{\bar{\mathbf{q}}} = -\underbrace{\begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_N \\ -Q_{\mathrm{tot}} \end{pmatrix}}_{\bar{\mathbf{c}}} \tag{10}$$

with the hardness matrix $\bar{\mathbf{H}}$, the charge vector $\bar{\mathbf{q}}$ and the electronegativity vector $\bar{\mathbf{c}}$. Here, the bar-notation is used to indicate that these arrays are extended by one dimension due to the Lagrange multiplier. The corresponding N-dimensional sub-arrays are indicated by $\mathbf{H}$, $\mathbf{q}$ and $\mathbf{c}$ in the following.

2.2. Kernel charge equilibration

The QEq approach as outlined above only requires three parameters, namely the electronegativity (χ_A), the non-classical contribution to the hardness (η_A) and the atomic size (α_A) for each species in the system. As a flipside of this elegant simplicity, the accuracy and transferability of the QEq method is limited, however. In the kQEq method proposed herein, we follow the basic idea of Goedecker and coworkers to overcome this limitation [32]. This is achieved by allowing the electronegativity of an atom to change as a function of its chemical environment. Importantly, taking advantage of the fact that both QEq and kernel ridge regression (KRR) are formulated as linear problems, we obtain a closed-form expression for training these models.

The environment-dependent electronegativity in kQEq is defined via a kernel regression Ansatz as:

$$\chi_A(\mathbf{p}_A) = \sum_{B=1}^{N_{\text{train}}} w_B k(\mathbf{p}_A, \mathbf{p}_B), \tag{11}$$

where k is a kernel function, $\mathbf{p}_A$ is a representation vector that encodes the chemical environment of atom A, w_B is a regression coefficient and N_{train} is the number of atoms in the training set. We use the SOAP kernel and representation vector, which are widely used in the construction of interatomic potentials and as descriptors of local environments [18]. We refer to the original literature and a recent review for a detailed account of the corresponding theory and implementation [1, 11, 48]. In general, the kernel function measures the similarity of chemical environments and is defined as:

$$k(\mathbf{p}, \mathbf{p}') = (\mathbf{p} \cdot \mathbf{p}')^2. \tag{12}$$

To derive an expression for the regression coefficients w, we begin by noting that the prediction of charges with QEq can be expressed as a matrix multiplication of the matrix $\mathbf{A}$ with the vector of electronegativities $\mathbf{c}$:

$$\mathbf{q} = -\mathbf{A}\mathbf{c} \tag{13}$$

where $\mathbf{A}$ is the N-dimensional submatrix of $\bar{\mathbf{H}}^{-1}$. Using equation (11), $\mathbf{c}$ can also be written in terms of a matrix-vector multiplication:

$$\mathbf{c} = \mathbf{K}\mathbf{w} \tag{14}$$

so that

$$\mathbf{q} = -\mathbf{A}\mathbf{K}\mathbf{w}. \tag{15}$$

Here, $\mathbf{K}$ is the kernel matrix (quantifying the similarity between the atoms in the system of interest and the atoms in the training set) and $\mathbf{w}$ is the vector of regression coefficients. There are in principle several options for defining the 'optimal' regression coefficients. One could, e.g. fit them to partial charges obtained from some partitioning of the DFT density. However, the choice of a partial charge model is necessarily somewhat arbitrary and does not guarantee an accurate description of electrostatic interactions. We therefore instead use molecular dipole moments $\boldsymbol{\mu}$ as a reference, which are unambiguously defined physical observables for finite systems and define the leading order term of molecular interactions in the long-range limit.

In a first-principles calculation, the dipole moment is calculated as:

$$\boldsymbol{\mu} = \int \mathbf{r} n(\mathbf{r}) d\mathbf{r}, \tag{16}$$

where $n(\mathbf{r})$ is the total charge density (obtained from the sum of the electron and nuclear charge distributions). For a charge equilibration model like kQEq, this simplifies to

$$\boldsymbol{\mu} = \sum_{A=1}^{N} \mathbf{r}_A q_A = \mathbf{R}\mathbf{q}, \tag{17}$$

using the $3 \times N$ matrix $\mathbf{R}$ with columns $\mathbf{r}_A$. Note that for an unambiguous definition of $\boldsymbol{\mu}$ for charged systems, center-of-mass shifted coordinates are used throughout.

Combining equations (17) and (15), we obtain:

$$\boldsymbol{\mu} = -\mathbf{R}(\mathbf{A}\mathbf{K}\mathbf{w}), \tag{18}$$

for the kQEq dipole moment. To determine the regression weights w, we then set up a regularized least-squares problem with the loss-function:

$$L = ||(\boldsymbol{\mu} - \boldsymbol{\mu}_{\text{ref}})||^2 + \sigma \mathbf{w}^T \mathbf{K} \mathbf{w} = || - \mathbf{R}(\mathbf{A}\mathbf{K}\mathbf{w}) - \boldsymbol{\mu}_{\text{ref}}||^2 + \sigma \mathbf{w}^T \mathbf{K} \mathbf{w}, \tag{19}$$

where σ is a regularization hyperparameter and the term $\mathbf{w}^T\mathbf{K}\mathbf{w}$ comes from the use of Tikhonov regularization in a kernel regression model. Note that we use the simplest form of regularization, with a single parameter σ to control for overfitting. In principle different regularization strengths could be used for each training point (e.g. proportional to the dipole magnitude).

In a final step, we set $\frac{dL}{dw} = 0$ to minimize L and solve for $\mathbf{w}$, to obtain:

$$\mathbf{w} = -(\mathbf{A}^T\mathbf{R}^T\mathbf{R}\mathbf{A}\mathbf{K} + \sigma \mathbb{1})^{-1}\mathbf{A}^T\mathbf{R}^T\boldsymbol{\mu}_{\mathbf{ref}}. \tag{20}$$

The above equations are formulated for a single kQEq problem (i.e. a single molecule or system). In practice we train on multiple systems at once. This can still be achieved with a single linear algebra equation by using blocked matrices for **A** and **R**, and by concatenating the dipole vector elements of all training systems into a single vector.

It should be noted that dipole moments in principle do not contain sufficient information to obtain a unique set of atomic partial charges. One advantage of the kQEq framework is therefore that it offers a natural way to enforce physical constraints on the partial charges. These constraints come in two forms, namely that atoms with similar chemical environments must display similar electronegativities (enforced via the Kernel and regularization) and that the charges must minimize the kQEq energy expression and sum to the total charge (enforced via the QEq framework). Importantly, the set of charges that minimizes a given kQEq energy expression is unique.

2.3. Hyperparameters

Up to now, an environment-dependent description of the atomic electronegativity χ_A is defined, which can be learned from data. It remains to specify the non-classical contribution to the atomic hardness η_A and the atomic radius α_A for each element. Herein, we choose these by very simple heuristics: α_A is set to be proportional to the original QEq radius of the element in question. These radii are tabulated for all elements up to Lawrencium ($Z = 103$) [30]. Empirically, we found that scaling these radii with single global scaling parameter $s_{\text{at}} = 0.75$ yields satisfactory results. Similarly, the non-classical hardness parameter η_A is set to zero throughout, as we found that this yields robust models while keeping the empiricism of the method as low as possible. These choices are quite simplistic and further optimization would certainly be possible. As shown below, already these simple defaults provide highly accurate results for the investigated systems though.

The main hyperparameters to be considered for SOAP are the cutoff radius r_{cut} within which the neighborhood is expanded and the broadness of the Gaussians used to smear out the atomic positions (σ_{atom}). The choice of these lengthscales governs the range in which the environment of an atom affects its electronegativity and how sensitive it is to changes of the atomic positions. In the following, we keep the ratio between these parameters constant ($\sigma_{\text{atom}} = \frac{r_{\text{cut}}}{8}$) so that for larger cutoffs, the atomic positions are smeared out more strongly. The idea of keeping this ratio fixed is based on the fact that the expressiveness of a given atom-centered basis set is limited by the number of basis functions, meaning that it can either provide a high-resolution picture that is short-ranged or a lower-resolution picture that is longer-ranged. Alternatively, one could increase the number of basis functions for larger cutoff radii instead, but this would lead to a significantly increased computational cost. The particular constant of proportionality we use was found to work well empirically. Note that the `Dscribe` implementation of SOAP is used [48]. Full details are provided in the supplementary information (available online at stacks.iop.org/MLST/3/015032/mmedia). The influence of the cutoff parameter is discussed below.

2.4. Error metrics

To quantify the performance of the kQEq models for predicting molecular dipole moments, we use two complementary metrics. On one hand, we use the mean absolute error (MAE) of predicted absolute dipole moments. This is a common measure of accuracy, which allows direct comparison with previous models. Additionally, we use the root mean squared regularized relative error (RRMSE) as used by Hait and Head-Gordon in [49]. This metric is defined as $\frac{|\mu - \mu_{\text{ref}}|}{\max(\mu_{\text{ref}}, 1\ \text{D})} \times 100\%$, with an arbitrary threshold of 1 Debye that discriminates between small and large dipole moments. In this way, a seamless transition from absolute error (for small dipoles) to relative error (for large dipoles) is achieved, which is necessary since the pure relative error is otherwise severely distorted toward systems with small dipoles.

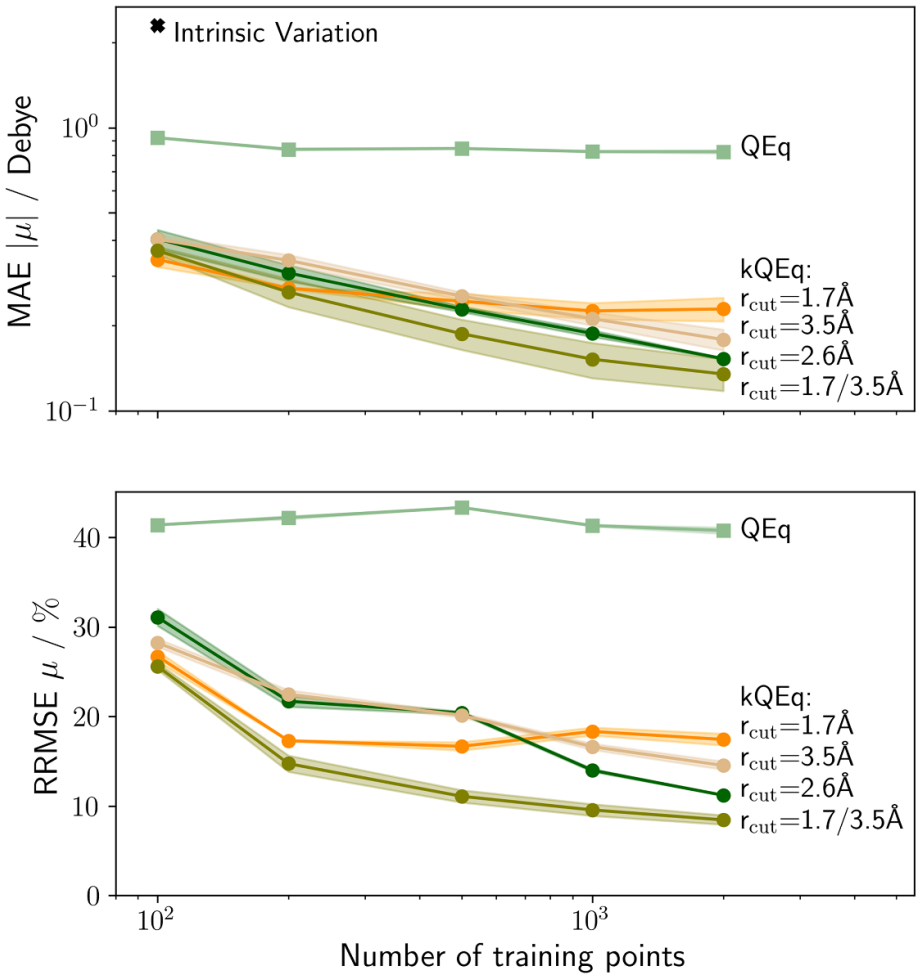


Figure 1. Learning curves for QEq and kQEq models trained on PBE0 dipole moments for the QM9 dataset. Top: MAEs of four different kQEq models compared to QEq. Three different randomized training sets are used and averaged for each model. Bottom: regularized relative root mean square error (RRMSE) of the same models.

3. Results

3.1. Molecular dipole moments

As a first benchmark we trained kQEq models for predicting dipole moments of organic molecules. As reference data, the dipole moments of 7500 random molecules from the QM9 database were calculated at the PBE0/def2-TZVP level, using ORCA (data provided in the SI) [50–52]. This set spans a wide range of small to medium sized molecules containing the elements C, H, N, O and F. From these structures, we randomly selected a validation set (used to optimize the regularization parameter σ) and a test set of 1000 molecules each. The training sets used below were drawn from the remaining 4000 molecules. Figure 1 depicts the learning curves of different kQEq models, using a range of SOAP cutoffs. For comparison, we also fitted a conventional QEq model to the same data.

All kQEq models clearly outperform the conventional QEq approach, underscoring the benefit of the additional flexibility obtained by using environment-dependent electronegativities. Furthermore, it can be seen that the kQEq models improve continuously when given more data, whereas the MAE of the conventional approach quickly saturates. The model with the smallest SOAP cutoff used here (1.7 Å) shows the best performance for small training sets but stops improving when training on larger sets. Meanwhile, the larger cutoffs we tested (2.6 and 3.5 Å) continuously improve and reach an excellent accuracy of 0.15 D (compared to an intrinsic standard deviation of ca. 2 D).

Following the experience with interatomic potentials based on SOAP, we further tested a 1.7/3.5 Å 'double SOAP' representation [7, 53]. This combines a short-ranged/high-resolution kernel with a longer-range/low-resolution one. The corresponding model reaches an even better accuracy and displays the most robust performance across different training set sizes. The latter is particularly evident when considering the RRMSE, which shows that this model consistently improves the relative error when

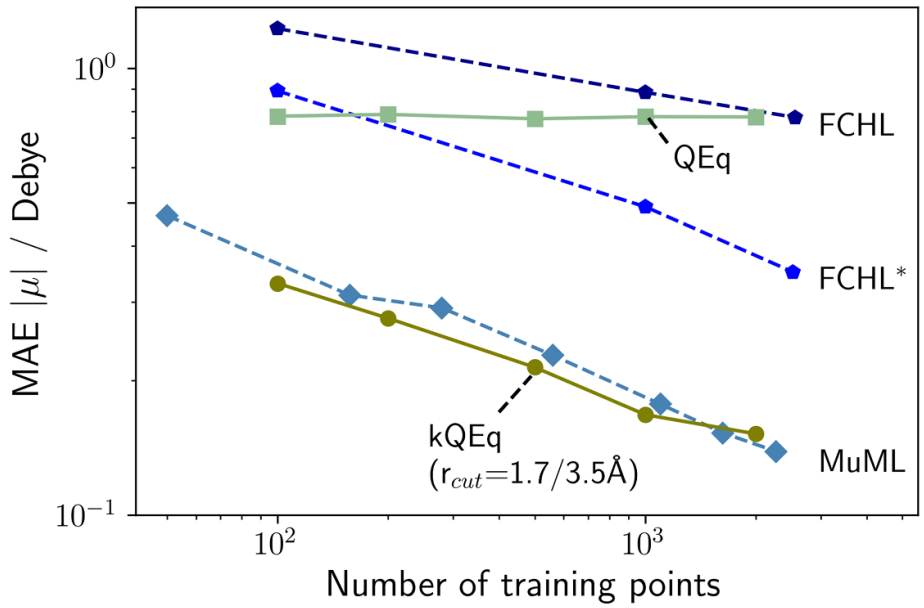


Figure 2. Comparison of literature reported dipole models (dashed lines) and this work (solid). QEq and kQEq models use the same B3LYP training/test data as the MuML model. Reference curves are taken from [20].

increasing the training set, whereas some of the other models improve the performance on total dipole moments at the expense of the relative error (i.e. by describing small dipole moments less accurately).

Overall, these results show that the physical description of long-range contributions in kQEq allows the use of rather small cutoffs for the ML part, effectively focusing on the nearest neighbors. This is beneficial both in terms of transferability of the models and the cost of computing the representation. Notably, the philosophically similar BpopNN model of Xie, Persson and Smalls uses a much larger cutoff radius of 13.2 Å [35]. Meanwhile, the MuML dipole model of Veit *et al* (which is also based on SOAP) uses a cutoff of 5 Å [20]. This indicates that kQEq does a good job of partitioning contributions into long-ranged physical terms and a short-ranged ML model.

To put this performance into perspective, we compare these models to two recent kernel ML models that are specifically tailored to predicting dipole moments, namely the operator ML approach of Christensen, Faber and Lilienfeld and the aforementioned MuML model of Veit *et al* (see figure 2) [20, 54]. The former uses a modified variant of the Faber-Christensen-Huang-Lilienfeld (FCHL) representation (FCHL*), which can incorporate the response of the ML model to applied electric fields and thus provides a physically rigorous and equivariant route to predicting dipole moments. Meanwhile, the latter uses a decomposition of the total molecular dipole into atomic monopole and dipole contributions, using the equivariant λ-SOAP approach [55]. For reference we also include the learning curve of a naive FCHL model, which simply predicts the total dipole moment as a scalar (taken from [54]).

As already discussed in [54], the FCHL* model is a significant improvement over the scalar approach. It also significantly outperforms conventional QEq for all but the smallest training sets. Meanwhile the MuML and kQEq models display remarkably similar learning curves and represent a further improvement over FCHL*. Note that for consistency the QEq and kQEq models in figure 1 were trained on the same reference data (calculated at the B3LYP level) as the MuML model from [20].

It is also instructive to consider the accuracy of the reference methods themselves. As mentioned above, Hait and Head-Gordon used the RRMSE to benchmark density functional methods against high-level Coupled Cluster [49]. According to this metric, the best kQEq model reaches an accuracy of 8.1% on the dipole moments of QM9. This can be compared with the reported errors of popular hybrid functionals like PBE0 (5.2%) and B3LYP (7%). However, it should be noted that the benchmark in [49] focuses on very small molecules and includes spin-polarized systems, so that this comparison should not be overinterpreted. Nonetheless, it indicates that kQEq models approach hybrid DFT accuracy at a much reduced cost.

3.2. Charge analysis

Having established the high accuracy of the kQEq predicted dipole moments, we next turn to the predicted charges themselves. It is well known that the electron density cannot be unambiguously partitioned into atomic partial charges. Consequently, there is no way to objectively measure the quality of such partitionings.

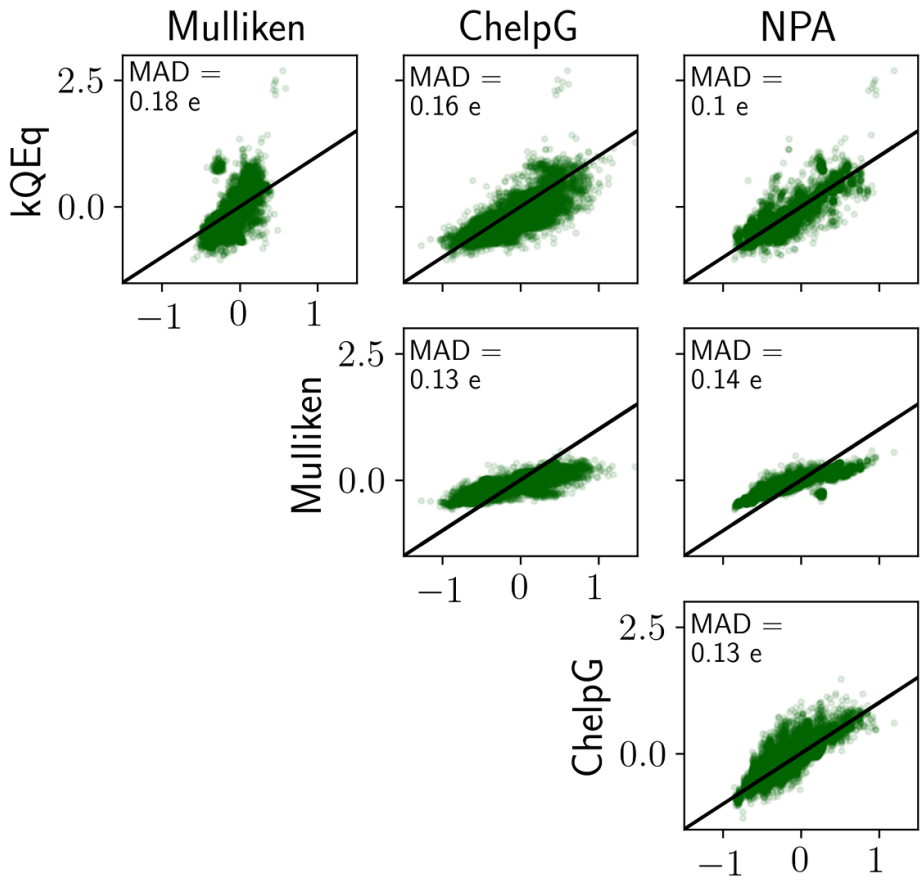


Figure 3. Correlations between all pairs of charge models. The mean absolute deviation (MAD) between each pairing is given in the inset.

Indeed, there is a fundamental tension between describing the local charge density around each atom versus global properties such as dipole moments or electrostatic potentials, when approximating a continuous charge density with atom-centered spherically symmetric charges. Nonetheless, it is worth considering whether the predicted charges are reasonably intuitive and how they compare to standard population analysis schemes like the one of Mulliken [56], restricted electrostatic potential fits like the ChelpG scheme [57], or localized molecular orbital approaches like the Natural Population Analysis [58, 59]. Here, a double SOAP model trained on 2000 training configurations (at the PBE0 level) is used to provide the kQEq charges (see figure 1).

Figure 3 depicts the pairwise correlations between these partial charge models. As has been noted previously, different charge models in general only display moderate agreement with each other [27]. In fact, the lowest mean absolute deviation (0.1 elementary charges) is found between NPA and kQEq, the largest between Mulliken and kQEq. The latter can be attributed to the fact that kQEq charges appear in a broader range (−1–2.5) while Mulliken charges lie between −0.5 and 0.5. Furthermore, kQEq charges display a much more pronounced clustering into functional groups. This is particularly evident for oxygen. While the Mulliken analysis assigns similar charges to ether and carbonyl oxygen atoms, kQEq predicts the carbonyl groups to be significantly more polar. Similarly, the polarity of nitrile and fluoride functional groups is much higher in kQEq (see SI for element-wise correlation plots with all charge models). In contrast, ChelpG and NPA charges mostly fall into the same range as the kQEq ones. In particular, the large differences for polar functional groups (e.g nitrile or carbonyl functional groups) are not observed here. Nonetheless, large deviations can be seen in other cases, particularly for Fluoride functional groups. In general, the relatively good agreement with NPA charges is in line with the observations of Bultinck and coworkers for conventional charge equilibration methods [44].

Two illustrative examples of these differences are shown for 1-(2-Methylcyclopropyl)-ethanone and 2-Fluoropyrazine in figure 4 (both of which are not part of the kQEq training set). At first glance, the partitionings are qualitatively similar, meaning that the signs of most charges match. However, kQEq, NPA, and ChelpG predict significantly larger absolute charges than Mulliken. In the case of 1-(2-Methylcyclopropyl)-ethanone, this is particularly evident for the carbonyl group, which is much more

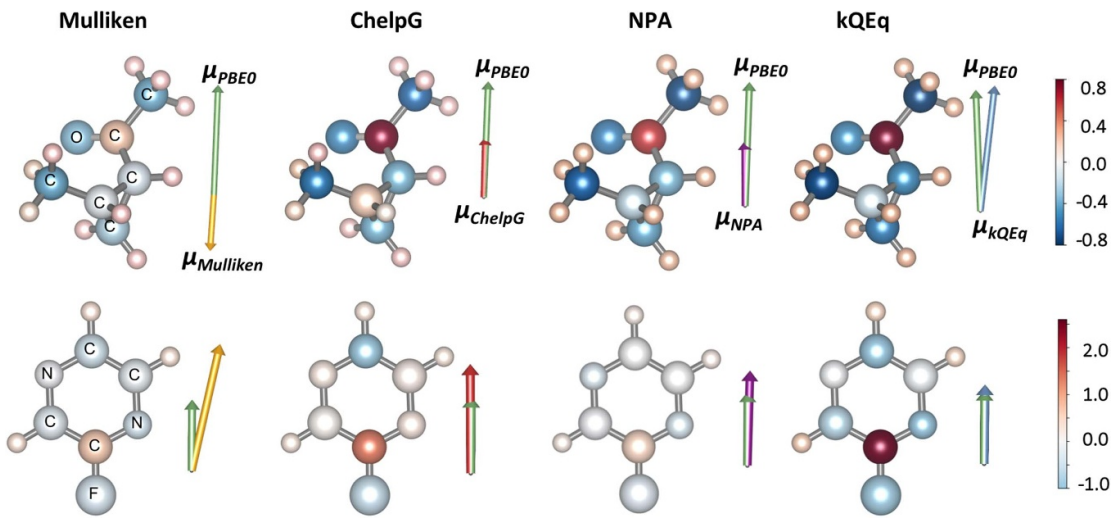


Figure 4. Comparison of Mulliken, ChelpG, NPA and kQEq partial charges for 1-(2-Methylcyclopropyl)-ethanone (top) and 2-Fluoropyrazine (bottom). Reference dipoles (at the PBE0 level) are depicted as arrows (green) alongside the dipoles of the corresponding charge partitioning (Mulliken (yellow), ChelpG (red), NPA(purple), kQEq (blue)).

strongly polarized. Importantly, these differences have strong consequences for the overall description of the molecular electrostatics. Indeed, the dipole moment calculated from the Mulliken charges points in the opposite direction of the actual dipole. In contrast, the ChelpG and NPA dipoles are well aligned with the reference, but too small in magnitude. Finally, the kQEq charges provide an accurate prediction of both the absolute dipole moment and its direction.

The case of 2-Fluoropyrazine is particularly interesting because it illustrates the observed discrepancy between the charges assigned to C-F groups. Here, kQEq predicts the largest absolute charges on the corresponding C and F atoms. Interestingly, the Mulliken, NPA, and ChelpG charges nonetheless overestimate the molecular dipole moment, while the kQEq prediction provides an excellent fit. This is because the large positive charge on the carbon atom stems not only from charge transfer to Fluorine, but also to the adjacent Carbon and Nitrogen atoms, which partially counterbalance the polarity of the C-F group.

The generally poor correlation between the charges obtained with different schemes raises some questions about the different roles and interpretations partial charges can have. On one hand, they can reflect a local partitioning of the electron density, as in the case of Mulliken, Hirshfeld or Bader charges. On the other hand, they can reflect the electrostatic potential on the surface of a molecules, as in the case of ChelpG and related schemes. While the latter is arguably less arbitrary (as it is directly tied to a physical observable) it has well-known issues with assigning meaningful charges to atoms that are 'buried' in bulky molecules.

The kQEq model proposed herein does not neatly fit into these categories. Firstly, it is not a charge partitioning scheme but an ML model. Mulliken, NPA, and ChelpG charges can only be obtained by running a full DFT calculation, whereas the kQEq prediction is much cheaper. Secondly, while kQEq models are trained to reproduce molecular dipole moments, the charges themselves are obtained by minimizing the electrostatic energy expression in equation (5). The Coulomb interactions between partial charges thus provide an important physical constraint on how the charge distribution is approximated. As a consequence, the kQEq derived molecular electrostatic potentials are in good agreement with the ChelpG and NPA ones (see SI). Nonetheless, the kQEq charges are comparatively large. Since bio-organic forcefields usually rely on ChelpG or scaled Mulliken charges, the current kQEq models may thus not be directly applicable in MD simulations. Indeed, the development of interatomic potentials based on kQEq will likely require more complex loss functions, e.g. incorporating energetic contributions or higher order moments. This will be the subject of future work.

3.3. Extrapolating beyond small molecules

A well-known drawback of the conventional QEq approach is that it suffers from an delocalization error akin to that observed for local density functionals. This is e.g. evident in the fact that QEq models incorrectly dissociate molecules into partially charged atoms, since electronegativity differences between isolated atoms lead to spurious charge transfer [39]. While kQEq could in principle cure this particular pathology (by assigning the same electronegativity to all isolated atoms), the more general delocalization tendencies of QEq will likely be inherited by kQEq to an extent. In this section we explore this issue by testing the performance of kQEq for the organic polymer chains dataset introduced by Veit *et al* [20] (see figure 5).

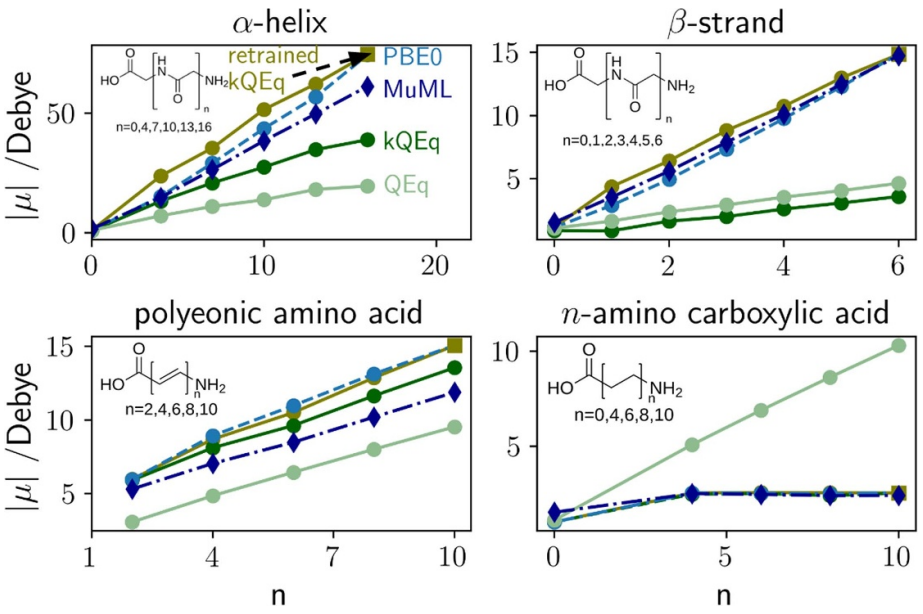


Figure 5. Dipole moment predictions for α-helix and β-strand polyglycine, polyenoic amino acids, and n-amino carboxylic acids [20]. The kQEq and QEq models are trained on 2000 molecules from QM9. The retrained kQEq model additionally includes the largest polymer in each series, highlighted by a square symbol.

This dataset consists of two types of structures. On one hand, glycine polypeptides in the α-helix and β-strand configurations are considered. On the other hand, polyenoic amino acids and n-amino carboxylic acids are included, which consist of a carboxylic acid and an amine group separated by a conjugated double bond or alkane spacer, respectively. Each of these systems shows characteristic changes of the total dipole moment as the polymer length increases. For the polypeptides, each additional amide bond is itself polar, so that the total dipole increases approximately linearly with the system size. However, the precise behavior depends on the spatial orientation of these bond dipoles and their interactions, so that the α-helix and β-strand configurations show different scalings. The polyeonic amino acid chains also display a linearly increasing dipole moment. In this case this is not due to the addition of polar bonds, however, but due to the polarization of the delocalized π-electrons in the spacer. Finally, the dipole of the n-amino carboxylic acids remains constant upon increasing the chain length, since no polar bonds or delocalized electrons are present.

We again use a 'double SOAP' kQEq model trained on 2000 randomly drawn QM9 molecules for comparison. This model has a mixed performance for this test. For polyenonic and n-amino carboxylic acids, the performance is very good. This is both in terms of the qualitative features (linear increase in dipole moments, vs. quick saturation) and the quantitative agreement. Furthermore, kQEq is a strong improvement over conventional QEq here. In contrast, the performance for the polypeptides is less satisfactory. While the linear trend is correctly captured, the magnitude of the dipoles is significantly underestimated, in particular for the β-strand. This behavior is analogous to the underestimation of dipoles typically observed with local functionals, due to the delocalization error [49]. While this shows that kQEq does not automatically cure all pathologies of QEq, a clear advantage of a ML approach is that it can be improved with more data. Indeed, by including the longest chains of each type explicitly in the training set, a retrained kQEq model can be generated that captures these trends more accurately, see figure 5. Potentially, an improved extrapolation behavior could be obtained by using a more flexible expression for the site-energy in equation (3), effectively tackling the problem at its root.

Figure 5 also shows the best MuML model from [20] for comparison. This model performs quite well overall, in particular for the polypeptide systems, with slightly worse performance for the carboxylic acids. In this context, it is worth discussing the fundamental differences between MuML and kQEq. The former uses a fully local decomposition of the overall dipole moment. This works very well for situations where the charge distribution of a large system is essentially just a sum of smaller fragments (as for the polypeptides). In contrast, long-range charge transfer and polarization effects cannot be described by such a model. Specifically, in a system where two functional groups A and B are separated by some spacer, the MuML charge on A will be independent of B. This is reflected in the underestimated dipole moments of longer polyenoic amino acid chains. QEq and kQEq are in principle able to describe such non-local effects. The QEq curve for n-amino carboxylic acid reveals that this is not necessarily an advantage however, as large

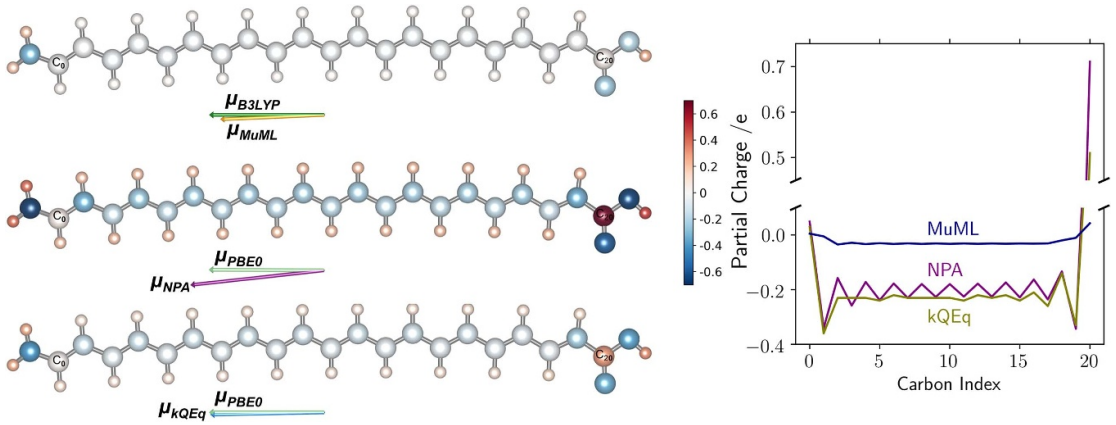


Figure 6. Left: Atomic partial charges for the longest polyenoic amino acid from figure 5, based on MuML (top), NPA (middle) and kQEq (bottom). Right: Comparison of carbon charges in all three models.

unphysical charge transfer is predicted in this case. In other words, a purely local model will generally lead to more systematic and less dramatic failures than a poor non-local one. Fortunately, kQEq and related methods now provide a framework for sophisticated non-local charge equilibration models.

More detailed insight into the charge distributions of the different models can be obtained by comparing the partial charges directly (see figure 6). Here, the MuML and kQEq charges for the longest polyenoic amino acid are shown. For comparison, quantum mechanical NPA charges are also included. This reveals that the kQEq charges display a much better qualitative agreement with NPA. In particular, the oscillatory behavior of the carbon charges (which can be attributed to polarization effects) is completely absent in MuML but clearly observable for kQEq. This is notable, given that both ML models are exclusively trained on dipole moments (i.e. without any atomistic detail) and supports the notion that the kQEq framework introduces useful physical constraints on the charges.

More generally, it should be noted that kQEq is not primarily intended as a stand-alone molecular dipole model. Since it is based on an energy functional, it can be used to model long-range electrostatic interactions in combination with local interatomic potentials. Indeed, it remains an open question whether dipole moments alone provide enough information for this purpose. Fortunately, the current approach can easily be expanded to also include higher moments, electrostatic potentials or reference partial charges. This will be explored in future work.

4. Conclusion

In this work, we introduced kQEq, a kernel-based approach to charge equilibration in molecules. In contrast to conventional charge equilibration methods like QEq, a data-driven, environment-dependent description of atomic electronegativities is introduced. kQEq models trained on molecular dipole moments display excellent performance, *en par* with or better than state-of-the-art kernel models, specifically tuned to predicting dipole moments [20, 54].

The formalism presented herein opens the door toward physics-based kernel ML models for predicting atomic charges, to be used in combination with reactive interatomic potentials. Importantly, the presented approach is quite general and can be extended to other fit targets (e.g. quadrupole moments and electrostatic potentials) and to more flexible density representations (e.g. using atom centered dipoles in addition to partial charges).

While this work focuses on molecular systems, the application to inorganic materials is also envisioned. For finite systems (e.g. nanoparticles) this is in principle straightforward. The use of other fit targets may become important for larger systems, however, as the dipole moment alone might contain too little information in this case. For periodic systems multipole moments are in general ill-defined. Here, the current approach will have to be extended to reproduce other electronic properties.

It should also be noted that QEq-based frameworks are likely not equally well suited for different kinds of materials. In principle, one would expect the best performance for metallic systems, where charges are strongly delocalized and mobile. In contrast, polar insulators or interfaces may be less well described due to the intrinsic delocalization error of charge equilibration models. More general site-energy expressions could be developed to overcome these tendencies. Ideally, such developments will ultimately converge with recent

developments in ML-based DFT, yielding a new generation of orbital-free density functionals [60–63]. Work in this direction is ongoing in our groups.

Data availability statement

All data and a reference implementation of the kQEq method is available at https://gitlab.com/jmargraf/kqeq.

Acknowledgments

The authors thank the DFG for financial support through Germany's Excellence Strategy—EXC 2089/1-390776260. S W and K R further acknowledge funding from Deutsche Forschungsgemeinschaft (DFG) under project RE1509/18-2.

ORCID iD

Johannes T Margraf https://orcid.org/0000-0002-0862-5289

References

[1] Deringer V L *et al* 2021 Gaussian process regression for materials and molecules *Chem. Rev.* **121** 10073–141
[2] Behler J 2021 Four generations of high-dimensional neural network potentials *Chem. Rev.* **121** 10037–72
[3] Unke O T *et al* 2021 Machine learning force fields *Chem. Rev.* **121** 10142–86
[4] Manzhos S and Carrington T J 2020 Neural network potential energy surfaces for small molecules and reactions *Chem. Rev.* **121** 10187–10217
[5] Rupp M, Tkatchenko A, Müller K-R and von Lilienfeld O A 2012 Fast and accurate modeling of molecular atomization energies with machine learning *Phys. Rev. Lett.* **108** 058301
[6] Deringer V L, Caro M A and Csányi G 2019 Machine learning interatomic potentials as emerging tools for materials science *Adv. Mater.* **31** 1902765
[7] Bartók A P *et al* 2017 Machine learning unifies the modeling of materials and molecules *Sci. Adv.* **3** e1701816
[8] Bartók A P, Payne M C, Kondor R and Csányi G 2010 Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons *Phys. Rev. Lett.* **104** 136403
[9] Behler J and Parrinello M 2007 Generalized neural-network representation of high-dimensional potential-energy surfaces *Phys. Rev. Lett.* **98** 146401
[10] Artrith N, Morawietz T and Behler J 2011 High-dimensional neural-network potentials for multicomponent systems: applications to zinc oxide *Phys. Rev.* B **83** 153101
[11] Bartók A P, Kondor R and Csányi G 2013 On representing chemical environments *Phys. Rev.* B **87** 184115
[12] Faber F A, Christensen A S, Huang B and von Lilienfeld O A 2018 Alchemical and structural distribution based representation for universal quantum machine learning *J. Chem. Phys.* **148** 241717
[13] Christensen A S, Bratholm L A, Faber F A and Anatolevon Lilienfeld O 2020 FCHL revisited: faster and more accurate quantum machine learning *J. Chem. Phys.* **152** 044107
[14] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 Quantum-chemical insights from deep tensor neural networks *Nat. Commun.* **8** 13890
[15] Hy T S, Trivedi S, Pan H, Anderson B M and Kondor R 2018 Predicting molecular properties with covariant compositional networks *J. Chem. Phys.* **148** 241745
[16] Klicpera J, Groß J and Günnemann S 2020 Directional message passing for molecular graphs (arXiv:200303123 [Phys. Stat])
[17] Klicpera J, Giri S, Margraf J T and Günnemann S 2020 Fast and uncertainty-aware directional message passing for non-equilibrium molecules (arXiv:201114115 [physics])
[18] Cheng B *et al* 2020 Mapping materials and molecules *Acc. Chem. Res.* **53** 1981–91
[19] Unke O T and Meuwly M 2019 PhysNet: a neural network for predicting energies, forces, dipole moments and partial charges *J. Chem. Theory Comput.* **15** 3678–93
[20] Veit M, Wilkins D M, Yang Y, DiStasio R A J and Ceriotti M 2020 Predicting molecular dipole moments by combining atomic partial charges and atomic dipoles *J. Chem. Phys.* **153** 024113
[21] Yao K, Herr J E, Toth D W, McKintyre R and Parkhill J 2018 The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics *Chem. Sci.* **9** 2261–9
[22] Staacke C G *et al* 2021 On the role of long-range electrostatics in machine-learned interatomic potentials for complex battery materials *ACS Appl. Energy Mater.* **4** 12562–9
[23] Mahmoud C B, Anelli A, Csányi G and Ceriotti M 2020 Learning the electronic density of states in condensed matter *Phys. Rev.* B **102** 235130
[24] Fabrizio A, Grisafi A, Meyer B, Ceriotti M and Corminboeuf C 2019 Electron density learning of non-covalent systems *Chem. Sci.* **10** 9424–32
[25] Zubatyuk R, Smith J S, Leszczynski J and Isayev O 2019 Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network *Sci. Adv.* **5** 10
[26] Gastegger M and Marquetand P 2020 Molecular dynamics with neural network potentials *Machine Learning Meets Quantum Physics (Lecture Notes in Physics)* ed K T Schütt *et al* (Cham: Springer Int. Publishing) pp 233–52
[27] Sifain A E *et al* 2018 Discovering a transferable charge assignment model using machine learning *J. Phys. Chem. Lett.* **9** 4495–501
[28] Nebgen B *et al* 2018 Transferable dynamic molecular charge assignment using deep neural networks *J. Chem. Theory Comput.* **14** 4687–98
[29] Ko T W, Finkler J A, Goedecker S and Behler J 2021 A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer *Nat. Commun.* **12** 398

[30] Rappe A K and Goddard W A 1991 Charge equilibration for molecular dynamics simulations *J. Phys. Chem.* **95** 3358–63
[31] Lehtola S 2019 Assessment of Initial Guesses for self-consistent field calculations. Superposition of atomic potentials: simple yet efficient *J. Chem. Theory Comput.* **15** 1593–604
[32] Ghasemi S A, Hofstetter A, Saha S and Goedecker S 2015 Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network *Phys. Rev.* B **92** 045131
[33] Faraji S *et al* 2017 High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride *Phys. Rev.* B **95** 1–11
[34] Ko T W, Finkler J A, Goedecker S and Behler J 2021 General-purpose machine learning potentials capturing nonlocal charge transfer *Acc. Chem. Res.* **54** 808–17
[35] Xie X, Persson K A and Small D W 2020 Incorporating electronic information into machine learning potential energy surfaces via approaching the ground-state electronic energy as a function of atom-based electronic populations *J. Chem. Theory Comput.* **16** 4256–70
[36] Gasteiger J and Marsili M 1980 Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges *Tetrahedron* **36** 3219–28
[37] Rick S W, Stuart S J and Berne B J 1994 Dynamical fluctuating charge force fields: application to liquid water *J. Chem. Phys.* **101** 6141–56
[38] van Duin A C T, Dasgupta S, Lorant F and Goddard W A 2001 ReaxFF: a reactive force field for hydrocarbons *J. Phys. Chem.* A **105** 9396–409
[39] Morales J and Martínez T J 2001 Classical fluctuating charge theories: the maximum entropy valence bond formalism and relationships to previous models *J. Phys. Chem.* A **105** 2842–50
[40] Morales J and Martínez T J 2004 A new approach to reactive potentials with fluctuating charges: quadratic valence-bond model *J. Phys. Chem.* A **108** 3076–84
[41] Nistor R A, Polihronov J G, Müser M H and Mosey N J 2006 A generalization of the charge equilibration method for nonmetallic materials *J. Chem. Phys.* **125** 094108
[42] Valone S M 2011 Quantum mechanical origins of the Iczkowski-Margrave model of chemical potential *J. Chem. Theory Comput.* **7** 2253–61
[43] Pracht P, Caldeweyher E, Ehlert S and Grimme S 2019 A robust non-self-consistent tight-binding quantum chemistry method for large molecules *ChemRxiv* 1–19
[44] Bultinck P *et al* 2002 The electronegativity equalization method II: applicability of different atomic charge schemes *J. Phys. Chem.* A **106** 7895–901
[45] Bultinck P *et al* 2002 The electronegativity equalization method I: parametrization and validation for atomic charge calculations *J. Phys. Chem.* A **106** 7887–94
[46] Kwon S, Naserifar S, Lee H M and Goddard W A 2018 Polarizable charge equilibration model for transition-metal elements *J. Phys. Chem.* A **122** 9350–8
[47] Parr R G and Pearson R G 1983 Absolute hardness: companion parameter to absolute electronegativity *J. Am. Chem. Soc.* **105** 7512–16
[48] Himanen L *et al* 2019 DScribe: library of descriptors for machine learning in materials science *Comput. Phys. Commun.* **247** 106949
[49] Hait D and Head-Gordon M 2018 How accurate is density functional theory at predicting dipole moments? An assessment using a new database of 200 benchmark values *J. Chem. Theory Comput.* **14** 1969–81
[50] Ramakrishnan R, Dral P O, Rupp M and von Lilienfeld O A 2014 Quantum chemistry structures and properties of 134 kilo molecules *Sci. Data* **1** 1–7
[51] Neese F 2012 The ORCA program system *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2** 73–8
[52] Neese F 2018 Software update: the ORCA program system, version 4.0 *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **8** e1327
[53] Stocker S, Csányi G, Reuter K and Margraf J 2020 Machine learning in chemical reaction space *Nat. Commun.* **11** 5505
[54] Christensen A S, Faber F A and von Lilienfeld O A 2019 Operators in quantum machine learning: response properties in chemical space *J. Chem. Phys.* **150** 064105
[55] Grisafi A, Wilkins D M, Csányi G and Ceriotti M 2018 Symmetry-adapted machine learning for tensorial properties of atomistic systems *Phys. Rev. Lett.* **120** 36002
[56] Mulliken R S 1955 Electronic population analysis on LCAO–MO molecular wave functions. i *J. Chem. Phys.* **23** 1833–40
[57] Breneman C M and Wiberg K B 1990 Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis *J. Comput. Chem.* **11** 361–73
[58] Reed A E, Weinstock R B and Weinhold F 1985 Natural population analysis *J. Chem. Phys.* **83** 735–46
[59] Nikolaienko T Y, Bulavin L A and Hovorun D M 2014 JANPA: an open source cross-platform implementation of the natural population analysis on the Java platform *Comput. Theor. Chem.* **1050** 15–22
[60] Brockherde F *et al* 2017 Bypassing the Kohn-Sham equations with machine learning *Nat. Commun.* **8** 872
[61] Bogojeski M, Vogt-Maranto L, Tuckerman M E, Müller K-R and Burke K 2020 Quantum chemical accuracy from density functional approximations via machine learning *Nat. Commun.* **11** 5223
[62] Dick S and Fernandez-Serra M 2020 Machine learning accurate exchange and correlation functionals of the electronic density *Nat. Commun.* **11** 3509
[63] Margraf J T and Reuter K 2021 Pure non-local machine-learned density functional theory for electron correlation *Nat. Commun.* **12** 344

nanomaterials

MDPI

Article

Tackling Structural Complexity in Li_2S-P_2S_5 Solid-State Electrolytes Using Machine Learning Potentials

Carsten G. Staacke [1,†], Tabea Huss [1,†], Johannes T. Margraf [1], Karsten Reuter [1] and Christoph Scheurer [1,2,*]

[1] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany
[2] Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, Fundamental Electrochemistry (IEK-9), Wilhelm-Johnen-Straße, 52428 Jülich, Germany
* Correspondence: scheurer@fhi.mpg.de
† These authors contributed equally to this work.

Abstract: The lithium thiophosphate (LPS) material class provides promising candidates for solid-state electrolytes (SSEs) in lithium ion batteries due to high lithium ion conductivities, non-critical elements, and low material cost. LPS materials are characterized by complex thiophosphate microchemistry and structural disorder influencing the material performance. To overcome the length and time scale restrictions of *ab initio* calculations to industrially applicable LPS materials, we develop a near-universal machine-learning interatomic potential for the LPS material class. The trained Gaussian Approximation Potential (GAP) can likewise describe crystal and glassy materials and different P-S connectivities P_mS_n. We apply the GAP surrogate model to probe lithium ion conductivity and the influence of thiophosphate subunits on the latter. The materials studied are crystals (modifications of Li_3PS_4 and $Li_7P_3S_{11}$), and glasses of the $x$$Li_2S$–$(100 - x)P_2S_5$ type (x = 67, 70 and 75). The obtained material properties are well aligned with experimental findings and we underscore the role of anion dynamics on lithium ion conductivity in glassy LPS. The GAP surrogate approach allows for a variety of extensions and transferability to other SSEs.

Keywords: machine learning; amorphous; Li-ion battery; high ionic conductivity solid electrolyte

Citation: Staacke, C.G.; Huss, T.; Margraf, J.T.; Reuter, K.; Scheurer, C. Tackling Structural Complexity in Li_2S-P_2S_5 Solid-State Electrolytes Using Machine Learning Potentials. *Nanomaterials* **2022**, *12*, 2950. https://doi.org/10.3390/nano12172950

Academic Editor: Hyun-Suk Kim

Received: 12 July 2022
Accepted: 18 August 2022
Published: 26 August 2022

1. Introduction

While lithium-ion batteries with liquid electrolytes entered the market in 1991, all-solid-state lithium-ion batteries (ASS-LIBs), although investigated for decades, are still not widely in use [1–6]. ASS-LIBs promise several advantages in comparison to liquid electrolyte batteries: higher power density, minimized safety and fire hazards, longer cycle lifetimes, more comprehensive temperature ranges, and enhanced energy density by potential usage of Li metal anodes [1,7,8]. Solid electrolytes of the Li_2S-P_2S_5 material class have gained substantial attention due to their favorable properties [6,9]. First, they possess high conductivities of up to 10^{-2} Scm^{-1}, which ranks them among the most conductive solid electrolytes such as $Li_{10}GeP_2S_{12}$ or $Li_{1.3}Al_{0.3}Ti_{1.7}(PO_4)_3$ [10,11]. Secondly, they are composed of the earth-abundant elements sulfur and phosphorous enabling sustainable applications at large scales.

However, this material class's potential is hampered by the poor understanding of the relevant structure-property relations. This manifests itself in huge deviations in Li-ion conductivity between theory and experiment. As such, β-Li_3PS_4 serves as an illustrating example. Experimental studies report a lithium ion conductivity of approximately 10^{-7} Scm^{-1}, making the material unsuitable for industrial battery applications [12]. In contrast, an *ab initio* study predicts a conductivity of 10^{-1} Scm^{-1}; a six orders of magnitude deviation from experiment that would make the material the new record holder in solid-state lithium ion conduction [13]. Such huge discrepancies often arise from computational limitations that constrain tractable system sizes and sampling times. In the LPS case, high-resolution TEM images for instance revealed the presence of crystalline

nanoparticles in otherwise amorphous regions, highlighting that conductivity calculations of ideal crystals are too short-sighted for this materials class [14]. The problem is further accentuated by the complex chemical structure of LPS [15,16]. A large structural variability at the molecular level, more precisely different thiophosphate poly-anions, are found in all crystalline and amorphous materials [6,17]. For a detailed description of the lithium ion conductivity in LPS we thus need to tackle these challenges: structural complexity of LPS glass compounds, chemical reactivity of thiophosphates, and the influence of anion composition on the lithium ion conductivity.

Here we tackle these challenges by replacing the computationally demanding direct first-principles calculations with a surrogate machine-learning (ML) model. Once trained, this Gaussian Approximation Potential (GAP) model allows for an upscaling of both time and length scale: molecular dynamics (MD) simulations covering up to several nanoseconds and system sizes of several thousand atoms become feasible. Furthermore, the flexibility offered by the ML approach allows one to implement a GAP model that is more versatile and can better represent the crucial complex chemistry than a classical force field [18]. We present a data-efficient iterative training approach to extend an earlier ML force field to yield a near-universal description of the LPS material class [19,20].

In the first part of this work we present our data-efficient training protocol and evaluate the GAP model on (a) its predictive accuracy for lithium ion conductivity and (b) its ability to reproduce two known phase transitions in crystalline Li_3PS_4. The second part focuses on the influence of anion composition on the lithium ion conductivity of different LPS glass compounds.

2. Methods

Computational Details

Reference density-functional theory (DFT) calculations are performed with the PBE functional, default 'light' integration grids and a 'tier 1' basis set of numerical atomic orbitals, as implemented in FHI-aims [21,22]. The Brillouin zone is sampled with a $1 \times 1 \times 1$ k-grid. Initial training configurations are generated with *ab initio* molecular dynamics (MD) using the Γ-point approximation for the k-grid. GAP-based MD and Nudged-Elastic-Band (NEB) simulations are performed using the LAMMPS [23] code and the corresponding interface to QUIP [24,25]. For training set construction and data analysis, the atomic simulation environment ASE, SciPy and scikit-learn are used [26–28].

3. Results

3.1. Lithium Ion Mobility

We obtain the reactive GAP model used to describe the LPS class by fitting to DFT training data computed with the FHI-aims full-potential package [22]. The underlying approach is based on three consecutive steps: defining the anion lattice, sampling of Li-sites, and fine-tuning the materials density. In the first place, only the dominant anion species (e.g., PS_4^{3-} and $P_2S_7^{4-}$) are taken into consideration and utilized in a ratio that represents the desired stoichiometry correctly. For a data-efficient sampling of lithium sites, we sample Li-ion distributions on stable and meta-stable Li sites in a quasi-Monte Carlo like fashion. The materials density is obtained by an iterative compression scheme. Convergence, a detailed step-by-step description of the underlying algorithm, and numerical error assessments of the training procedure are given in Sections A–C in the Supplementary Materials. The benefit of this approach is that it allows the free tuning of stoichiometries and polyanion ratios. In contrast to previous work on crystalline $Li_7P_3S_{11}$ we use a purely short-ranged GAP. In Ref. [19] we combined a GAP model with an electrostatic baseline in order to study the role of long-range electrostatics in machine-learned interatomic potentials for complex battery materials. We previously showed that neglecting long-range electrostatics is unproblematic for describing lithium ion transport in isotropic bulk-like systems [19].

As a first validation of our GAP model, we turn to the Li-ion conductivity of crystalline LPS materials (α, β, γ Li_3PS_4 and $Li_7P_3S_{11}$) at finite temperature, predicted from MD

simulations via the Nernst-Einstein equation (see section F in the Supplementary Materials for details). Using the GAP model we evaluate the ionic conductivity from the mean-square-displacement (MSD) sampled during 2 ns MD simulations for every crystalline compound at various temperatures between 400 and 800 K. Room temperature (RT) conductivities are extrapolated from a linear fit. Note that for crystalline $Li_7P_3S_{11}$ we required longer simulation times of up to 13 ns to reach converged conductivities, i.e., time scales that would essentially be prohibitive for direct *ab initio* MD. While Li-ion conductivity in LPS is usually dominated by diffusion of defects (Li^+ vacancies), $Li_7P_3S_{11}$ exhibits a more collective Li^+ motion yielding the observed high conductivity [19,29,30]. As seen in Figure 1, a broad range of Li-ion conductivities are exhibited in LPS.

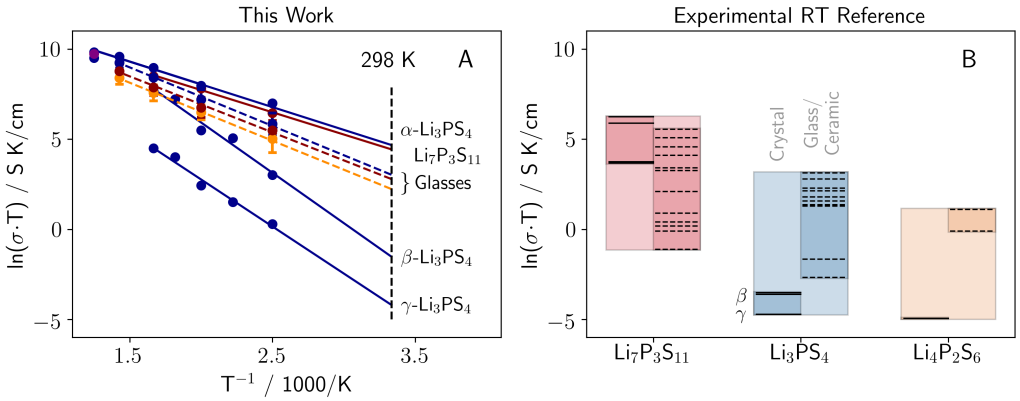


Figure 1. (**A**) Computational Arrhenius plots for $Li_7P_3S_{11}$ (red solid line) and α, β and γ phase of Li_3PS_4 (blue solid lines), as well as the glasses of $Li_4P_2S_6$ (orange dashed line), $Li_7P_3S_{11}$ (red dashed line), and Li_3PS_4 (blue dashed line). (**B**) Reference conductivity data from literature. A tabulated form including references can be found in Table S3 in the Supplementary Materials. Solid lines refer to nominal crystalline materials, dashed lines to glasses/ceramics.

While high RT conductivities of up to 3.6×10^{-3} Scm^{-1} are found for α-Li_3PS_4 and $Li_7P_3S_{11}$, β and γ-Li_3PS_4 exhibit poor RT conductivities of 10^{-5} to 10^{-7} Scm^{-1}. These crystalline RT conductivities are in good agreement with experimental literature, although the extrapolated RT conductivity of β-Li_3PS_4 is somewhat overestimated [30]. For ensemble averaging, we generated 20 structurally uncorrelated glass geometries for each specific temperature and stoichiometry. Hence, each data point in Figure 1 is an average over 20 MD calculations [31]. We consider three different stoichiometries in the analysis that cover the range from fully tetrahedral (Li_3PS_4) via mixed ($Li_7P_3S_{11}$) to fully bridged tetrahedral ($Li_4P_2S_7$) thiophosphate moieties. These three dominant anion subunits are depicted below. As apparent from Figure 1, the ion conductivity over the whole temperature range and the extrapolated RT conductivities increase with growing Li_2S content of the glass material, almost tripling conductivity from $Li_4P_2S_7$ (Li_2S = 67 mol%) to Li_3PS_4 (Li_2S = 75 mol%). Hence, for an increasing Li_2S content an increase RT conductivity is observed. These findings are again in good agreement with experimental studies.

3.2. Li_3PS_4 Phase Transition

As a final validation step, we test the GAP's predictive power on the known phase transitions in Li_3PS_4. As we show in Figure S8 in the Supplementary Materials, the Arrhenius curves of β and γ-Li_3PS_4 exhibit a change of slope at roughly 700 K. Above 700 K, conductivities of β and γ-Li_3PS_4 even match those of α-Li_3PS_4. This change of slope is caused by the phase transition to α-Li_3PS_4, involving a rotation of 25 % of the PS_4^{3-} tetrahedra by 180° for both structures [18].

We can probe the phase transition quantitatively by studying the radial distribution functions (RDFs) of the sulfur sublattice as a function of simulation temperature (Figure 2). The β- and γ-phase share a HCP (hexagonal close-packed) sulfur sublattice, which is

transformed to a BCC (body centered cubic) lattice in the α-phase [18]. For both sublattices, the S-S RDF displays a distinct peak at 3.4 Å, attributed to the intramolecular S-S distance. In the HCP sublattice, a second distinct peak at 4.3 is observed. The latter is missing in the BCC structure. Both β- and γ-phase show the characteristic double-peak in the low-temperature RDF, while the second peak vanishes for temperatures above 650 K. This same phase transition has also been observed in experimental studies and *ab initio* simulations [32,33]. Conceptually, these three phases can be distinguished by their different arrangement of PS_4^{3-}. These are either all pointing in the same direction (γ), are arranged in a zig-zag fashion in one (α) or two directions (β) in space. A visualization can be found in Figure S6 in the Supplementary Materials [32]. The here obtained temperature between 600 and 700 K for the phase transition again matches fairly well with the experimentally reported 746 K [33].

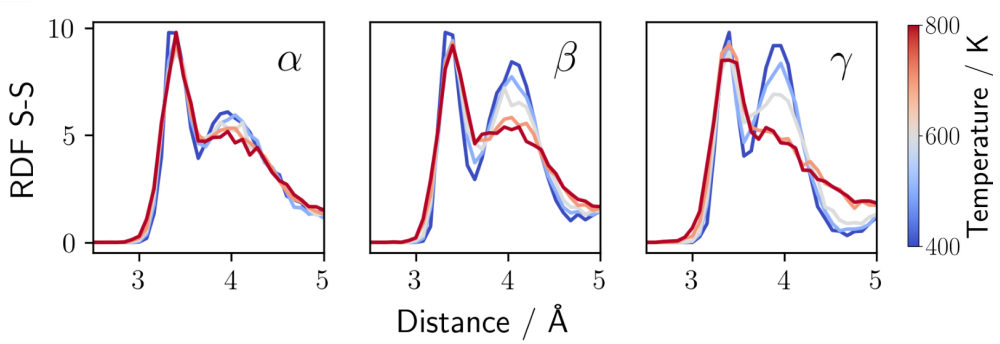


Figure 2. S-S radial distribution functions (RDFs) for MD snapshots of α-Li_3PS_4 (**left** panel), β-Li_3PS_4 (**middle** panel), and γ-Li_3PS_4 (**right** panel) at different temperatures. The disappearance of the peak at 4.3 Å, occurring for β and γ at 700K, corresponds to the phase transition to the α-phase.

3.3. The Role of Anion Composition in Li_2S-P_2S_5 Glasses

Concluding that we can correctly describe the lithium ion dynamics and structural changes in crystalline Li_3PS_4 we now turn to the influence of anion composition on the lithium ion conductivity in LPS glasses. As shown above, the RT conductivity generally increases with Li_2S content of the glass material. The increasing Li-ion conductivity is partly attributable to the lithium mass percentage increase at equal densities. This larger concentration of charge carriers yields higher conductivities for similar diffusion coefficients, accounting for an increase in conductivity of ~30%. As this is much less than the above described rough tripling of the conductivity, we suspect the different anion compositions in the sulfur sub-lattice to be another, dominant factor.

Existing data on the origin of ion conductivity suppression by the anion lattice is quite ambiguous. For example, experimental studies report a strong conductivity suppression by $P_2S_6^{4-}$, attributed to meso-scale precipitation of the non-conducting $Li_2P_2S_6$ phase [34,35]. On the contrary, density of state calculations report that $P_2S_7^{4-}$ should suppress ion conduction at the atomic scale [36]. The charge transfer along the covalent bond between the P and the bridging S lowers the positive partial charge of the P centers, which supposedly attracts Li^+ ions to the $P_2S_7^{4-}$ anions more strongly than the other thiophosphate anions. These are just two illustrative examples discussed as possible origins of ion conductivity suppression by the anion lattice.

First, we analyse the anion composition at different temperatures for all three stoichiometries. Violin plots depicting the building block distributions at different temperatures within the structure ensembles are displayed in Figure 3. For the Li_3PS_4 glass, the simple PS_4^{3-} *ortho*-thiophosphate is as intuitively expected the predominant species over the whole temperature range. *Hypo*-thiodiphosphate $P_2S_6^{4-}$ occurs only in small concentrations $\leq$10 at.% and shows no strong temperature dependence. Up to 25 at.% of *pyro*-thiodiphosphate $P_2S_7^{4-}$ occur at the lower temperature but gradually disappear between 600 and 700 K. These anion ratios are in agreement with experimental ratios found for

Li_3PS_4 [37]. In both, the $Li_7P_3S_{11}$ and $Li_4P_2S_7$ glasses, the $P_2S_6^{4-}$ content instead increases between 400 and 700 K, even though the increase is not too pronounced in comparison to the width of the distribution in the ensemble. The found $P_2S_6^{4-}$ contents in $Li_7P_3S_{11}$ and $Li_4P_2S_7$ are slightly higher compared to experimental data [17].

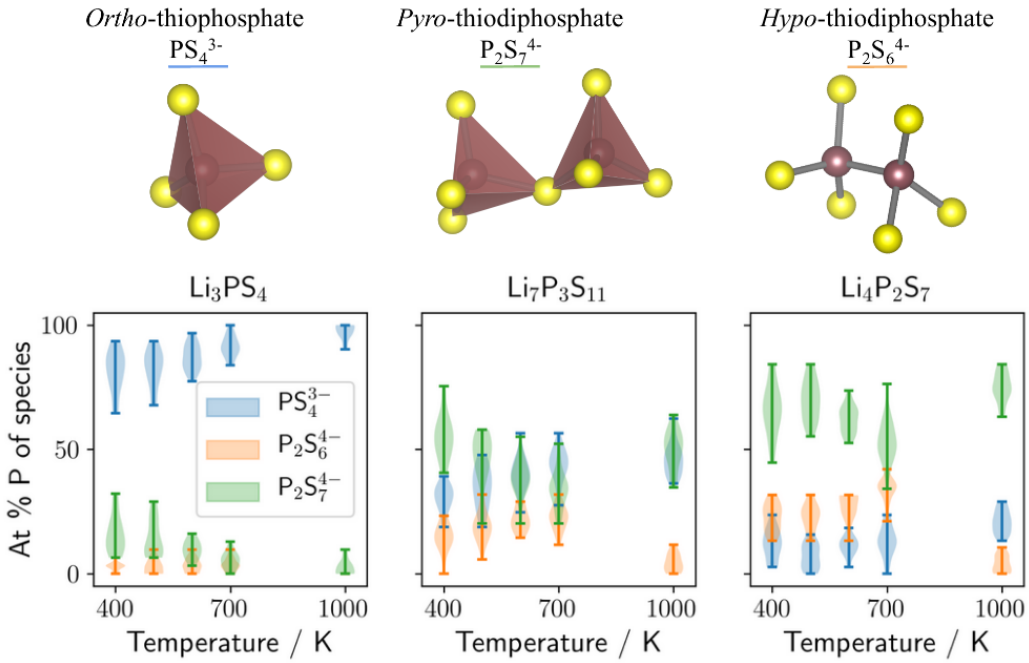


Figure 3. Top: Dominant anions in the Li_2S–P_2S_5 material class. **Bottom**: Anion compositions for different MD temperatures, displayed for Li_3PS_4 (**left** panel), $Li_7P_3S_{11}$ (**middle** panel), and $Li_4P_2S_7$ (**right** panel) glasses.

Next, we analyze the number of Li-positions occupied during MD simulations at finite temperatures, by calculating the isosurface of the probability density distribution of Li-positions (exemplary visualizations see Figure S9 in the Supplementary Materials). When referencing the volume enclosed by the isosurface to the total volume of the simulation cell, we identify the relative accessible volume for all Li-ions for a given stoichiometry. As shown in the left panel of Figure 4 for the $Li_7P_3S_{11}$ and $Li_4P_2S_7$ glasses, the same relative volume is accessed by Li, while Li_3PS_4 exhibits a 10 % higher accessible Li-volume at all temperatures. This is intuitive as $P_2S_n^{4-}$ moieties have a larger surface/volume of the anion, allowing for a smaller number of Li-sites in the material compared to smaller PS_4^{3-} anions.

In order to explore the effect of the anion lattice motion on the Li-ion conductivity we either constrain the sulfur positions, or the phosphorous positions, and compare the Li ion conductivity obtained within MD simulations with these two frozen lattices to the unconstrained Li-ion conductivity. As seen in the right panel of Figure 4, the Li-ion conductivity decreases for all three glass stoichioemetries for both frozen lattices. However, while in the case of frozen phosphorus we observe only a slight decrease in conductivity, freezing the sulfur degrees of freedom reduces the conductivity by approximately two orders of magnitude. We observe the largest decrease of the conductivity for the stoichiometry consisting of the highest PS_4^{3-} (Li_3PS_4) content, and the smallest change for the lowest PS_4^{3-} content ($Li_4P_2S_7$). This suggests that the motion of sulfur throughout the Li-ion conduction plays a significant role.

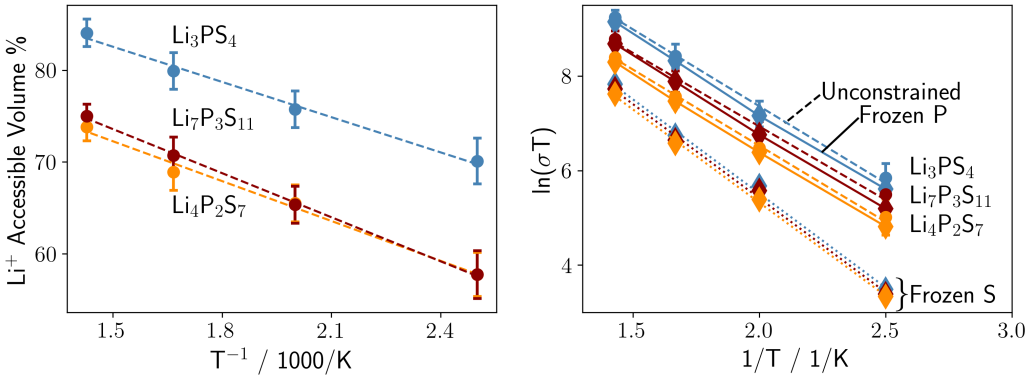


Figure 4. (**Left** panel) Accessible volume of lithium during MD simulations at various temperatures: Li_3PS_4 (blue dashed line), $Li_7P_3S_{11}$ (red dashed line), and $Li_4P_2S_7$ (orange dashed line). (**Right** panel) Li ion conductivity with frozen sulfur lattice (diamonds with dotted lines), with frozen phosphor lattice (diamonds with solid line) and without constraints on the sulfur lattice (dots with dashed line): Li_3PS_4 (blue lines), $Li_7P_3S_{11}$ (red lines), and $Li_4P_2S_7$ (orange lines).

Smith and Siegel showed that in glassy Li_3PS_4, lithium migration occurs via a mechanism that combines a concerted motion of lithium ions with re-orientations of PS_4^{3-} anions [38]. This effect, known as the 'paddlewheel' mechanism, can directly attribute the increasing Li-ion conductivity with increasing PS_4^{3-} content. So far, the paddlewheel effect has only been shown in Li_3PS_4, but our results confirm this effect occurs as long as PS_4^{3-} is present. Hence, the conductivity of $Li_4P_2S_6$ and $Li_4P_2S_7$ decrease as well when the sulfur lattice is frozen, but the effect is not as pronounced as in Li_3PS_4. As the Li-ion conductivity of all three stoichiometries is almost identical when the sulfur lattice is frozen, this actually suggests that the higher accessible volume of Li in Li_3PS_4 arises from re-orientations of PS_4^{3-} anions. Hence, both effects can not be decoupled, but rather the re-orientation of PS_4^{3-} anions generates new Li sites. Together with the increased overall Li content, this thus fully rationalizes why the Li-ion conductivity increases with higher Li_2S content.

4. Conclusions

All of the herein described effects, collective Li-ion motion of crystalline $Li_7P_3S_{11}$, phase transitions of crystalline Li_3PS_4, and the conductivity/anion-composition relation in glassy LPS, could not be studied before by a single interatomic potential, preventing the relative identification of trends and common origins. While not only this can now be achieved by our machine learning surrogate model, the general structure of the training protocol furthermore allows for a variety of extensions, including additional selection criteria [20,39], using an electrostatic baseline in the model [40], doping with transition metals, and modeling of solid/solid interfaces [41,42]. We correspondingly see much prospects in the use of ML potentials to further elucidate atomic scale processes in complex battery materials.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/nano12172950/s1, Figure S1: Anion Voronoi Tesselation, Figure S2: Glass sampling approach, Figures S3 and S4: Sampling of P-P distances, Figures S5 and S6: Force parity plot, Figure S7–S10: Radial distribution functions, Figure S11: Li_3PS_4 crystal configuration, Figure S12: Density dependence of the conductivity, Figure S13: Computational Arrhenius plots, Figure S14: Li accessible volume. Table S1: Technical Hyperparameters for the GAP, Table S2: Coordination-resolved force RMSEs, Table S3: Lithium ion conductivity. References [43–49] are cited in the supplementary materials

Author Contributions: The project has been conceptualized by C.G.S. and K.R. C.G.S. developed the general fitting procedure, T.H. refined the iterative training scheme. C.S. and J.T.M. guided the formal analysis. The manuscript was jointly written and edited by all authors. All authors have read and agreed to the published version of the manuscript.

Funding: Open access funded by Max Planck Society. We acknowledge funding from the German Research Foundation (DFG) through its Cluster of Excellence *e*-conversion EXC 2089/1.

Data Availability Statement: Data and codes can be found under Ref. [31].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LPS	Lithium thiophosphate
SSE	Solid-state electrolytes
GAP	Gaussian Approximation Potential
ASS-LIB	All-solid-state lithium-ion batteries
NEB	Nudged-Elastic-Band
ML	Machine Learning
MD	Molecular Dynamics
DFT	Density-functional theory
MSD	Mean-square-displacement
RDF	Radial distribution function
HCP	Hexagonal close-packed

References

1. Janek, J.; Zeier, W.G. A solid future for battery development. *Nat. Energy* **2016**, *1*, 16141. [CrossRef]
2. Zheng, F.; Kotobuki, M.; Song, S.; Lai, M.O.; Lu, L. Review on solid electrolytes for all-solid-state lithium-ion batteries. *J. Power Sources* **2018**, *389*, 198–213. [CrossRef]
3. Luntz, A.C.; Voss, J.; Reuter, K. Interfacial challenges in solid-state Li ion batteries. *J. Phys. Chem. Lett.* **2015**, *6*, 4599–4604. [CrossRef] [PubMed]
4. Oudenhoven, J.F.; Baggetto, L.; Notten, P.H. All-solid-state lithium-ion microbatteries: A review of various three-dimensional concepts. *Adv. Energy Mater.* **2011**, *1*, 10–33. [CrossRef]
5. Shoji, M.; Cheng, E.J.; Kimura, T.; Kanamura, K. Recent progress for all solid state battery using sulfide and oxide solid electrolytes. *J. Phys. D Appl. Phys.* **2019**, *52*, 103001. [CrossRef]
6. Kudu, Ö.U.; Famprikis, T.; Fleutot, B.; Braida, M.D.; Le Mercier, T.; Islam, M.S.; Masquelier, C. A review of structural properties and synthesis methods of solid electrolyte materials in the Li_2S-P_2S_5 binary system. *J. Power Sources* **2018**, *407*, 31–43. [CrossRef]
7. Robinson, A.L.; Janek, J. Solid-state batteries enter EV fray. *MRS Bull.* **2014**, *39*, 1046–1047. [CrossRef]
8. Kim, K.J.; Balaish, M.; Wadaguchi, M.; Kong, L.; Rupp, J.L. Solid-state Li–metal batteries: Challenges and horizons of oxide and sulfide solid electrolytes and their interfaces. *Adv. Energy Mater.* **2021**, *11*, 2002689. [CrossRef]
9. Guo, H.; Wang, Q.; Urban, A.; Artrith, N. Artificial Intelligence-Aided Mapping of the Structure–Composition–Conductivity Relationships of Glass–Ceramic Lithium Thiophosphate Electrolytes. *Chem. Mater.* **2022**, *34*, 6702–6712. [CrossRef]
10. Kamaya, N.; Homma, K.; Yamakawa, Y.; Hirayama, M.; Kanno, R.; Yonemura, M.; Kamiyama, T.; Kato, Y.; Hama, S.; Kawamoto, K.; et al. A lithium superionic conductor. *Nat. Mater.* **2011**, *10*, 682–686. [CrossRef]
11. Aono, H.; Sugimoto, E.; Sadaoka, Y.; Imanaka, N.; Adachi, G.-y. Ionic Conductivity of Solid Electrolytes Based on Lithium Titanium Phosphate. *J. Electrochem. Soc.* **1990**, *137*, 1023–1027. [CrossRef]
12. Tachez, M.; Malugain, J.; Mercire, R.; Robert, G. Ionic conductivity of and phase transition in lithium thiophosphate Li_3PS_4. *Solid State Ion.* **1984**, *14*, 181–185. [CrossRef]
13. Yang, Y.; Wu, Q.; Cui, Y.; Chen, Y.; Shi, S.; Wang, R.Z.; Yan, H. Elastic Properties, Defect Thermodynamics, Electrochemical Window, Phase Stability, and Li+ Mobility of Li_3PS_4: Insights from First-Principles Calculations. *ACS Appl. Mater. Interfaces* **2016**, *8*, 25229–25242. [CrossRef] [PubMed]
14. Tsukasaki, H.; Mori, S.; Morimoto, H.; Hayashi, A.; Tatsumisago, M. Direct observation of a non-crystalline state of Li2S–P2S5 solid electrolytes. *Sci. Rep.* **2017**, *7*, 4142. [CrossRef]
15. Liu, Z.; Fu, W.; Payzant, E.A.; Yu, X.; Wu, Z.; Dudney, N.J.; Kiggans, J.; Hong, K.; Rondinone, A.J.; Liang, C. Anomalous High Ionic Conductivity of Nanoporous β-Li_3PS_4. *J. Am. Chem. Soc.* **2013**, *135*, 975–978. [CrossRef]
16. Wenzel, S.; Weber, D.A.; Leichtweiss, T.; Busche, M.R.; Sann, J.; Janek, J. Interphase formation and degradation of charge transfer kinetics between a lithium metal anode and highly crystalline $Li_7P_3S_{11}$ solid electrolyte. *Solid State Ion.* **2016**, *286*, 24–33. [CrossRef]
17. Dietrich, C.; Weber, D.A.; Sedlmaier, S.J.; Indris, S.; Culver, S.P.; Walter, D.; Janek, J.; Zeier, W.G. Lithium ion conductivity in Li_2S-P_2S_5 glasses–building units and local structure evolution during the crystallization of superionic conductors Li_3PS_4, $Li_7P_3S_{11}$ and $Li_4P_2S_7$. *J. Mater. Chem. A* **2017**, *5*, 18111–18119. [CrossRef]

18. Kim, J.S.; Jung, W.D.; Choi, S.; Son, J.W.; Kim, B.K.; Lee, J.H.; Kim, H. Thermally Induced S-Sublattice Transition of Li_3PS_4 for Fast Lithium-Ion Conduction. *J. Phys. Chem. Lett.* **2018**, *9*, 5592–5597. [CrossRef]
19. Staacke, C.G.; Heenen, H.H.; Scheurer, C.; Csányi, G.; Reuter, K.; Margraf, J.T. On the Role of Long-Range Electrostatics in Machine-Learned Interatomic Potentials for Complex Battery Materials. *ACS Appl. Energy Mater.* **2021**, *4*, 12562–12569. [CrossRef]
20. Timmermann, J.; Lee, Y.; Staacke, C.G.; Margraf, J.T.; Scheurer, C.; Reuter, K. Data-efficient iterative training of Gaussian approximation potentials: Application to surface structure determination of rutile IrO_2 and RuO_2. *J. Chem. Phys.* **2021**, *155*, 244107. [CrossRef]
21. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865. [CrossRef] [PubMed]
22. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196. [CrossRef]
23. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19. [CrossRef]
24. Bartók, A.P.; Payne, M.C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403. [CrossRef] [PubMed]
25. Bartók, A.P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B* **2013**, *87*, 184115. [CrossRef]
26. Larsen, A.H.; Mortensen, J.J.; Blomqvist, J.; Castelli, I.E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M.N.; Hammer, B.; Hargus, C.; et al. The atomic simulation environment—A Python library for working with atoms. *J. Phys. Condens. Matter* **2017**, *29*, 273002. [CrossRef]
27. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. [CrossRef]
29. Wang, Y.; Richards, W.D.; Ong, S.P.; Miara, L.J.; Kim, J.C.; Mo, Y.; Ceder, G. Design principles for solid-state lithium superionic conductors. *Nat. Mater.* **2015**, *14*, 1026–1031. [CrossRef]
30. Chu, I.H.; Nguyen, H.; Hy, S.; Lin, Y.C.; Wang, Z.; Xu, Z.; Deng, Z.; Meng, Y.S.; Ong, S.P. Insights into the performance limits of the $Li_7P_3S_{11}$ superionic conductor: A combined first-principles and experimental study. *ACS Appl. Mater. Interfaces* **2016**, *8*, 7843–7853. [CrossRef]
31. Scheurer, C. Model Structures for Glass Phases of the Solid-State Electrolyte LPS. 2022. Available online: https://edmond.mpdl.mpg.de/dataset.xhtml?persistentId=doi:10.17617/3.VZHSXS (accessed on 17 August 2022). [CrossRef]
32. Homma, K.; Yonemura, M.; Kobayashi, T.; Nagao, M.; Hirayama, M.; Kanno, R. Crystal structure and phase transitions of the lithium ionic conductor Li_3PS_4. *Solid State Ion.* **2011**, *182*, 53–58. [CrossRef]
33. Phuc, N.H.H.; Totani, M.; Morikawa, K.; Muto, H.; Matsuda, A. Preparation of Li_3PS_4 solid electrolyte using ethyl acetate as synthetic medium. *Solid State Ion.* **2016**, *288*, 240–243. [CrossRef]
34. Busche, M.R.; Weber, D.A.; Schneider, Y.; Dietrich, C.; Wenzel, S.; Leichtweiss, T.; Schröder, D.; Zhang, W.; Weigand, H.; Walter, D.; et al. In Situ Monitoring of Fast Li-Ion Conductor $Li_7P_3S_{11}$ Crystallization Inside a Hot-Press Setup. *Chem. Mater.* **2016**, *28*, 6152–6165. [CrossRef]
35. Hayashi, A.; Minami, K.; Ujiie, S.; Tatsumisago, M. Preparation and ionic conductivity of $Li_7P_3S_{11}$-z glass-ceramic electrolytes. *J. Non. Cryst. Solids* **2010**, *356*, 2670–2673. [CrossRef]
36. Ohara, K.; Mitsui, A.; Mori, M.; Onodera, Y.; Shiotani, S.; Koyama, Y.; Orikasa, Y.; Murakami, M.; Shimoda, K.; Mori, K.; et al. Structural and electronic features of binary Li_2S-P_2S_5 glasses. *Sci. Rep.* **2016**, *6*, 21302. [CrossRef]
37. Dietrich, C.; Weber, D.A.; Culver, S.; Senyshyn, A.; Sedlmaier, S.J.; Indris, S.; Janek, J.; Zeier, W.G. Synthesis, Structural Characterization, and Lithium Ion Conductivity of the Lithium Thiophosphate $Li_2P_2S_6$. *Inorg. Chem.* **2017**, *56*, 6681–6687. [CrossRef]
38. Smith, J.G.; Siegel, D.J. Low-temperature paddlewheel effect in glassy solid electrolytes. *Nat. Comm.* **2020**, *11*, 1483. [CrossRef]
39. Timmermann, J.; Kraushofer, F.; Resch, N.; Li, P.; Wang, Y.; Mao, Z.; Riva, M.; Lee, Y.; Staacke, C.; Schmid, M.; et al. IrO_2 Surface Complexions Identified through Machine Learning and Surface Investigations. *Phys. Rev. Lett.* **2020**, *125*, 206101. [CrossRef]
40. Staacke, C.G.; Wengert, S.; Kunkel, C.; Csányi, G.; Reuter, K.; Margraf, J.T. Kernel charge equilibration: Efficient and accurate prediction of molecular dipole moments with a machine-learning enhanced electron density model. *Mach. Learn. Sci. Technol.* **2022**, *3*, 015032. [CrossRef]
41. Stegmaier, S.; Schierholz, R.; Povstugar, I.; Barthel, J.; Rittmeyer, S.P.; Yu, S.; Wengert, S.; Rostami, S.; Hans, K.; Reuter, K.; et al. Nano-Scale Complexions Facilitate Li Dendrite-Free Operation in LATP Solid-State Electrolyte. *Adv. Energy Mater* **2021**, *11*, 2100707. [CrossRef]
42. Türk, H.; Schmidt, F.P.; Götsch, T.; Girgsdies, F.; Hammud, A.; Ivanov, D.; Vinke, I.C.; de Haart, L.; Eichel, R.A.; Reuter, K.; et al. Complexions at the Electrolyte/Electrode Interface in Solid Oxide Cells. *Adv. Mater. Interfaces* **2021**, *8*, 2100967. [CrossRef]
43. Mehrer, H. *Diffusion in Solids: Fundamentals, Methods, Materials, Diffusion-Controlled Processes*; Springer Science & Business Media: Cham, Switzerland, 2007; Volume 155.
44. Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Elsevier: Alpharetta, GA, USA, 2001; Volume 1.

45. Seino, Y.; Ota, T.; Takada, K.; Hayashi, A.; Tatsumisago, M. A sulphide lithium super ion conductor is superior to liquid ion conductors for use in rechargeable batteries. *Energy Environ. Sci.* **2014**, *2*, 627–631. [CrossRef]
46. de Klerk, N.; van der Maas, E.; Wagemaker, M. Analysis of Diffusion in Solid-State Electrolytes through MD Simulations, Improvement of the Li-Ion Conductivity in β-Li_3PS_4 as an Example. *ACS Appl. Energy Mater.* **2018**, *2*, 3230–3242. [CrossRef] [PubMed]
47. Baba, T.; Kawamura, Y. Structure and ionic conductivity of Li_2S–P_2S_5 glass electrolytes simulated with first-principles molecular dynamics. *Front. Energy Res.* **2016**, *4*, 22. [CrossRef]
48. Shiotani, S.; Ohara, K.; Tsukasaki, H.; ; Mori, S.; Kanno, R. Pair distribution function analysis of sulfide glassy electrolytes for all-solid-state batteries: Understanding the improvement of ionic conductivity under annealing condition. *Sci. Rep.* **2017**, *7*, 6972. [CrossRef]
49. Seino, Y.; Nakagawa, M.; Senga, M.; Higuchi, H.; Takada, K.; Sasaki, T. Analysis of the structure and degree of crystallisation of $70Li_2S$–$30P_2S_5$ glass ceramic. *J. Mater. Chem. A* **2015**, *3*, 2756–2761. [CrossRef]