Multiscale Modelling and Simulation of Advanced Battery Materials

Mauricio R. Bonilla, Fabián A. García Daza, Mario Fernández-Pendás, Javier Carrasco and Elena Akhmatskaya

This is a preprint of the following chapter: M. R. Bonilla, F. García Daza, M. Fernández-Pendás, J. Carrasco, E. Akhmatskaya, Multiscale Modelling and Simulation of Advanced Battery Materials, published in Progress in Industrial Mathematics: Success Stories, edited by M. Cruz, C. Pares, P. Quintela (Eds), Springer, 69-113 (2021), ISBN 978-3-030-61843-8, reproduced with permission of Springer. The final authenticated version is available online at:
$https://link.springer.com/chapter/10.1007/978-3-030-61844-5\_6$

Abstract Development of efficient strategies for the rational design of materials involved in the production and storage of renewable energy is essential for accelerating the transition to a low-carbon economy. To contribute to this goal, we propose a novel workflow for the assessment and optimization of battery materials. The approach effectively combines quantum and atomistic modelling/simulations, enhanced by

---

Mauricio R. Bonilla
Basque Center for Applied Mathematics, Alameda de Mazarredo 14, Bilbao 48009, Spain

Fabián A. García Daza
Department of Chemical Engineering and Analytical Science, The University of Manchester, Sackville Street, Manchester M13 9PL, United Kingdom

Javier Carrasco
Centre for Cooperative Research on Alternative Energies (CIC energiGUNE), Basque Research and Technology Alliance (BRTA), Alava Technology Park, Albert Einstein 48, Vitoria-Gasteiz 01510, Spain

Mario Fernández-Pendás
Basque Center for Applied Mathematics, Alameda de Mazarredo 14, Bilbao 48009, Spain
Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU) and Donostia International Physics Center (DIPC), Donostia 20080, Spain

Elena Akhmatskaya
Basque Center for Applied Mathematics, Alameda de Mazarredo 14, Bilbao 48009, Spain
IKERBASQUE, Basque Foundation for Science, Bilbao 48013, Spain e-mail: `akhmatskaya@bcamath.org`

efficient sampling, Bayesian parameterization, and experimental information. It is implemented to study prospective materials for lithium and sodium batteries.

1 Introduction

Nearly 30 years ago Sony Co. launched the first commercial lithium-ion battery (LIB) and changed the world [25]. LIBs powered the revolution in portable electronics, allowing the transformation of mobile phones into general purpose computers within a decade. Moreover, governments around the globe have gained awareness of the role greenhouse gases play in climate change, leading to incentives for the development of renewable energy technologies (solar, wind, etc.) and electric vehicles (EVs). For all of these technologies to be effectively implemented, energy storage systems are key enablers. As an example, the European Commission has set the target of achieving emission-free urban passenger transportation by 2050 (i.e. abandoning the use of conventionally fuelled cars in cities) and emission-free urban freight transportation by 2030 [15]. Consequently, research and development on LIBs have exploded (in the span of 7 years, researchers around the globe have added at least 119188 new publications on batteries from 2010 to 2017 [54]). A recognition of the current importance of the field comes from the 2019 Nobel Prize in Chemistry, awarded to John Goodenough, M. Stanley Whittingham and Akira Yoshino for the invention of the rechargeable LIB.

While the bulk of battery research is directed towards LIBs, it is important to consider that lithium is not regarded as an abundant element (the relative abundance of lithium in the Earth's crust is only 20 parts per million). Moreover, lithium resources are greatly concentrated in South America, which means that regional politics can have an outsized effect on its price. The second-smallest alkali metal, sodium, constitutes 1% of the earth's crust and is a strong candidate to substitute lithium in rechargeable batteries [77]. For this reason, sodium-ion batteries (NIBs) are one of the most widely investigated alternatives to LIBs, and commercial units are being produced already for systems with low energy density (e.g. e-bikes, e-scooters) [11]. For most applications, though, LIBs are yet to find a credible contender.

While advances in materials and battery design have been impressive, the fundamental purpose of battery research has remained unchanged over the years: to decrease the weight and volume of the battery, increase its durability (i.e. the number of charge/discharge cycles of the battery) and minimize its cost while maintaining safety. Achieving these objectives requires an interdisciplinary approach involving material scientists, chemists, physicists, and a growing army of applied mathematicians and computer scientists. Indeed, increasing computing power, and improving the algorithms for quantum, atomistic and continuous simulations are making the feedback loop between the chemistry lab and the laptop screen a practical reality. Still, the complex multiscale nature of LIBs and NIBs, involving solid state ionics, interface science, polymer physics and engineering, means that plenty of open problems remain to be satisfactorily tackled.

From a mathematical and computational perspective, the goals stated above often translate into answering the following questions: *(i)* is it possible to accurately and efficiently estimate the electronic and/or ionic conductivities of the materials involved, based on prior knowledge of their atomistic structure?; *(ii)* can the chemical composition be optimized to enhance desirable features such as conductivity, cyclability or energy density?; *(iii)* can the structure and transport properties of the interface between different components be predicted from models of the individual phases?; and *(iv)* how can all this information be integrated into the macroscopic model of a realistic battery? Our research program in the recent years has attempted to contribute to answering these questions.

One of the critical challenges involving the atomistic simulation of battery materials is the slow diffusion of charge carriers at room temperature. To tackle this issue, we have developed enhanced sampling methods and multi-stage integration schemes that are suitable for solid-state systems including polarizable atoms, allowing us to simulate the ionic transport in materials for LIBs and NIBs at realistic operating conditions [2, 6, 30, 32, 60]. In order to produce experimentally meaningful results, an accurate model of the interatomic interactions (a *force field*) must be first developed or adapted. For the specific case of polarizable systems, the development of a force field can be quite challenging and the resulting model might not be accurate or lead to instability at certain chemical compositions of interest. We have overcome this shortcomings by introducing composition dependant force fields, adjusted with respect to accurate quantum mechanical simulations and/or experimental data [18, 19, 34]. Also, we provided a simple solution to deal with the additional degrees of freedom posed by the core-shell polarizability model [18].

One promising way to reduce the costs associated with experimental synthesis and characterization of new materials is in-silico material discovery and optimization. Despite considerable progress and the seemingly unstoppable growth of computing power, novel techniques to perform this task that can efficiently incorporate experimental information are still needed [57, 38]. On this front, we have designed Bayesian inference algorithms coupled with multi-stage integrators that will combine incoming data from atomistic simulations and information from experiments to produce the distribution of compositions that is most likely to maximize the expectation of a desired set of macroscopic properties [6, 61].

The article summarizes these contributions and presents some key examples involving state-of-the-art LIB and NIB components. In Sect. 2, we offer a short but necessary summary of the key fundamental aspects of rechargeable batteries. In Sect. 3, we provide a brief introduction to traditional atomistic simulation methods and introduce our enhanced sampling techniques for molecular simulation and computational statistics [61], along with the in-house advanced adaptive integration schemes for Hamiltonian dynamics [6]. In Sect. 4, we discuss the development of force fields for polarizable and non-polarizable materials. Sect. 5 presents our published applications of the discussed algorithms for modelling of sodium intercalation cathodes for NIBs and solid electrolyte garnets for LIBs [32, 18, 19, 41]. Finally, Sect. 6 provides concluding remarks and some future research directions. Specifically, we discuss the incorporation of the in-house Bayesian parameterization technique Mix & Match Hamiltonian Monte Carlo (MMHMC) [61] into the proposed framework for Bayesian materials screening and multiscale simulation of composite materials using the novel version of our Generalized Shadow Hamiltonian Monte Carlo (GSHMC) [2] that is specially adapted for sampling of coarse-grained systems (meso-GSHMC) [4].

2 Fundamental Aspects of LIBs and NIBs

We provide a summary of the key fundamental properties of the rechargeable batteries essential for this study.

2.1 Structure and Chemistry of Electrochemical Cells

The basic building blocks of any battery are the electrochemical cells (Fig. 1), in which charge carriers move from the anode (negative electrode) to the cathode (positive electrode) during the discharge. In LIBs, the carrier consists of positively charged lithium ions (Li^+) and the process is reversible: when an external voltage above the cell potential is applied, Li^+ is transferred from the cathode to the anode,

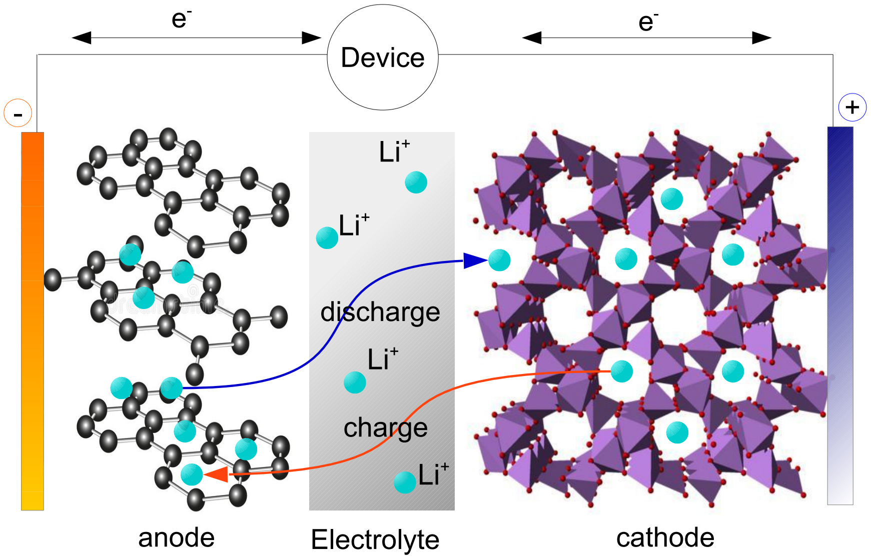

Fig. 1 Simplified structure of the electrochemical cell in a LIB. Each LIB consists of the anode (negative electrode) and the cathode (positive electrode) separated by the electrolyte containing dissociated lithium salts, which enables transfer of lithium ions between the two electrodes

charging the cell once again. The cathode and the anode are separated by the electrolyte containing dissociated lithium salts. In most commercial LIBs, electrolytes are comprised of organic liquid salts contained within a porous separator membrane, which avoids direct contact between the electrodes. When a lithium atom leaves the anode during the discharge, an electron is released to the outer circuit and performs electrical work. During the charge, electrons are injected into the anode from the external source, driving the release of Li^+ from the cathode. Hence, it is clear that the electrolyte must be electronically insulating while constituting a good medium for the diffusion of Li^+. In NIBs, the charge carriers are sodium ions (Na^+).

Currently, cathode materials are made of electrochemically active metal oxide particles (such as $XCoO_2$ and $XFePO_4$, with X = Li for LIBs or X = Na for NIBs), an inert binder "gluing" them together and some electronically conductive coating (typically carbon). The key feature of the active particles is their ability to accept and release Li^+ ions (lithiation/delithiation) with minimal changes to their crystal structure, through a process which Whittingham called the *intercalation* mechanism [75]. Anodes are most commonly made of graphite, because the lithiation/delithiation reaction is quite reversible without altering the mechanical and electrical properties of the material. On the first charge of the battery, a *passivation* layer, the so-called *solid electrolyte interphase* (SEI), forms from the decomposition of the electrolyte on the anode surface. This layer is of crucial importance for the battery operation in terms of safety, as it prevents the carbon from reacting with the electrolyte and helps avoiding graphite exfoliation [37, 54].

Important efforts have been made to replace liquid electrolytes with solids, due to the safety problems associated with the flammability of the organic salts that comprise them. Another driver is the impossibility to use metallic lithium (Li^0) as the anode along with liquid electrolytes, because the fast and disorder electrodeposition of Li^+ leads to the formation of Li^0 filaments (dendrites) that can penetrate the separator and connect both electrodes, producing a short-circuit. Since Li^0 anodes have a much higher energy density, their safe and efficient incorporation could greatly reduce battery volume/weight. Several solid and semi-solid electrolytes have been developed in recent years, including crystalline materials, polymers, gels, and composites mixing these three families [80]. Still, none has yet achieved the combination of ionic conductivity, thermal, mechanical and chemical stability and low cost necessary for widespread adoption.

While there are many factors influencing battery performance, we will focus on the one of particular importance: ionic conductivity.

2.2 Ionic Conduction in Battery Materials

The rate of X^+ ion transfer between the electrodes is a multiscale variable depending on *(i)* the ionic diffusivity within the active electrode particle; *(ii)* the grain boundary resistance between adjacent electrode particles; *(iii)* the rate of interfacial exchange between electrode and electrolyte; *(iv)* the ionic diffusivity within the electrolyte and, for solid electrolytes, the grain boundary resistance between adjacent electrolyte particles; and *(iv)* the ionic diffusivity within the SEI (Fig. 2). It is still possible to subdivide these resistances further, as electrode particles may be agglomerated through a binder and contain some coating particles to enhance the electronic properties (see Fig. 2).

2.2.1 Ionic Conduction in Cathodes

Within the active particles of the cathode, insertion of an X^+ leads to the reduction (acceptance of an electron) by the transition metal (M) (e.g. in $LiFePO_4$, iron is reduced from Fe^{3+} to Fe^{2+}). Clearly, the ease with which electrons can reach and reduce M and the ease with which X^+ can travel from the surface to the bulk of the particle are of paramount importance. The latter is captured by the ionic conductivity tensor (σ_X), which is defined as

$$\mathbf{J} = \sigma_X \mathbf{E},$$

where $\mathbf{J}$ is the electric current density transported by the charge carrier and $\mathbf{E}$ the electric field. Most active cathode particles, however, are orthotropic at the nanoscale

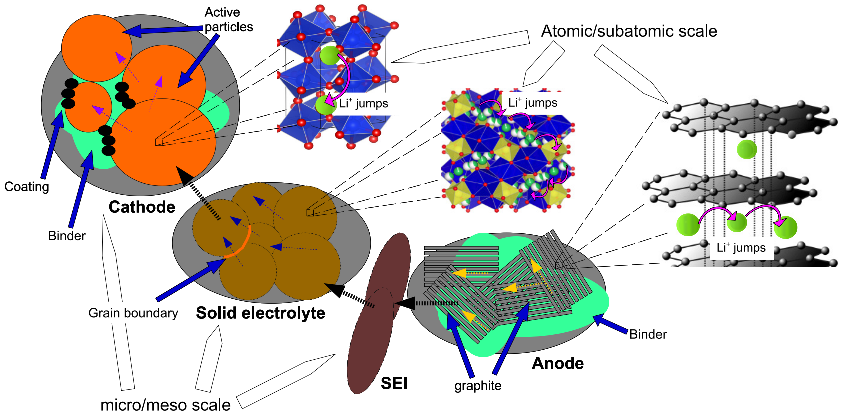

Fig. 2 The several scales of ionic conduction in an electrochemical cell for LIBs. Cathode, anode and electrolyte (a solid one in the sketch) are composed of active particle agglomerates, presenting grain boundary resistances at the particle/particle interfaces. Often, these particles are glued by a polymeric binder and may be coated to enhance electronic conduction. These micro/meso structures ($\sim 10^{-6}$ - 10^{-5} m) are generally isotropic. Within single active particle crystals, structural anisotropy can create preferential diffusion paths for Li^+. In addition, only particular sites are capable of accepting Li^+ ions, which means that intracrystalline Li^+ motion occurs in discrete jumps between those sites. Individual jumps occur at the atomistic scale ($\sim 10^{-10}$ - 10^{-9} m) and long range conduction happens over many such jumps ($\sim 10^{-8}$ - 10^{-7} m). However, subatomic scale phenomena such as oxygen polarization ($\sim 10^{-12}$ m) can strongly influence the energy barriers associated with the jumps

($\sim 10^{-9}$ m) and only the diagonal components of $\boldsymbol{\sigma}_{\mathrm{X}}$ are needed. In addition, because cathodes are made of randomly oriented collections of active particles, mesoscopic ionic conduction is isotropic and the conductivity tensor at this scale reduces to a simple scalar $\sigma_{\mathrm{X,meso}}$.

Most experimental measurements are microscopic ($\sim 10^{-6}$ m). Moreover, electron movement and reduction of M is comparatively fast compared to ionic diffusion, which means that the cathode conductivity reported in the literature normally corresponds to $\sigma_{\mathrm{X,meso}}$ [10]. There are, nonetheless, single crystal measurement methods capable of capturing conductivity anisotropy [7]. In most atomistic studies (ours included), it is assumed that intraparticle diffusion dominates over interparticle diffusion, and $\sigma_{\mathrm{X,meso}}$ is simply estimated as $\sigma^{*}_{\mathrm{X,meso}} = \frac{1}{3}\mathrm{tr}(\boldsymbol{\sigma}_{\mathrm{X}})$ [30, 70, 74, 18, 19]. This is not entirely correct, as grain boundary and other interfacial resistances can be quite significant and neglecting them leads to overestimated conductivities. In general terms, if the atomistically estimated conductivity is within an order of magnitude above the experimental values, the estimation is considered good enough. Extending atomistic methods to the mesoscale simulation of cathodes (coarse-graining) will provide powerful means to examine the impact of large-scale interfacial effects.

The self-diffusion coefficient for X^{+} in an orthotropic crystal is a diagonal tensor $\boldsymbol{D}_{\mathrm{X}}$ whose i-th entry is given by

$$\langle \Delta^2 r_i(t) \rangle = 2 D_{\mathrm{X},i} t, \qquad t >> t_c, \tag{1}$$

where $\langle \Delta^2 r_i(t) \rangle$ is the mean square displacement of the ions in the i-th direction after time t, and t_c is a characteristic diffusion time (for the cathode, it corresponds to the average interval between consecutive ionic jumps). $D_{\mathrm{X},i}$ is related to $\sigma_{\mathrm{X},i}$ through the corrected Nernst-Einstein equation:

$$\sigma_{\mathrm{X},i} = c_{\mathrm{X}} (Q_{\mathrm{X}} F)^2 \frac{D_{\mathrm{X},i}}{H_r k_B T}, \tag{2}$$

where c_{X} is the charge carrier concentration, Q_{X} is its ionic charge, F is the Faraday's constant, k_B is the Boltzmann's constant, T is the temperature and $H_r \leq 1$ is the Haven's ratio. H_r corresponds to the ratio of the charge carrier collective diffusivity to the self-diffusivity; it is unity in the small concentration limit. Above that, ionic displacements are correlated due to long-range Coulomb interactions and this correction is required [56].

2.2.2 Ionic Conduction in Electrolytes

In most electrolytes, positive ions (cations) move isotropically in all directions. However, in liquid and polymer electrolytes, X^{+} usually originates from the dissociation of a salt, which means that negative counterions (anions) are also present (e.g. standard Li salt $\mathrm{LiPF_6}$ dissociates into Li^{+} and $\mathrm{PF_6^{-}}$). Since the mobility of anions and cations is different (the anion is always much heavier), the diffusivity of both species must be considered in the calculation of the conductivity:

$$\sigma_i = c_{\mathrm{tot}} F^2 \frac{Q_+^2 D_{+,i} + Q_-^2 D_{-,i}}{H_r k_B T}, \tag{3}$$

where c_{tot} is the total ionic concentration, Q_v is the ionic charge of species v and $D_{v,i}$ its self-diffusivity in the i-th direction. $v = +$ for the cation and $v = -$ for the anion. Using atomistic simulations, it is straightforward to estimate the terms in equations (2) and (3) [20]. While for single ionic carriers it is often a reasonable approximation to take $H_r = 1$, this almost always leads to large errors for dissolved salts, because of the strong tendency cations and anions have to cluster and behave in a collective fashion. Finally, in crystalline solid electrolytes such as $Li_7La_3Zr_2O_{12}$, X^+ moves through a partially unoccupied lattice embedded within a crystalline framework [19, 41]. In these systems, ionic conduction occurs through discrete jumps between X sites. However, because the lattice contains vacancies, diffusion is somewhat more *fliud* than it is in typical cathode particles (there, diffusion requires the prior formation of a *defect* in order to produce a vacant site to which X^+ can jump to). Since no counterions are present in crystalline solid electrolytes, equation (2) can be used to obtain the ionic conductivity.

2.2.3 Ionic Conduction in Anodes

Since we have not tackled yet the study of conduction in graphite anodes, we will not discuss it in detail. Suffice it to say that ion intercalation in graphitic anodes involves ion absorption and diffusion through the SEI, and the occurrence of phase transitions within the ordered graphite structure. As a consequence, it is rather complex and multiscale in nature. A recent review on advances in the modelling of anodes can be found in [58].

3 Sampling in Atomistic Simulation

Molecular dynamics (MD) and Metropolis Monte Carlo (MC) are traditional sampling techniques broadly used for solid-state atomistic simulation [39]. Recently, we have proposed the class of enhanced sampling methods [5] which combine the advantages of both MD and MC offering improved sampling efficiency and accuracy vital for the success of a molecular simulation. The methods utilize the modified Hamiltonians for sampling enhancement and hence the name of the class - Modified Hamiltonian Monte Carlo (MHMC). In this section, we briefly review the molecular dynamics and Monte Carlo approaches and discuss the main features of MHMC as well as the benefits of employing them in the study of battery materials.

3.1 Molecular Dynamics

Molecular Dynamics (MD) is a computer simulation approach in which the evolution in time of a group of interacting atoms is followed by integrating the Newtonian equations of motion. In general, MD is not trying to generate physically accurate trajectories: this is often an impossible task. Instead, quantities of interest are statistical averages computed during the sampling of phase space.

Let us denote the positions of a system of n particles as $\mathbf{r} \in \mathbb{R}^d$, where $d = 3n$, and the velocities as $\mathbf{v} \in \mathbb{R}^d$ and assume that the system is conservative, i.e. a potential energy function U exists such that

$$F(\mathbf{r}) = -\nabla_{\mathbf{r}} U(\mathbf{r}), \tag{4}$$

where F are the forces. The function U is also called a *force field* and it contains the contribution to the potential energy from all interatomic interactions. Then, for the classical n-body problem in MD-simulations, Newton's second law can be expressed as

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}, \quad M\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = F(\mathbf{r}), \tag{5}$$

where $M \in \mathbb{R}^{d\times d}$ is the (symmetric positive definite) mass matrix.

The Hamiltonian formulation of Newtonian mechanics (5) reads

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{p}} = M^{-1}\mathbf{p}, \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{r}} = -\nabla_{\mathbf{r}} U(\mathbf{r}). \tag{6}$$

Here $\mathbf{p} = M\dot{\mathbf{r}}$,[1] $\mathbf{p} \in \mathbb{R}^d$ is the linear momenta and the *Hamiltonian* or energy functional is defined as

$$H(\mathbf{r}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + U(\mathbf{r}), \tag{7}$$

where the super index T denotes the transpose vector. The equations in (6) are the *Hamilton equations of motion*. By introducing the $2d\times 2d$ matrix $J = \begin{bmatrix} 0_{d\times d} & -I_{d\times d} \\ I_{d\times d} & 0_{d\times d} \end{bmatrix}$, (6) can be rewritten as (cf. [65])

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \mathbf{r} \\ \mathbf{p} \end{bmatrix} = J^{-1}\nabla H(\mathbf{r}, \mathbf{p}), \tag{8}$$

where

$$\nabla H(\mathbf{r}, \mathbf{p}) = \left[\frac{\partial H}{\partial \mathbf{r}}, \frac{\partial H}{\partial \mathbf{p}}\right]^T .$$

The space defined by the vectors $(\mathbf{r}, \mathbf{p})$ satisfying equation (8) is called the *phase space*. Under suitable conditions, there is a unique exact solution at a time t for every initial point $(\mathbf{r}(0), \mathbf{p}(0))$, and, for any time t, a *flow map* ϕ_t can be defined as

---

[1] For simplicity, we use here the dot-notation for the derivatives in time.

$$(\mathbf{r}(t), \mathbf{p}(t)) = \phi_t(\mathbf{r}(0), \mathbf{p}(0)). \tag{9}$$

Hamiltonian flows have a number of important properties, including (cf. [65]):

1. **Symplecticity**: For all t, a flow map ϕ_t is symplectic if, at each point $(\mathbf{r}, \mathbf{p})$ in phase space,
$$\phi_t'(\mathbf{r}, \mathbf{p})^T J \phi_t'(\mathbf{r}, \mathbf{p}) = J,$$
where $\phi_t'(\mathbf{r}, \mathbf{p})$ is the $2d \times 2d$ Jacobian matrix of ϕ_t. An important consequence of symplecticity is phase volume preservation (oriented volume): If A represents a region in phase space, then the volume of $\phi_t(A)$ is invariant in t.
2. **Time reversibility**: The Hamiltonian system described by equations (7) - (9) is time reversible; i.e., for all t, $\phi_t^{-1} = \phi_{-t}$.

There are another two characteristics that do not apply generically to all Hamiltonian flows. However, they have been empirically observed in the flows arising from many systems in MD simulations (cf. [53]):

3. **Ergodicity**: In statistical physics, ergodicity is defined to be the property that for any observable Ω, the time averages of Ω eventually converge to the average of Ω over the phase space. More precisely, for a probability measure π preserved under the Hamiltonian flow, ergodicity implies that
$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \Omega(\phi_t(\mathbf{r}, \mathbf{p})) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \Omega(\mathbf{r}, \mathbf{p}) \mathrm{d}\pi(\mathbf{r}, \mathbf{p}). \tag{10}$$
This is very useful for simulations, as it allows approximating the spatial averages by means of time-averages available through MD.
4. **Sensitive dependence on initial conditions (SDIC)**: In high-dimensional Hamiltonian systems, there are arbitrarily close points $(\mathbf{r}', \mathbf{p}')$ to $(\mathbf{r}(0), \mathbf{p}(0))$, such that the positive orbit starting at $(\mathbf{r}', \mathbf{p}')$ eventually diverges from the one starting at $(\mathbf{r}(0), \mathbf{p}(0))$. As a consequence, MD trajectories cannot be computed for long periods of time in a classical sense (that is, point-to-point matching between exact value and computed value), because even the unavoidable round-off error will naturally be magnified, thus potentially rendering computed values that are far from the actual values.

SDIC and ergodicity are two phenomena playing complementary roles in computing long-term trajectories of MD flow-maps: while SDIC prevents the computation of accurate trajectories, ergodicity allows the extraction of statistical information from the same computations. From a numerical standpoint, if a discretization $\phi_{L\Delta t}$ of the flow map ϕ_t (where $L\Delta t$ is a trajectory length and Δt is a time step) is symplectic, then averages of an observable Ω calculated along orbits of $\phi_{L\Delta t}$ will converge to the desired spatial average as the time step Δt tends to 0. Efficient and simple integrators providing symplecticity and time reversibility are available, the most popular of which are the Störmer-Verlet method and its extensions (leapfrog, velocity Verlet, etc.). Nonetheless, MD still suffers from several drawbacks:

(i) As stated in equation (10), making use of ergodicity (if it exists) requires sufficiently long trajectories. In systems with large energy barriers such as the ones between adjacent sites in cathode materials, meaningful trajectories may need to be hundreds of nanoseconds long ($\sim 10^{-7}$ s) to sample atomic jumps. Because time steps need to be small (typically $\sim 10^{-15}$ s), the number of steps required can be as high as $L \sim 10^8$. For a tens-of-thousands atoms system this can be very demanding, even with robust computing architectures.

(ii) Because MD only allows for slow exploration of configurational space through a sequence of many small steps, systematic discretization error can be significant.

(iii) In the absence of external forces, the flow defined by (9) is energy preserving: the dynamics occurs on a surface in $\mathbb{R}^{2d}$ for which H is constant. The probability density arising from this condition is the so-called *microcanonical ensemble*. However, most practical applications occur either at constant number of particles, N, temperature, T, and volume V (the *canonical* or NVT ensemble) or at given N, T and pressure, P (the *isothermal-isobaric* or NPT ensemble). In the latter two ensembles, the Hamiltonian must be modified by adding additional degrees of freedom, which are coupled to the particle velocities (thermostatting) and simulation domain dimensions (barostatting), increasing computational cost and introducing additional sources of error.

One significant advantage of MD is the ability to capture dynamical properties, such as the diffusion coefficients in equations (2) and (3). In addition, even though the *true* evolution of the system may not be captured, qualitative information regarding the underlying mechanisms of the physical processes can still be extracted.

3.2 Monte Carlo Simulations

While MD tries to estimate the value of the integral on the r.h.s. of equation (10) using a deterministic scheme, Monte Carlo (MC) simulations attempt to approximate the value of such integral using a stochastic approach. For simplicity, we will refer only to the canonical ensemble, for which the probability density distribution is

$$\pi(\mathbf{r}, \mathbf{p}) \propto \exp(-\beta H(\mathbf{r}, \mathbf{p})), \tag{11}$$

where $\beta = (k_B T)^{-1}$ is the thermodynamic beta. The position and momenta contributions to the Hamiltonian in equation (7) are separable. Moreover, the momenta part can be analytically integrated, which means that only the target distribution

$$\pi(\mathbf{r}) \propto \exp(-\beta U(\mathbf{r})), \tag{12}$$

needs to be sampled to obtain the desired system configurations. In order to collect N samples, this distribution is sampled using the Metropolis-Hastings Markov chain MC algorithm [42]:

- The current configuration in the chain $\mathbf{r}$ is randomly perturbed to generate a trial atomic configuration $\mathbf{r}'$. A typical perturbation might be a single-particle displacement: randomly pick an atom and displace its position by a small amount.
- Accept the proposal configuration to become $\mathbf{r}^{\text{new}}$ with probability

  $$\alpha = \min\left\{1, \frac{\pi(\mathbf{r}')}{\pi(\mathbf{r})}\right\} = \min\left\{1, \exp[-\beta(U(\mathbf{r}') - U(\mathbf{r}))]\right\}.$$

  Otherwise, dismiss $\mathbf{r}'$ and set $\mathbf{r}^{\text{new}}$ to $\mathbf{r}$.
- Repeat the two steps until the number of configurations reaches N.

Following this procedure, due to the law of the large numbers, the average of an observable Ω can be estimated as

$$\langle\Omega\rangle = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \Omega(\mathbf{r}^i),$$

where $\{\mathbf{r}^i\}$ is the set of accepted configurations.

Metropolis MC has a number of advantages with respect to MD. Thus, MC does not have an equivalent of a time step error, its sampling is exact with only statistical errors involved. The method is flexible in the choice of sequences of steps which potentially can lead to rapid exploration of configurational space. However, it also has a number of important drawbacks: for systems with large numbers of degrees of freedom, such as biomolecules, it may become impractical due to the difficulties in specifying a physically meaningful move leading to high acceptance rates. In addition, the method does not provide dynamical information, which means that quantities such as diffusivities or conductivities (equations (1) - (3)) cannot be estimated.

3.3 Modified Hamiltonian Monte Carlo Methods

Our choice of simulation techniques for the effective study of ion transport in bulk and nanostructured materials is based on four requirements critical for the success of the project. Such methods should *(i)* sample efficiently multidimensional space; *(ii)* reproduce dynamical properties of a simulated system; *(iii)* be able to detect rare events; and *(iv)* be easily extended to simulations on meso-scales.

The Modified Hamiltonian Monte Carlo (MHMC) methods were originally developed for atomistic simulations of complex systems [46, 68, 1, 2] and then adapted for multiscale [31] and mesoscale simulations [4]. They proved to be successful in the study of rare events in complex biological processes [73, 31, 4, 5] though never have been applied to solid-state chemistry until recently when we proposed using them for the simulation of battery materials [32, 18, 41, 19].

MHMC are importance sampling Hybrid Monte Carlo (HMC) methods [28] that achieve higher efficiency than MC, HMC or MD by sampling with respect to a modified Hamiltonian. Such methods are especially appropriate when exploring high-dimensional configurational spaces, helping in finding energy global minima and simulating rare events such as slow chemical reactions or phase transitions. Furthermore, the MHMC methods originated from the generalized hybrid Monte Carlo (GHMC) [44, 49] can keep the dynamic information throughout the sampling process similar to stochastic Langevin and Brownian dynamics simulations.

In general, to simulate the properties of physical systems, one collects samples with respect to a target distribution such as $\pi(\mathbf{r})$ in (12). For that purpose, HMC methods are a suitable choice as a sampling technique. Given a separable Hamiltonian (7), HMC generates samples of configurations from the augmented canonical distribution (11). Then, marginalizing momenta variables out, one can obtain the samples with respect to the target distribution $\pi(\mathbf{r})$ in (12).

Instead of sampling from the canonical distribution (11), MHMC methods sample from an importance canonical density

$$\tilde{\pi}(\mathbf{r}, \mathbf{p}) \propto \exp\left(-\beta \tilde{H}^{[k]}(\mathbf{r}, \mathbf{p})\right), \tag{13}$$

where $\tilde{H}^{[k]}$ is the kth order truncation of the modified Hamiltonian which is preserved exactly by a symplectic integrator [65]. For certain $H_2, H_3, \ldots$, a modified Hamiltonian reads as

$$\tilde{H} = H + \Delta t H_2 + \Delta t^2 H_3 + \ldots,$$

where Δt is an integration time step. For an integrator of order m, $\tilde{H} = H + O(\Delta t^m)$, so that $H_2, \ldots, H_m$ vanish. Thus, for $k > m$,

$$\tilde{H}^{[k]} = H + \Delta t^m H_{m+1} + \cdots + \Delta t^{k-1} H_k. \tag{14}$$

According to [6, 14], for a system of D particles, the expectation of the increments of H and $\tilde{H}^{[k]}$ in an integration leg satisfy

$$\mathbb{E}[\Delta H] = O\left(D \Delta t^{2m}\right), \tag{15}$$

and

$$\mathbb{E}\left[\Delta \tilde{H}^{[k]}\right] = O\left(D \Delta t^{2k}\right). \tag{16}$$

respectively.

Let us now describe the algorithm of a generic MHMC method. Given a sample $(\mathbf{r}, \mathbf{p})$ from the distribution $\tilde{\pi}$, the next configuration $(\mathbf{r}^{\text{new}}, \mathbf{p}^{\text{new}})$ is obtained as follows:

- Set up the new momentum $\mathbf{p}^*$ by applying a momentum refreshment procedure that preserves the importance density $\tilde{\pi}$ in (13).
- Generate a proposal configuration $(\mathbf{r}', \mathbf{p}')$ by simulating the Hamiltonian dynamics (6) with the Hamiltonian H in (7) and the initial condition $(\mathbf{r}, \mathbf{p}^*)$ using a symplectic and time-reversible numerical integrator Ψ (cf. [65]).
- Accept the proposal configuration $(\mathbf{r}', \mathbf{p}')$ to become $(\mathbf{r}^{\text{new}}, \mathbf{p}^{\text{new}})$ with probability

$$\alpha = \min\left\{1, \frac{\tilde{\pi}(\mathbf{r}', \mathbf{p}')}{\tilde{\pi}(\mathbf{r}, \mathbf{p}^*)}\right\}.$$

Otherwise, reject the proposal and perform a momentum flip, i.e. $(\mathbf{r}^{\text{new}}, \mathbf{p}^{\text{new}}) = (\mathbf{r}, -\mathbf{p}^*)$.

The reader should notice that

$$\frac{\tilde{\pi}(\mathbf{r}', \mathbf{p}')}{\tilde{\pi}(\mathbf{r}, \mathbf{p}^*)} = \exp\left(-\beta\left(\tilde{H}^{[k]}(\mathbf{r}', \mathbf{p}') - \tilde{H}^{[k]}(\mathbf{r}, \mathbf{p}^*)\right)\right) = \exp\left(-\beta\Delta\tilde{H}^{[k]}\right). \tag{17}$$

For comparison, in the Hybrid Monte Carlo algorithm, the corresponding ratio is given by

$$\frac{\pi(\mathbf{r}', \mathbf{p}')}{\pi(\mathbf{r}, \mathbf{p})} = \exp\left(-\beta\Delta H\right). \tag{18}$$

Therefore, since $k > m$, one can see from (15) - (18) that MHMC methods may provide higher acceptance rates than regular HMC methods. The fewer rejections has a positive impact not only on the sampling efficiency but also on the accuracy of dynamics since less momentum flips occur [2, 3]. Moreover, for very big systems, where performance of HMC degrades dramatically, MHMC algorithms may counterbalance the error introduced by the size D and maintain a high acceptance rate by choosing a bigger truncation order k without increasing the order of the integrator m [6].

Since in MHMC methods the samples are generated with respect to the importance density (13), the computation of averages of any observable Ω with respect to the canonical distribution (11) requires a *reweighting*. Given N samples of the observable Ω_i, $i = 1, 2, \ldots, N$ along a trajectory $(\mathbf{r}^i, \mathbf{p}^i)$ drawn from (13), the average of Ω with respect to (11) is calculated as

$$\langle\Omega\rangle_\pi = \frac{\sum_{i=1}^N \omega_i \Omega_i}{\sum_{i=1}^N \omega_i}, \tag{19}$$

where the importance weights are computed as

$$\omega_i = \exp\left(-\beta\left(H(\mathbf{r}^i, \mathbf{p}^i) - \tilde{H}^{[k]}(\mathbf{r}^i, \mathbf{p}^i)\right)\right) \quad \forall i = 1, 2, \ldots, N.$$

A big variability among weights would mean that the canonical density π in (11) and the importance density $\tilde{\pi}$ in (13) are not close. This would lead to errors in the averages, as many samples would not contribute significantly to the computation in (19). Such a situation is well controlled, however, in the MHMC methods through the appropriate choice of an integration step and an order of a modified Hamiltonian (cf. (14)). This makes samples of (13) an efficient means towards computing expectations with respect to (11).

The various algorithms in the MHMC class of methods may differ in the elements described above: refreshing the momenta, simulating the Hamiltonian dynamics or computing the modified Hamiltonians.

For our purposes here, the Generalized Shadow Hybrid Monte Carlo (GSHMC) method [2] is of special interest. As its predecessor the Targeted Shadowing Hybrid Monte Carlo (TSHMC) method [1], GSHMC takes advantage of a partial momentum update based on the ideas of Horowitz [44], which helps preserving partially dynamics and enhancing sampling efficiency:

$$\begin{aligned}\mathbf{p}^r &= \cos\varphi\,\mathbf{p} + \sin\varphi\,\mathbf{u},\\ \mathbf{u}^r &= -\sin\varphi\,\mathbf{p} + \cos\varphi\,\mathbf{u},\end{aligned} \tag{20}$$

where $\varphi \in (0, \pi/2]$ is a parameter and $\mathbf{u}$ a random noise drawn from $\mathcal{N}(0, \beta^{-1}M)$. The proposed refreshed momentum $\mathbf{p}^r$ is accepted, i.e. $\mathbf{p}^* = \mathbf{p}^r$, with the probability

$$\alpha_p = \min\left\{1, \frac{\hat{H}(\mathbf{r}, \mathbf{p}^r, \mathbf{u}^r)}{\hat{H}(\mathbf{r}, \mathbf{p}, \mathbf{u})}\right\}, \tag{21}$$

where

$$\hat{H}(\mathbf{r}, \mathbf{p}, \mathbf{u}) = \tilde{H}^{[k]}(\mathbf{r}, \mathbf{p}) + \frac{1}{2}\mathbf{u}^T M^{-1}\mathbf{u}.$$

In the case of rejection, we set $\mathbf{p}^* = \mathbf{p}$.

GSHMC was originally introduced for atomistic simulations in canonical ensembles [2, 73, 30] and then adapted to simulations in isobaric-isothermal ensembles [35].

Different extensions of GSHMC aiming at specific applications appeared recently. For meso-scale simulations, the meso-GSHMC method [4] has been proposed. In coarse-grained mesoscopic models, the fluctuation-dissipation contributions should mimic the impact of non-resolved finer details of an atomistic model on the coarse grained length and time scales while maintaining the system at a desired temperature. For this purpose, the stochastic thermostat called Dissipative Particle Dynamics (DPD) [33] is used. The meso-GSHMC puts DPD in the GSHMC framework by introducing a DPD-type momentum update step which conserves total linear and angular momenta. Unlike original DPD, meso-GSHMC samples exactly from the canonical distribution (11) while dealing with the fluctuation-dissipation terms.

For multiscale simulations, a naturally possessed weak stochasticity of GSHMC can be combined with the modified impulse multi-time-stepping (MTS) molecular dynamics and the modified Hamiltonians specially derived for modified MTS integrators. The resulting MTS-GSHMC method [31] demonstrates the improved stability of the MTS integrators and superior sampling performance over multi-time-stepping molecular and Langevin dynamics.

More recently, the MHMC method for solid-state atomistic simulations called Randomized Shell Mass GSHMC (RSM-GSHMC) has been proposed [32]. In this algorithm, a mass randomization is implemented as a part of the momentum update step in order to reduce the negative effect of a shell mass within a dynamical shell model on the kinetics of a simulated system. Before updating the momenta, a fraction of the atomic mass is redistributed between the core and the shell, maintaining the total mass constant. Apart from the study where the methodology was introduced [32], it has been also used in [18].

All variants of GSHMC described above posses the strongest features of GSHMC such as the enhanced sampling, ability to reproduce dynamical properties of simulated systems and rigorous temperature control. As demonstrated in the studies cited above (see also [6] and references therein) for different problems, the use of GSHMC methods resulted in a systematic improvement of the sampling efficiency and accuracy with respect to those observed with conventional sampling techniques, such as MD, MC, HMC or Langevin dynamics.

While GSHMC methods were designed for applications in the field of molecular simulation, the Mix & Match Hamiltonian Monte Carlo (MMHMC) [61] method has been recently proposed for Bayesian inference problems. In simple terms, MMHMC is an adaptation of GSHMC to computational statistics. It proved to be successful in the simulation of popular statistical models such as multivariate Gaussian, Bayesian Logistic Regression and Stochastic volatility [61].

Two possible strategies for enhancing performance of MHMC methods are (i) a suitable choice of the method-specific parameters; and (ii) an optimization of conservation properties of numerical integrators used for simulating the Hamiltonian dynamics. The latter is explored in the following section.

3.4 Modified Adaptive Integration Approach

In this study, we limit ourselves to *separable* Hamiltonians, such as the one in (7), meaning that the Hamiltonian can be written as a sum $H \equiv A + B$ of two partial functions

$$A(\mathbf{r}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p}, \quad B(\mathbf{r}, \mathbf{p}) = U(\mathbf{r}),$$

which correspond to the kinetic and potential energies, respectively. Thus, the Hamilton equations of motion (6) can be rewritten as

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \nabla_{\mathbf{p}} A(\mathbf{r}, \mathbf{p}) = M^{-1}\mathbf{p}, \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\nabla_{\mathbf{r}} B(\mathbf{r}, \mathbf{p}) = -\nabla_{\mathbf{r}} U(\mathbf{r}).$$

These equations can be integrated analytically and their solution flows at a time t are given by (cf. (9))

$$(\mathbf{r}(t), \mathbf{p}(t)) = \phi_t^A(\mathbf{r}(0), \mathbf{p}(0)) = (\mathbf{r}(0) + tM^{-1}\mathbf{p}(0), \mathbf{p}(0)), \tag{22}$$

and

$$(\mathbf{r}(t), \mathbf{p}(t)) = \phi_t^B(\mathbf{r}(0), \mathbf{p}(0)) = (\mathbf{r}(0), \mathbf{p}(0) - t\nabla_{\mathbf{r}} U(\mathbf{r}(0))). \tag{23}$$

These solution flows of the partial systems, for a time t, associate the exact solution value $(\mathbf{r}(t), \mathbf{p}(t))$ with each initial condition $(\mathbf{r}(0), \mathbf{p}(0))$.

A splitting integrator can be constructed as a palindromic composition of the solutions flows (22) and (23) of the partial systems (cf. [17]). Here, we limit our attention to the one-parameter family of two-stage[2] splitting integrators studied in detail in previous works [16, 36, 6]. The map that advances the solution over one step Δt is

$$\psi_{\Delta t} = \phi^B_{b\Delta t} \circ \phi^A_{\Delta t/2} \circ \phi^B_{(1-2b)\Delta t} \circ \phi^A_{\Delta t/2} \circ \phi^B_{b\Delta t}, \tag{24}$$

where $b \in (0, 1/4]$ is the parameter that fully characterizes a two-stage integrator. Then, for a time $\tau = L\Delta t$ with $L \in \mathbb{N}$, one can also define the transformation Ψ_τ as the composition

$$\Psi_\tau = \underbrace{\psi_{\Delta t} \circ \cdots \circ \psi_{\Delta t}}_{L \text{ times}}.$$

The integrators in (24) are symplectic because they are a composition of Hamiltonian flows [65]. Moreover, they are also time-reversible due to the palindromic structure of (24) [16].

The optimal choice of a parameter b in (24) has been thoroughly investigated for the HMC [36] and MHMC methods [6, 60]. The suggested criteria for such "optimality" is the maximization of acceptance rates in HMC or MHMC due to the conservation properties of a chosen integrator.

For MHMC, the focus is on those two-stage integration schemes ψ that provide the lowest expected error $\Delta\tilde{H}^{[k]}$ in modified Hamiltonians with respect to a modified density (13). Here,

$$\Delta\tilde{H}^{[k]} = \tilde{H}^{[k]}(\Psi_\tau(\mathbf{r}, \mathbf{p}^*)) - \tilde{H}^{[k]}(\mathbf{r}, \mathbf{p}^*), \tag{25}$$

for a given time $\tau = L\Delta t$. Thus, the acceptance criterion in (17) is integrator-dependent.

In contrast to the well established practice in molecular simulation to use the same integrator, usually Verlet / leapfrog, for a broad range of the simulated systems, in [6], it has been proposed to construct *a system specific* two-stage integrator prior to an MHMC simulation. The Modified Adaptive Integration Approach (MAIA), which realises this idea, adapts the parameter b to a given simulated system and a chosen by a user value of step size Δt in such a way that the expected value of (25), $\mathbb{E}\left[\Delta\tilde{H}^{[k]}\right]$, is minimal. Based on the detailed analysis of the one-dimensional harmonic oscillator, the MAIA approach provides the upper bound $\rho(h, b)$ of $\mathbb{E}\left[\Delta\tilde{H}^{[k]}\right]$[3], where h is a dimensionless step size related to Δt by

$$h = \sqrt{3}\tilde{\omega}\Delta t.$$

Here, $\sqrt{3}$ is a safety factor used to avoid nonlinear resonances (cf. [6, 36, 67]) and $\tilde{\omega}$ is the fastest angular frequency of the two-body interactions. Then, the value of b is found as the one that minimizes

$$\max_{0<h<\tilde{h}} \rho(h, b),$$

---

[2] The *stage* notation first suggested by Blanes et al. [16] indicates the number of force evaluations performed per step.

[3] The expression for $\rho(h, b)$ as well as the details of its derivation can be found in [6].

where $\tilde{h}$ is the dimensionless time step associated to a user defined Δt.

One should note that, for a realistic physical model consisting of D harmonic oscillators, $(0, \tilde{h})$ is the shortest interval that contains all $h_i = \sqrt{3}\omega_i \Delta t$, where ω_i are the D frequencies in the problem.

It has to be stated that MAIA is a never-fail strategy, meaning that it guarantees, by construction, that the numerical stability is not lost. For the biggest allowed integration step sizes, the parameter b in (24) becomes 1/4, which means that the integrator of choice is the two-stage version of the classic velocity Verlet. This integration scheme is known as the integrator with the largest stability interval among two-stage integrators, namely $(0, 4)$.

The extended MAIA (e-MAIA) [6] is capable of keeping control of the momenta acceptance rate (21) at the user-desired level for any given problem through the intelligent choice of the parameter φ in the momentum update (20).

Both MAIA and e-MAIA proved their efficiency in simulations of complex molecular systems in Materials Science, Biology and Chemistry [32, 6, 18, 41, 19].

3.5 Software Implementation

The GSHMC methods for molecular simulation, i.e. GSHMC, meso-GSHMC, RSM-GSHMC, described in Sect. 3.3 have been implemented in the in-house software package MultiHMC-GROMACS [30, 35, 36] which is built on the top of the popular suite of programs for molecular dynamics simulations GROMACS [12, 43]. GROMACS supports state-of-the-art molecular simulation algorithms and is well known for its computational efficiency and parallel scaling properties. These features are carefully maintained in MultiHMC-GROMACS.

Two-stage integrators (24) are implemented in MultiHMC-GROMACS as a concatenation of velocity Verlet steps[4] [36]. The implementation is general enough to allow the use of all members of the family (24), including the adaptive integration schemes MAIA and e-MAIA (Sect. 3.4). The analysis required for the MAIA and e-MAIA methods is done within the GROMACS preprocessing module [6]. Thus, no computational overheads are introduced. Once the integrator parameter b and the angle φ (in the e-MAIA case) are obtained, they are passed to the running module prior to a simulation.

The package offers the efficient implementation of various formulations of modified Hamiltonians derived for the range of GSHMC methods and numerical integrators. In addition to the GSHMC methods, the HMC samplers have been included in MultiHMC-GROMACS as particular cases of GSHMC [36]. The user can select any GSHMC or HMC method from the code simply by predetermining appropriate parameters in the modified GROMACS input file. The same applies to numerical integrators.

---

[4] The main idea is that, for every b, an integration step $\psi_{\Delta t}$ can be rewritten as $\psi_{\Delta t} = \left(\phi^B_{b\Delta t} \circ \phi^A_{\Delta t/2} \circ \phi^B_{(1-2b)\Delta t/2}\right) \circ \left(\phi^B_{(1-2b)\Delta t/2} \circ \phi^A_{\Delta t/2} \circ \phi^B_{b\Delta t}\right)$.

A detailed description of the MultiHMC-GROMACS package may be found in [30, 35, 36, 6, 32].

The Mix & Match Hamiltonian Monte Carlo (Sect. 3.3), proposed for applications in computational statistics, is implemented in the in-house software package HaiCS [61]. The package is designed for statistical sampling of high dimensional, complex distributions and parameter estimation in different models by means of Bayesian inference using HMC-based methods. The code benefits from the efficient implementation of modified Hamiltonians, multi-stage splitting integrators [60], performance analysis tools compatible with CODA toolkit [59] and provides a user-friendly interface for implementing alternative HMC-type samplers, splitting integrators and statistical models.

Both MultiHMC-GROMACS and HaiCS are written in C and targeted to computers running UNIX certified operating systems.

4 Force Fields

As briefly introduced in Sect. 3.1, a force field is a mathematical expression describing the dependence of the energy of a system on the positions of its particles. It is fully defined by a functional form of the interatomic potential energy, $U(\mathbf{r})$, and a set of parameters. There are several force fields in use addressed to particular families of materials (e.g. AMBER, for proteins and DNA [27] or CHARMM for polymers [24]). However, it is clear that for many specific materials and applications, custom force fields must be developed. Force field development has three fundamental requirements:

- an appropriate functional form for the interaction potentials,
- a training data set of structural, mechanical and/or thermodynamic properties (e.g. experimental, computed, etc.) to fit the model, and
- an optimization strategy to perform the fitting.

Ideally, a force field must be simple enough to be evaluated quickly, but sufficiently detailed to reproduce the properties of interest. Below, we elaborate on each of these requirements, with particular focus on those interactions that are typical to intercalation in cathodes and solid electrolyte materials. Hence, metallic systems are not included in this discussion.

4.1 Functional Form of the Interaction Potentials

Let us denote a position of atom i in an n-particle system as $\mathbf{r}_i$, and the Euclidean distance between atoms i and j as r_{ij}. The general expression for most force fields has the form:

$$U(\mathbf{r}_1, \ldots, \mathbf{r}_n) = U_{\text{non-bonded}}(\mathbf{r}_1, \ldots, \mathbf{r}_n) + U_{\text{bonded}}(\mathbf{r}_1, \ldots, \mathbf{r}_n).$$

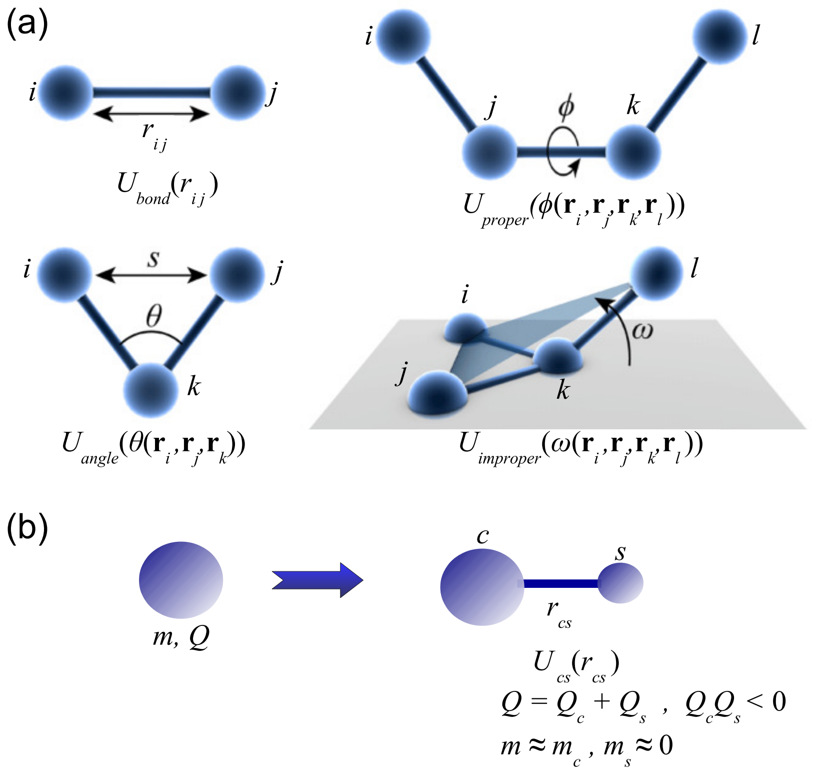
(a)
i
j
r_{ij}
$U_{bond}(r_{ij})$
i
l
j
ϕ
k
$U_{proper}(\phi(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k,\mathbf{r}_l))$
i
s
j
θ
k
$U_{angle}(\theta(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k))$
l
i
ω
j
k
$U_{improper}(\omega(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k,\mathbf{r}_l))$
(b)
m, Q
c
s
r_{cs}
$U_{cs}(r_{cs})$
$Q = Q_c + Q_s\ ,\ Q_c Q_s < 0$
$m \approx m_c\ ,\ m_s \approx 0$

Fig. 3 **a** Typical bonded interactions: bonding (U_{bond}), angular (U_{angle}), proper dihedral (U_{proper}) and improper dihedral (U_{improper}) interactions. $U_{\text{bonded}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{proper}} + U_{\text{improper}}$. **b** Core-shell model of polarizability: the polarizable ion of mass m and charge Q is represented as a core (c) and a shell (s) with opposite charges Q_c and Q_s, respectively, interacting through a bond potential U_{cs} (usually harmonic, $U_{cs} = k_{cs} r_{cs}$). The mass of the ion is concentrated in the core ($m = m_c$), to allow for instantaneous thermalization of the shell. In practice, however, it is better to choose a small shell mass ($m_s > 0$, $m_s << m_c$) to boost computational efficiency

Contributions to $U_{\text{bonded}}(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ come from groups of chemically linked atoms (sharing electrons), and are typically divided into two-body bond-stretching potentials, three-body angular-bending potentials, and four-body dihedral potentials (Fig. 3a). For instance, the form of U_{bonded} in the popular OPLS [48] and AMBER [27] force fields is:

$$U_{\text{bonded}}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \sum_{b=1}^{n_{\text{bonds}}} k_b (r_{ij} - r_{0,b})^2_{i,j\in b} + \sum_{a=1}^{n_{\text{angles}}} k_a (\theta_{ijk} - \theta_{0,a})^2_{i,j,k\in a}$$
$$+ \sum_{d=1}^{n_{\text{dihedrals}}} \sum_{m=1}^{3} \frac{k_{d,m}}{2} (1 + \cos m\phi_{ijkl})_{i,j,k,l\in d},$$

where angle $\theta_{ijk} = \theta(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k)$ (analogously for proper dihedral ϕ_{ijkl}) and $\{r_{0,b}\}$, $\{\theta_{0,a}\}$, $\{k_a\}$, $\{k_b\}$ and $\{k_{d,m}\}$ are positive force field parameters. There are no improper dihedrals in the original version of these force fields, but later updates have included them.

Polarizability is the ability for some atoms to form instantaneous dipoles. In many cases, its effect can be strong enough to significantly influence thermodynamic and transport properties [74, 78]. One of the simplest way to incorporate polarizability into the force field is the core–shell model suggested by Dick and Overhauser [26], in which a central core of a charge Q_c and a shell of a charge Q_s, with $Q_s Q_c < 0$, are introduced in such a way that the sum of these charges $Q_c + Q_s$ equals the ion charge Q (Fig. 3b). The core and shell are coupled together in a unit via a bond potential (usually harmonic), which allows the shell to move with respect to the core, thus simulating a dielectric polarization. In the original proposal by Dick and Overhauser, the core mass m_c was equal to the mass of the ion, m (i.e. the shell mass m_s was zero), in order to enable instant thermalization of the shell. Mitchell [55] showed that taking a sufficiently small value for m_s allowed simulating the core-shell system using conventional MD, as the thermal contribution from the shell was insignificant. This is called the cs-adiabatic scheme (cs-adi), which we have employed in our previous work [18]. The force field parameters for the harmonic cs-adi model, $U_{cs} = k_{cs} r_{cs}^2$, are the spring constant k_{cs}, Q_s and m_s.

The non-bonded interactions present in a typical force field are pairwise Coulomb, U_{Coul}, and van der Waals, U_{vdW}, interactions

$$U_{\text{non-bonded}} = U_{\text{Coul}} + U_{\text{vdW}},$$

acting over all atom pairs (i,j). In general, however, contributions from pairs already interacting through bonded potentials are totally or partially excluded. U_{Coul} is given by

$$U_{\text{Coul}} = \frac{1}{4\pi\epsilon} \sum_{i,j=1}^{n} \frac{Q_i Q_j}{r_{ij}}, \tag{26}$$

where ϵ is the vacuum permittivity and Q_i the charge of atom i. Q_i is a force field parameter, but in solid state systems its value is often taken as eZ_i, where e is the electron charge and Z_i the oxidation state of atom i. U_{vdW} can take many possible forms, the simplest non-trivial of which is the hard-sphere potential. For battery materials, two types of vdW potentials are particularly common: the Lennard-Jones (LJ)-type potentials, U_{LJ} (for liquid and polymeric electrolytes) and the softer Born potential U_{Born} (for crystalline materials) [40]. The LJ potentials between i and j atoms is given by

$$U_{\text{LJ},ij} = \frac{A_{ij}}{r_{ij}^n} - \frac{B_{ij}}{r_{ij}^m},$$

where $A_{ij} > 0$ and $B_{ij} > 0$ are force field parameters. Powers (n, m) typically assume the values (9,6) or (12,6), but can also be taken as adjustable parameters. $U_{\text{Born},ij}$, on the other hand, is a softer potential, allowing for greater interatomic overlap:

$$U_{\mathrm{Born},ij} = A_{ij} \exp\left(- \frac{r_{ij}}{\rho_{ij}} \right) - \frac{C_{ij}}{r_{ij}^{6}} + \frac{D_{ij}}{r_{ij}^{8}}, \tag{27}$$

where A_{ij}, ρ_{ij}, C_{ij} and D_{ij} are positive force field parameters. U_{vdW} can, of course, be written as a combination of several types of potentials: e.g. when modelling polymer/garnet composite solid electrolytes, the LJ potentials can be employed to model the interactions in the polymer phase, while the Born potential can be used to model interactions within the garnet. In any scenario, $U_{\mathrm{vdW}} = \frac{1}{2} \sum_{i,j=1}^{n} U_{\mathrm{vdW},ij}$.

4.2 Training Dataset

It is important to note that force fields partition the total electronic energy into well defined atom-atom contributions, such as Coulomb, polarization, dispersion, etc. However, it is impossible to fully separate the intricate electronic effects this way. The proper choice of a functional form of a force field and a set of data used to parameterize it (the training dataset, TD) usually makes it possible to generate a force field that is valid at least for specific problems and within the chemical and environmental constraints from the TD.

The TD typically comprises data obtained either from *ab initio* or semi-empirical quantum mechanical (QM) calculations, or from experimental measurements such as neutron, X-ray and electron diffraction, NMR, Raman spectroscopy, etc. Naturally, experiments can only probe a limited set of properties and their reliability depends on the quality of the experimental setup and the examined sample. The synthesis of many battery materials is by no means an exact science, and structural defects and impurities can be significant. That is why, it is expected that TDs from extensive QM calculations lead to more generally applicable force fields compared to those from experimental data. However, QM calculations still contain plenty of approximations, and a QM-based force field is of little or no use if it cannot reproduce experimental values or trends. In many cases, force fields trained through QM-based TDs are further fine-tuned with respect to experimental measurements.

4.3 Optimization Strategy to Perform the Fitting

The optimization strategy refers to the technique through which the set of optimal parameters is determined. When using an experimental TD, the general idea is to find the set of force field parameters $\gamma_1, \ldots, \gamma_M$ that minimizes the scoring function

$$f(\gamma_1, \ldots, \gamma_M) = \sum_{i=1}^{L} \parallel \omega_i (\langle \Omega \rangle_{\mathrm{sim},i} - \Omega_i^0) \parallel, \tag{28}$$

where Ω_i^0 is the ith reference observation in the TD, $\langle\Omega\rangle_{\mathrm{sim},i}$ its atomistically simulated value with the trial force field, L the total number of observations, ω_i is the weight given to observation i and $\|\cdot\|$ a suitable norm.

In exploiting a QM-based TD, the scoring function has a rather different form. While experimental TD's are invariably composed of average properties (e.g. unit cell sizes, Young's moduli, etc.) that result from a configurational average in a particular ensemble (even the fastest neutron scattering measurement results from averaging over an infinite number of configurations), QM-based TDs are normally composed of instantaneous snapshots of the equilibrated system. Hence, one can set up the particles in the exact positions provided by the snapshots, and then find the values of $\gamma_1, \dots, \gamma_M$ that achieve, for example, the total energy closest to the QM energy for each snapshot. This in fact is what most studies have done. A much more effective strategy is to additionally adjust $\{\gamma_k\}$ to multidimensional variables such as forces and stresses, which are also available from QM calculations. This is precisely what the so-called *force-matching algorithm* (FMA) does and what we employ in our study. For a training dataset consisting of L configurations, a suitable FMA scoring function is given by [22]

$$f(\gamma_1, \dots, \gamma_M) = \sum_{l=1}^{L} \left[3\omega_e (e_l - e_l^0)^2 + \frac{1}{2}\omega_s \sum_{p=1}^{6} (s_{p,l} - s_{p,l}^0)^2 + \sum_{i=1}^{3n_l} (f_{i,l} - f_{i,l}^0)^2 \right], \tag{29}$$

where n_l is the number of particles in configuration l. $\{e_l\}$, $\{s_{p,l},\ p = 1, \dots, 6\}$ and $\{f_{i,l},\ i = 1, \dots, 3n_l\}$ are the energies, stresses and forces obtained with the trial force field, respectively. The superindex 0 denotes the reference values from the TD. ω_e and ω_s are user-defined weights to balance the amount of available information for each quantity. Such a scoring function has been proposed and implemented in the open software *potfit*, developed by Brommer et al. [22, 23]. We have adjusted this function to parametrization of a broader range of models by incorporating in it core-shell polarizable potentials and angular terms [32, 18].

The scoring functions (28) and (29) can be highly multidimensional and non convex. Occasionally, some of the force field parameters can be reliably set to known values from similar force fields, simplifying the problem and allowing the use of local optimization strategies [19]. In most cases, however, stochastic global optimization techniques, such as simulated annealing, are necessary to avoid local minima. This, of course, comes with the hurdle of finding suitable adjustable parameters for the algorithm itself.

5 Applications

Below, we summarize the highlights of recent work performed in our group on the atomistic simulation of battery materials for LIBS and NIBS, using the numerical schemes described above. Specifically, we focus on cathode material olivine $NaFePO_4$ for NIBs, and solid electrolyte material $Li_7La_3Zr_2O_{12}$ for LIBs.

5.1 Na Diffusion in Cathode Material Olivine $NaFePO_4$ for NIBS

Olivine $LiFePO_4$ is likely the most studied commercial cathode for LIBs [72]. It provides high stability, rate capability, and sustained high voltage throughout the discharge cycle. In principle, one could expect $NaFePO_4$ to inherit these properties from its isostructural lithium counterpart, making it a suitable cathode for NIBs. The first fundamental studies revealed that the *intercalation mechanisms* (see Sect. 2.1) of Li^+ and Na^+ in $FePO_4$ are, however, significantly different [81]. While Li_xFePO_4 $(0 \leq x \leq 1)$ is stable at all values of x, Na_xFePO_4 is only stable at discrete Na^+ contents: $x \in \{0.0, 0.66, 0.83, 1.0\}$ [66]. This means that during the discharge, the structure of Na_xFePO_4 evolves through combinations of these stable, highly organized phases (Fig. 4), adding another level of complexity to the study of Na^+ intercalation as compared to that of Li^+. This issue is particularly relevant when studying dynamic properties such as Na^+ mobility using simulations, requiring large supercells and long simulation times to account for possible sodium orderings. These requirements make the use of first-principles methods, such as density functional theory (DFT), prohibitive in terms of computational cost. We addressed this computational challenge using atomistic simulations. Two key elements are required for this task: an accurate force field and an efficient sampling technique.

Whiteside et al. [74] developed a force field for the structural simulation of $NaFePO_4$, incorporating the polarizability of the oxygen atoms through the core-shell model. We will denote such force field as **Whiteside-ff** from here on. However, removal of one Na^+ ion leads to the oxidation of one iron (i.e. iron goes from Fe^{2+} to Fe^{3+}), which means that Na_xFePO_4, with $x < 1$, contains a combination of both Fe^{2+} and Fe^{3+} (Fig. 4b). Whiteside-ff does not consider the presence of Fe^{3+} and, as a consequence, is not suitable for a general study of charge transport in Na_xFePO_4. To overcome this issue, we developed a new force field for this system, the **Na_xFePO_4-ff**, using accurate DFT calculations at the stable Na^+ contents. The force field employs the cs-adi model described in Sect. 4.1 to incorporate polarizability in oxygen and iron atoms.

For configurational sampling of the system, conventional MD (see Sect. 3.1) has two key disadvantages. On the one hand, it requires many simulation steps for the estimation of the room temperature diffusivities. On the other hand, the cs-adi version of the core-shell polarizability model allows only for small time steps, increasing the computation time even further. In sight of this, we selected instead the in-house RSM-GSHMC technique as our simulation tool [32] (see Sect. 3.3), which simultaneously provides enhanced sampling while incorporating a scheme to minimize the computational impact of the small shell mass.

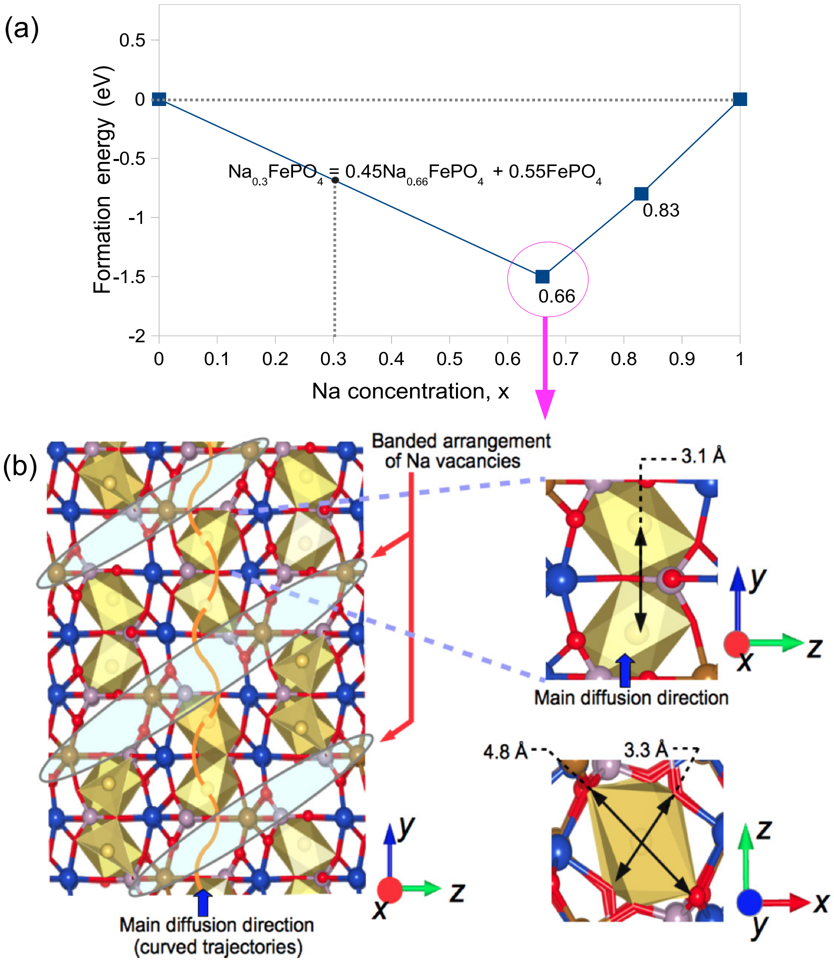

Fig. 4 **a** The convex-hull (blue line) of Na$_x$FePO$_4$. At any x, all possible system structures will have an energy of formation above the convex hull. As a consequence, the most stable structure for a given x will be a segregated mixture of the structures corresponding to the closest vertices in the convex-hull. These vertices are located at $x = 0.0, 0.66, 0.83$, and 1.0 [66]. For instance, a mol of Na$_{0.3}$FePO$_4$ will comprise 0.45 moles of Na$_{0.66}$FePO$_4$ and 0.55 moles of FePO$_4$. **b** *Left*: Structure of Na$_{0.66}$FePO$_4$: O^{2-} in green, P^{5+} in purple, Fe^{2+} in blue and Fe^{3+} in red. Na$^+$ is coordinated by six O^{2-} atoms (yellow octahedra). The curved trajectory followed by Na$^+$ along one of the channels in the main diffusion direction, y, is depicted as the solid orange line. Na$^+$ vacancies and Fe^{3+} ions follow a highly organized banded arrangement (similarly for $x = 0.83$). *Right-top*: close up showing the characteristic Na$^+$ jumping distance. *Right-bottom*: Cross section of a diffusion channel along y. The channels along x and z are much narrower, which means that jumps in these directions are rare. (Adapted with permission from [18]. Copyright (2018) American Chemical Society)

5.1.1 Force Field Development

Derivation

Ionic polarizability plays an important role in the dynamics of Na_xFePO_4. Similarly to Whiteside-ff [74], we divided the Fe^{2+} and O^{2-} ions into polarizable core-shell units. Hereafter, O_{sh} and Fe_{sh} will refer to the O^{2-} and Fe^{2+} shells, respectively. For the bonded interactions, we considered only the angular potential between the covalently linked O-P-O ions in PO_4 tetrahedra, following

$$U_{\text{bonded}} = U_{\text{angle}} = k_a \sum_{i=1}^{n_a} \frac{1}{2}(\theta_k - \theta_0)^2,$$

where n_a is the number of angles, and k_a and θ_0 are force field parameters. Non-bonded interactions were represented by a sum of the Coulomb (26) and Born (27) potentials. Only interactions between X - O_{sh}^{2-} pairs were considered in the Born contribution (X = Fe_{sh}^{2+}, Fe^{3+}, Na^{+}, P^{5+}, O_{sh}^{2-}). The cores of Fe^{2+} and O^{2-} only interacted electrostatically. The TD (see Sect. 4.2) comprised 58 olivine Na_xFePO_4 configurations consisting of (a) 11 quantum mechanically optimized structures at 0 K, ranging from 72 to 84 atoms, and (b) 58 equilibrated snapshots taken out of *ab initio* MD trajectories of an ideal $NaFePO_4$ crystal at 500 and 1000 K. All TD configurations were computed with the plane wave DFT code VASP [50]. As an optimization strategy, we used the force-matching algorithm described in Sect. 4.3 with the scoring function (29) and the adaptive simulated annealing scheme implemented in the in-house extended version of the open source software *potfit* [22, 23, 18].

Since the core-shell unit is simply a construct to introduce polarizability within an empirical force field, the *ab initio* dataset does not provide information on how to specify the core and shell coordinates. For the core, a sensible solution consists in setting its position to that of the actual ion, because the core carries most of the core-shell unit mass. For the shell, its position should be selected in such a way that the total force on it is zero. We proposed a consistent approach to treat the shells positions, which have usually been randomly chosen in previous works [18]. We fix the spring constant and charge distribution for all core-shell units prior to starting the parameter estimation. Then, given that dispersion interactions in a core-shell unit l are assigned to a shell only, the total force on a core l, $\mathbf{F}_{l,\text{core}}$, becomes

$$\mathbf{F}_{l,\text{core}} = \mathbf{F}_{l,\text{core}}^{\text{elect}} + k_l(\mathbf{r}_{l,\text{shell}} - \mathbf{r}_{l,\text{core}}),$$

where $\mathbf{F}_{l,\text{core}}^{\text{elect}}$ is the electrostatic force on a core l, and $\mathbf{r}_{l,,\text{core}}$ and $\mathbf{r}_{l,\text{shell}}$ are the respective core and shell positions. Since the force on a shell must in principle be zero, $\mathbf{F}_{l,\text{core}}$ should correspond to the force on the lth ion $\mathbf{F}_{l,\text{ion}}$, estimated from DFT. Thus, the shells must be allocated such that

$$\mathbf{r}_{l,\text{shell}} - \mathbf{r}_{l,\text{core}} + (\mathbf{F}_{l,\text{ion}} - \mathbf{F}_{l,\text{core}}^{\text{elec}})/k_l = 0, \quad \text{for } l = 1, \ldots, L. \tag{30}$$

$\mathbf{F}^{\mathrm{elec}}_{l,\mathrm{core}}$ depends only on the charge distribution and atom positions. Hence, expression (30) represents a system of $3L$ force field-independent equations and $3L$ unknown variables for the shells coordinates. In this work, the charge distribution and spring constants were taken from [74]. The equations in (30) were solved by a simple fixed-point iteration scheme. The schematic of the method is presented in Fig. 5.

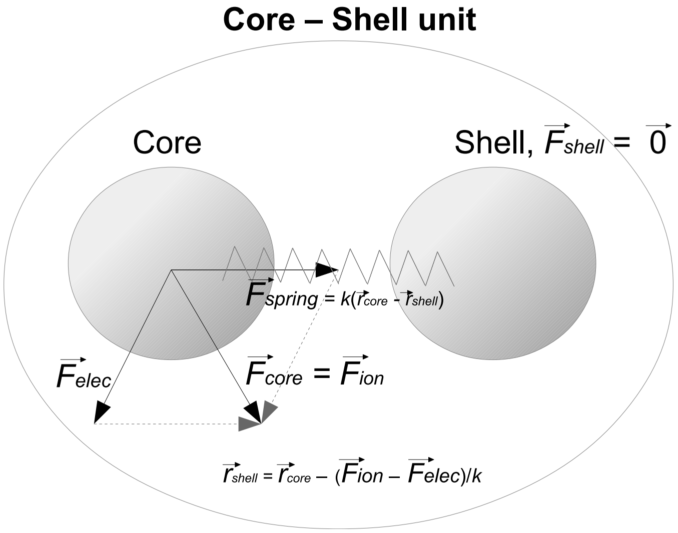
Core – Shell unit
Core
Shell, $\vec{F}_{shell} = \vec{0}$
$\vec{F}_{spring} = k(\vec{r}_{core} - \vec{r}_{shell})$
$\vec{F}_{elec}$
$\vec{F}_{core} = \vec{F}_{ion}$
$\vec{r}_{shell} = \vec{r}_{core} - (\vec{F}_{ion} - \vec{F}_{elec})/k$

Fig. 5 Correct positioning of the shell within a core-shell unit. If the charges and the spring constant k are set, the fact that dispersive forces do not act on the core uniquely determines the exact position of the shell

Validation

The validation involved a comparison of experimental data *outside* the TD with the results obtained through atomistic simulations incorporating the proposed force field. To simulate the present system we chose the RSM-GSHMC simulation method, which helped to reduce potential negative effects on the kinetic properties of the system caused by the use of the cs-adi model. We further combined RSM-GSHMC with the MAIA integrator (Sect. 3.4) to achieve the highest possible integration accuracy when using two-stage splitting methods, including Verlet.

The final set of parameters of $\mathrm{Na}_x\mathrm{FePO}_4$-ff can be found in [18]. Fig. 6 compares the accuracy of $\mathrm{Na}_x\mathrm{FePO}_4$-ff and Whiteside-ff. Fig. 6a depicts the relative error ×100 for the unit cell parameters (a, b, c) with respect to experimental measurements at room temperature [66]. In the absence of Na^+ ($x = 0$), Whiteside-ff leads to numerical instability and values cannot be extracted. For $x = 0.66$ and $x = 0.83$ both force fields display comparable levels of accuracy, and at $x = 1.0$, $\mathrm{Na}_x\mathrm{FePO}_4$-

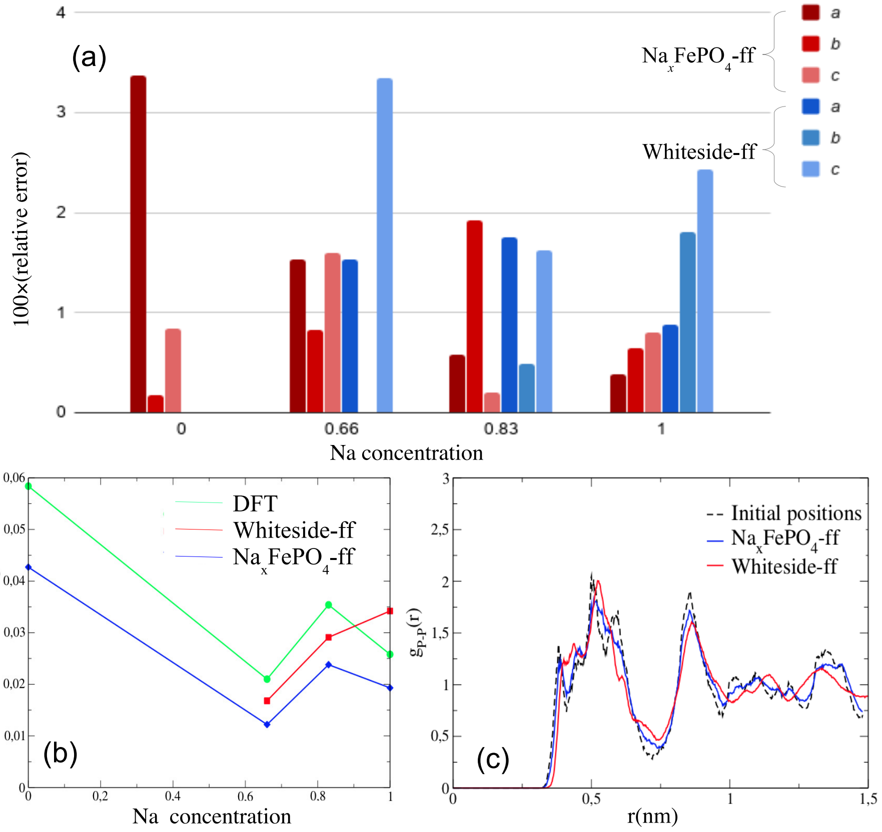

Fig. 6 Force field validation. All simulated values in this figure were obtained by combining Na$_x$FePO$_4$-ff or Whiteside-ff with the RSM-GSHMC technique. In Whiteside-ff, the parameters involving Fe^{3+} were taken to be the same as those for Fe^{2+}. **a** Relative error ×100 for the unit cell parameters (a, b, c) with respect to experimental measurements at room temperature [66]. **b** Relative volume divergence with respect to experimental results at room temperature. **c** Radial distribution functions for phosphorous atoms in a Na$_{0.58}$FePO$_4$ system after 10 ps at 300 K. The dashed line corresponds to the initial positions, which is the same for both force fields

ff is clearly better. When comparing the volume divergence at room temperature, Fig. 6b shows that Na$_x$FePO$_4$-ff is clearly superior. The DFT data is calculated at 0 K and thus, its divergence is high. However, it is clear that Na$_x$FePO$_4$-ff follows the same trend as the DFT values. Finally, Fig. 6c presents the radial distribution functions between phosphorous atoms, $g_{P-P}(r)$, in the Na$_{0.58}$FePO$_4$ system after 10 ps at 300 K. The dashed line corresponds to the initial configurations, which are the same for both force fields. After 10 ps of simulation, there is a distortion in the position of the peaks for Whiteside-ff with respect to the initial $g_{P-P}(r)$, evidencing the additional repulsion between oxygens of adjacent tetrahedrals when a Na$^+$ ion is removed. This is an anomalous effect not observed with Na_xFePO_4-ff due to the reoptimized parameters. Such an anomaly leads to errors in the diffusivities when using Whiteside-ff.

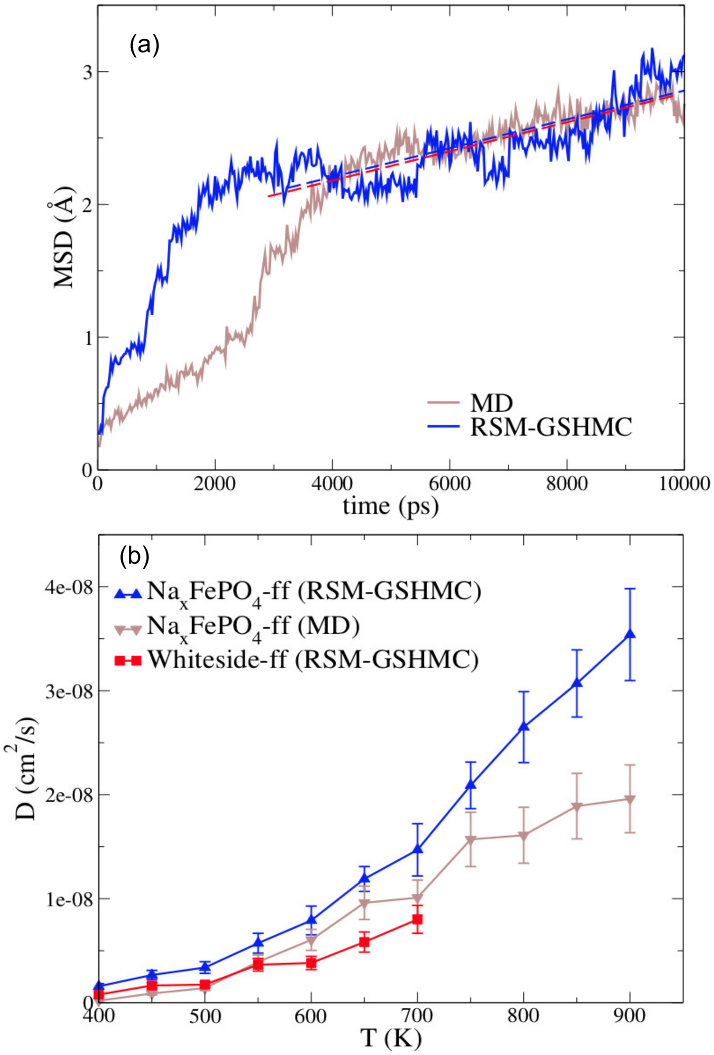

Fig. 7 **a** MSD at 700 K in the main diffusion direction, y, using both traditional MD and RSM-GSHMC combined with Na_xFePO_4-ff. Convergence to the equilibrium slope is reached faster through RSM-GSHMC. **b** Three-dimensional diffusion coefficient as a function of temperature (legends indicate force field and sampler combinations). (Adapted with permission from [18]. Copyright (2018) American Chemical Society)

5.1.2 MD vs. RSM-GSHMC

Fig. 7a presents the MSD for Na^+ at 700 K in the main diffusion direction, y, obtained from both traditional MD and RSM-GSHMC. Convergence to the equilibrium slope that provides the diffusion coefficient through equation (1) is reached nearly 2 ns quicker through RSM-GSHMC, demonstrating its ability to sample the configurational space more efficiently than MD. The noise in the MSD curves in Fig. 7 is at least partially due to the fact that a fraction of the Na^+ atoms do not perform any jump or jump only once. While the situation could be improved by extending the simulations, the cs-adi model severely limits the simulation time-step, and further increasing the simulation time comes with a significant increase in the computation time.

Fig. 7b shows the three-dimensional diffusion coefficients of Na^+, $D = \frac{1}{3}(D_x + D_y + D_z)$, for a range of temperatures using RSM-GSHMC with both force fields. We also plot results for similar simulations using standard MD with Na_xFePO_4-ff. Comparison of the three curves suggests that not only does the force field affects the diffusion coefficients but the sampling efficiency of the chosen sampler also does. Indeed, at higher temperatures the computed diffusions using MD and RSM-GSHMC combined with the same force field (for equal simulation times) diverge visibly. This is a consequence of the higher sampling efficiency of GSHMC methods [2]. For longer simulation times, both methods are expected to converge [32].

5.1.3 Mechanism of Na^+ Diffusion

Because RSM-GSHMC does a good job in preserving dynamical information, it allows an atom-eye perspective of the physics behind the diffusion dynamics of Na^+ in Na_xFePO_4. We confirmed that the main transport mechanism involves single Na-ion hops through the one-dimensional channels along the y axis (the [010] crystallographic direction). Furthermore, we identified the novel Na-ion diffusion dynamics involving the formation and annihilation of Na^+/Fe^{2+} antisite defects (the exchange of positions between Na^+ and Fe^{2+} ions), which effectively facilitate the migration of Na-ions between adjacent y-oriented channels (see Fig. 8 for a detailed explanation). Na^+/Fe^{2+} antisite defects were expected to block y channels and, therefore, hinder Na-ion mobility in the bulk material. In contrast, our results revealed that such defects can indeed favor Na-ion exchange between parallel y channels. This result is in general agreement with simulations of two-dimensional diffusion in $LiFePO_4$ [69]: occasional jumps of charge carriers in the y axis can overcome (at a cost of reduced diffusivity) the damaging effect of antisite defects.

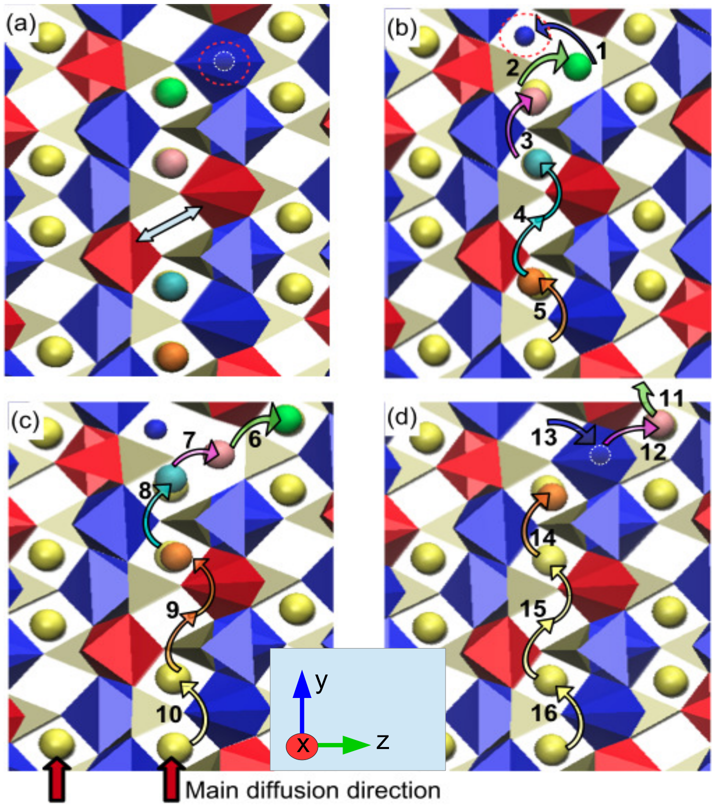

Fig. 8 Two-dimensional trajectory at 700 K. At this temperature, the coordination polyhedra are slightly distorted. The red octahedra are $Fe^{3+}O_6$ units, the blue octahedra are $Fe^{2+}O_6$ units, and the tetrahedrals correspond to $(PO_4)^{-3}$ units. The green, pink, cyan and orange atoms are tagged Na^+ ions; the remaining Na^+ ions are colored in yellow. The numbers in the figure specify the order in which atomic jumps occurred. **a** Initial, defect-free configuration. The double arrow shows the diametrically opposed Fe^{3+} ions around an extrinsic vacancy. **b** Antisite-like defect. A Fe^{2+} atom (encircled in red in **a** and **b**) moves to the adjacent extrinsic vacancy (jump 1), allowing the green Na^+ to jump to the empty Fe^{2+} site (jump 2). **c** Since the central diffusion channel is blocked, the green Na^+ transfers to an adjacent channel (jump 6), leading to the cascade of jumps 7 – 10. **d** After the defect in the Fe^{2+} site is removed (jump 13), one-dimensional diffusion is reestablished. Notice that Na^+ diffuses through the extrinsic vacancies in single, long curvilinear jumps (4, 9, and 15). (Adapted with permission from [18]. Copyright (2018) American Chemical Society)

5.2 Li^+ diffusion in cubic Ga/Al substituted $Li_7La_3Zr_2O_{12}$ (LLZO) solid electrolyte

LLZO garnet is a strong solid electrolyte contender for all-solid-state LIBs, which can in principle eliminate the safety problems associated with current liquid electrolyte-based batteries [9, 62]. LLZO garnets can crystallize in at least two different crystal structures, a poorly Li-ion conductive tetragonal structure (t-phase) and a significantly more conductive cubic one (c-phase) (Fig. 9). The c-phase structure contains partially occupied Li-sites (at any given instant they may or may not contain a Li^+) and is unstable at room temperature, but can be stabilized by substitution of, for instance, Li^+ by Al^{3+} and Ga^{3+} ions [64]. Despite having the same formal charge, gallium substitution is found to lead to higher room temperature conductivities ($\sim 10^{-3}$ S/cm) than aluminium substitution [62, 13, 63] ($\sim 10^{-4}$ S/cm). However, the origin and degree of this phenomenon is yet to be fully understood.

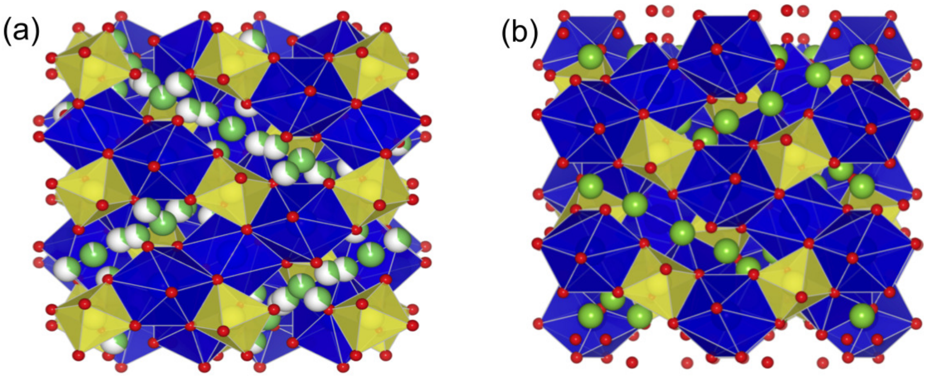

Fig. 9 Crystal structures of garnet-type LLZO. **a** c-phase. Li atoms are represented by partially filled spheres, indicating partial occupancy. **b** t-phase, containing an ordered distribution of fully occupied Li-sites (Li^+ ions are shown in green). La dodecahedra and Zr octahedra are shown in blue and yellow, respectively. The red spheres correspond to oxygen atoms. (Reproduced with permission from [19]. Copyright (2019) Elsevier)

We tackled this question using atomistic simulations, performed with traditional MD and our in-house GSHMC simulation method (Sect. 3.3). For this study, the chosen force field was based on that from Jalem et al. [47], but further adjusted with respect to experimental structural data reported in [63]. Details on the fitting procedure can be found in [41]. Below, we will use the notation Ga_yAl_x-LLZO to denote $Li_{7-3(x+y)}Ga_yAl_xLa_3Zr_2O_{12}$.

5.2.1 MD vs. GSHMC

To compare the sampling efficiency of MD and GSHMC methods for the simulation of the substituted LLZO systems, we calculated the integrated autocorrelation functions ($IACF$) of the potential energy U

$$IACF^U = \sum_{l=0}^{K'} ACF(l), \tag{31}$$

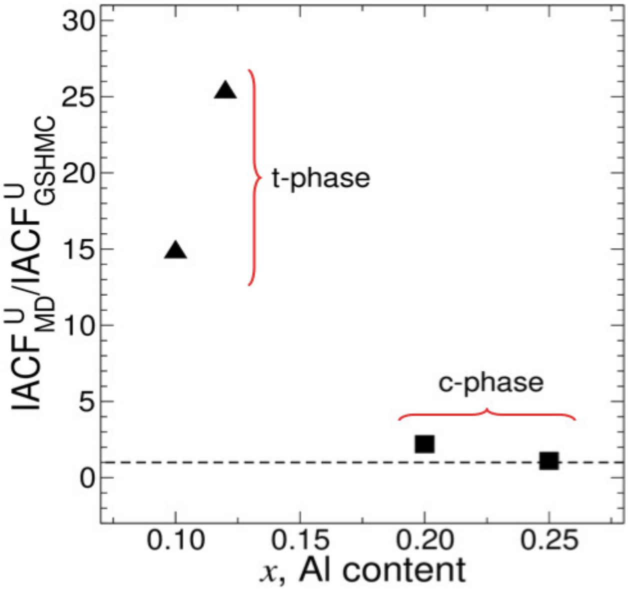

Fig. 10 Relative $IACF^U$ with respect to GSHMC at T = 298 K achieved when calculating the potential energy during the equilibration stage with the MD method. (Reproduced with permission from [19]. Copyright (2019) Elsevier)

where $ACF(l)$, $l = 0, \ldots, K \leq K'$ is the standard autocorrelation function for the time series U_k of K samples, $k = 1, \ldots, K$. Lower values of $IACF$ indicate a more efficient sampling, since it estimates the time required, on average, to generate an uncorrelated sample. In Fig. 10, we present the values of $IACF^U_{\mathrm{MD}}/IACF^U_{\mathrm{GSHMC}}$ for $\mathrm{Ga_0Al_{0.10}}$-LLZO and $\mathrm{Ga_0Al_{0.12}}$-LLZO in the t-phase and $\mathrm{Ga_0Al_{0.20}}$-LLZO and $\mathrm{Ga_0Al_{0.25}}$-LLZO in the c-phase at room temperature in the NVT ensemble. GSHMC simulations exhibit a sampling performance up to 25 times better than MD for the tested systems in the t-phase. In contrast, for the c-phase the performance of both methods are comparable, because the intersite energy barriers are relatively low. Based on this analysis, we have used GSHMC for simulations of the t-phase [41], as well as for simulations of the c-phase below room temperature, where the intersite barriers are much more significant [19]. Otherwise, we chose to use the conventional MD approach in order to minimize the efforts required for tuning of the GSHMC parameters.

5.2.2 Simulation Results

Length and Time Scales of Li$^+$ Diffusion in Cubic LLZO

A simple way to understand the Li$^+$ lattice in the c-phase is to consider it as a three-dimensional diffusion network, in which each node is connected to its closest

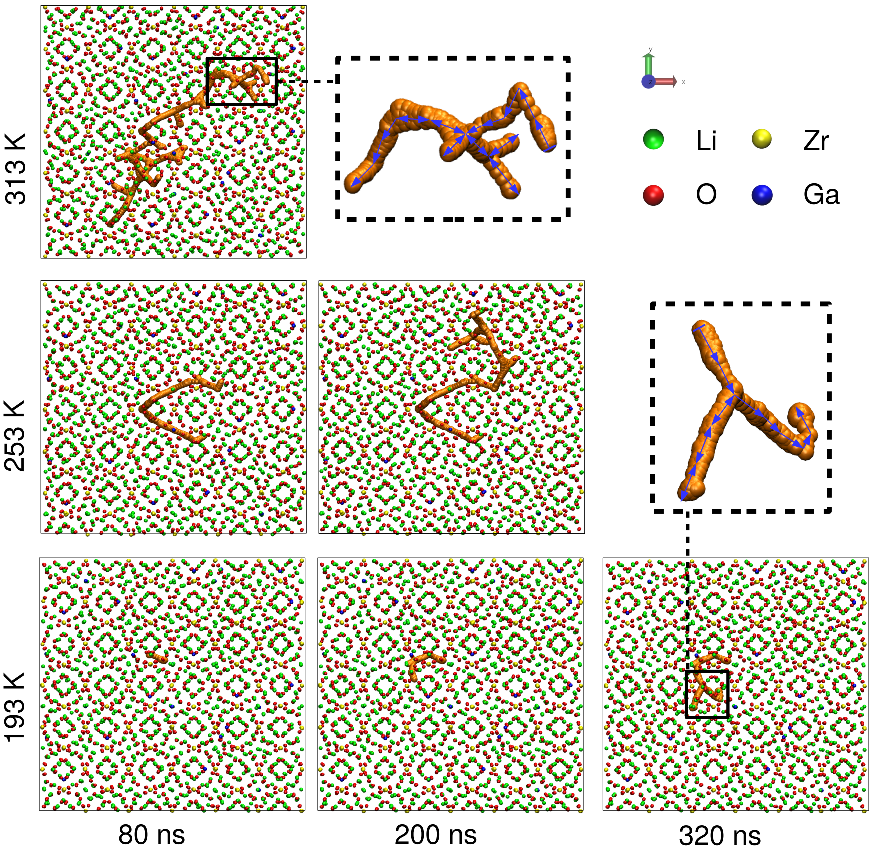

Fig. 11 Simulated trajectories of an individual Li^+ ion (orange) in the $Ga_{0.2}Al_{0.0}$-LLZO system at different temperatures. From *top* to *bottom* the temperature decreases, whereas from *left* to *right* the simulation time increases. Li, Zr, O and Ga ions are depicted in green, yellow, red and blue, respectively (La ions are not shown for simplicity). (Reproduced with permission from [41]. Copyright (2019) American Chemical Society)

four neighboring nodes through one intermediate site. Those intermediate sites are coordinated by six oxygen atoms, and thus are called Octahedral Li-sites, O_h. The nodes, in turn, are coordinated by four oxygen atoms, and thus are called tetrahedral sites, T_d. Hence, a typical Li^+ diffusion path involves a succession of jumps of the form $O_h \rightarrow T_d \rightarrow O_h \rightarrow T_d \rightarrow \ldots$. In Fig. 11, we provide an atom's eye view of Li-ion transport by considering a complete trajectory of a selected Li^+ ion across simulations at different temperatures and total time lengths.

Li-ion transport is isotropic. At 313 K, the selected Li^+ crosses frequently different T_d and O_h sites in a relatively short lapse of time (80 ns). In addition, we notice that a Li^+ ion can visit the same T_d site several times. This is revealed by the presence of branches centered in the nodal T_d sites along the diffusive path as shown in the inset of Fig. 11. When the temperature is decreased to 253 K, the distance covered by the Li^+ ion in 80 ns decreases. In fact, even after 200 ns the number of T_d and O_h sites visited has decreased considerably in comparison with the simulation at 313 K, resulting as well in a lower number of branches along the diffusive path. Finally, at 193 K we see that diffusion is severely hindered, with the Li^+ ion barely visiting more than one T_d site in less than 80 ns.

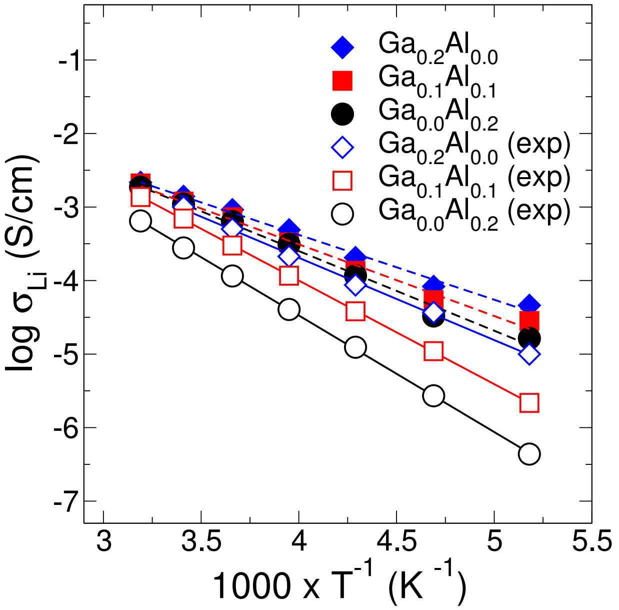

Fig. 12 Calculated Li-ion conductivity as a function of inverse temperature for $Ga_{0.2-x}Al_x$ ($x = 0.0, 0.1, 0.2$). Filled data points refer to simulation data. Experimental values, represented by empty symbols, were taken from Ref. [63]. Notice that the simulated data overestimates the experimental values, because the boundary resistances discussed in Sect. 2.2.1 cannot be easily incorporated. Those resistances are particularly high in $Ga_{0.0}Al_{0.2}$-LLZO, which has a powdery macrostructure. However, the trends in the simulated and experimental curves are clearly consistent

Variation of Li^+ Conductivity with Ga/Al in Cubic LLZO

Fig. 12 depicts the variation of the ionic conductivity σ_Li (calculated through (2)) with the inverse of temperature for $Ga_{0.2-x}Al_x$-LLZO and $x = 0.0$, 0.1 and 0.2. It is clear that for a constant concentration of substituents, increasing the ratio of Ga^{3+} to Al^{3+} has a positive effect on the conductivity. However, as temperature increases, the difference between Ga^{3+} and Al^{3+} vanishes, in agreement with recent experimental work [63]. This observation was explained on the grounds that a Ga^{3+} content above 0.15 leads to a phase change, in which shortcuts to the typical $O_\mathrm{h} \rightarrow T_\mathrm{d} \rightarrow O_\mathrm{h} \rightarrow \ldots$

diffusion path appear. However, if this was the only explanation, the curves would not display the observed temperature dependance.

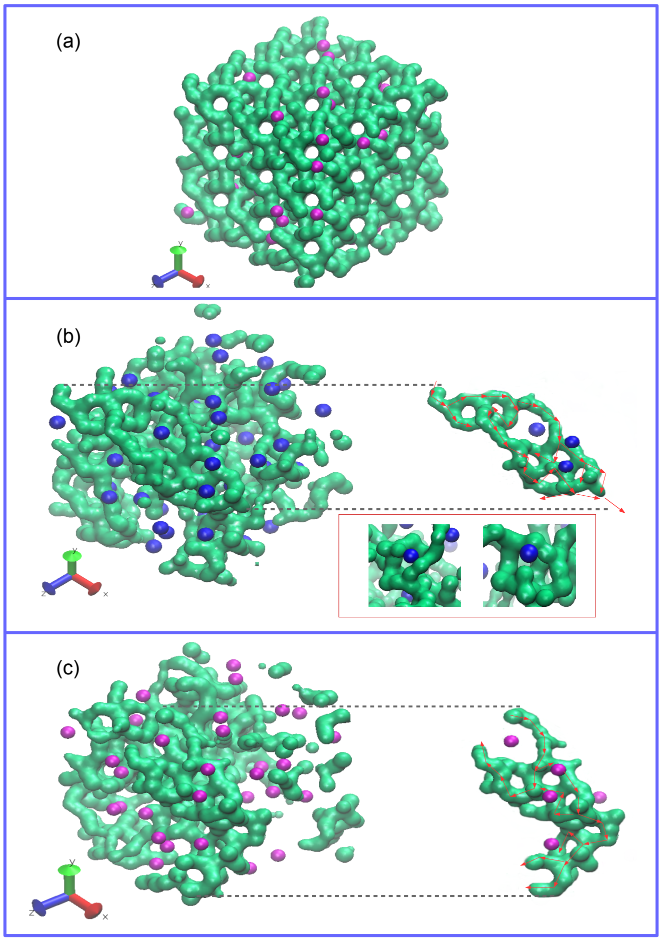

Fig. 13 Density volume maps at 233 K for all Li^+ ions in $Ga_{0.0}Al_{0.2}$-LLZO **a**, and only those Li^+ ions jumping at least one T_d site-T_d site length after 190 ns (>4 Å) in $Ga_{0.2}Al_{0.0}$-LLZO **b** and $Ga_{0.0}Al_{0.2}$-LLZO **c**. Al^{3+} and Ga^{3+} ions are represented by magenta and blue spheres, respectively. **a** Can be interpreted as the total available volume for Li-ion diffusion, while **b** and **c** correspond to the active volume where diffusion is actually occurring. Clearly, active diffusion paths avoid the vicinity of Al^{3+} and Ga^{3+} (see insets to the *right* of **b** and **c**) but the disruption of diffusion paths in $Ga_{0.0}Al_{0.2}$-LLZO is much more severe. (Reproduced with permission from [41]. Copyright (2019) American Chemical Society)

Through a detailed analysis of the radial distribution functions (see [41] for details) we were able to determine that the local environment around Ga^{3+} and Al^{3+} cations disrupts active diffusion paths. Hence, regions within the crystal volume that do not contribute to Li-ion conduction arise. Fig. 13 depicts volume density maps of mobile Li$^+$ ions at 233 K in Ga$_{0.2}$Al$_{0.0}$-LLZO and Ga$_{0.0}$Al$_{0.2}$-LLZO, showing a depletion of active diffusion paths around Ga^{3+} and Al^{3+}. However, because of the particular parameters of the Al^{3+} - O^{2-} interaction potential, the disruptive effect of Al^{3+} is more pronounced. As temperature increases, Li$^+$ ions acquire enough thermal energy to overcome the barriers surrounding Ga^{3+} and Al^{3+} ions, eliminating the difference in performance between both cations.

Variation of conductivity in the experimental data for Ga$_{0.0}$Al$_x$-LLZO

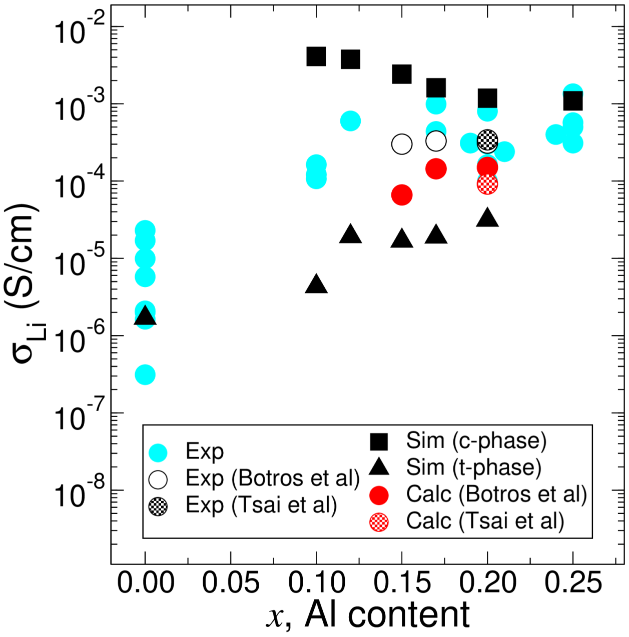

Fig. 14 Experimental and simulated Li-ion conductivities, σ_{Li}, of Al$_x$-LLZO for x = [0.0, 0.10, 0.12, 0.15, 0.17, 0.20, 0.25] at room temperature. Cyan circles correspond to experimental data taken from several experimental studies (the complete list of references can be found in the caption of Fig. 8 in [19]). Empty and dotted circles correspond to the experimental conductivities from Botros et al. [21] and Tsai et al. [71], respectively, who provide the percentages of c- and t-phases present in their samples. Filled squares (triangles) correspond to σ_{Li} for c-phase (t-phase) obtained in this work. Red filled and dotted circles correspond to the conductivities obtained by combining Bruggeman's model in (32) with the simulated σ_{Li} to represent the coexistence of t- and c-phases at the proportions reported in Botros et al. [21] and Tsai et al. [71]. (Reproduced with permission from [19]. Copyright (2019) Elsevier)

There is significant variability in the literature regarding the conductivity of Gal$_0$Al$_x$-LLZO. Given the recently reported gradients in Al content within a single sample

[79], one potential scenario is that Al-rich regions crystallize in the c-phase, while the Al-poor regions form the t-phase. The second potential scenario is that thermodynamic equilibrium can be established between the two phases at the same x. To the best of our knowledge, neither scenario can be categorically dismissed. Presuming that thermodynamic equilibrium exists between the c- and t-phases at some Al contents, we can estimate the proportion of each phase through an adequate mixing model. Assuming that both phases mix randomly and homogeneously at the microscopic scale ($\sim 10^{-6}$ m), Bruggeman's effective medium model [51] predicts that

$$\nu_c \frac{\sigma_c - \sigma_m}{\sigma_c + 2\sigma_m} + \nu_t \frac{\sigma_t - \sigma_m}{\sigma_t + 2\sigma_m} = 0, \tag{32}$$

where ν_i is phase volume fraction, σ_i is conductivity and $i = \{c, t, m\}$ denotes the cubic, tetragonal, and mixed phases, respectively. By equating the c-phase volume fraction ν_c to the experimental values, it is possible to determine the theoretical conductivity of the mixed c- and t-phases using (32). Considering the experimental works by Botros et al. [21] and Tsai et al. [71], who report weight percentages for mixed cubic/tetragonal samples, we estimated the corresponding mixed phase conductivities by using conductivities simulated over the range of Al content. The predictions are shown in Fig. 14 as red circles, while the measurements from [21] and [71] are highlighted through the empty circles. Our predictions are in good agreement with the conductivities provided by [21] and [71], reproducing the experimental results better than either of the pure phases at all Al contents. Thus, we provide a sound basis for the analysis of phase coexistence in substituted LLZO and similar polycrystalline solid electrolyte materials.

6 Conclusions

In this work, we present a brief review of the work performed in our team towards multiscale modelling and simulation of advanced battery materials. We have attempted to make the current material self-contained, and thus we have included sections on the basic science and engineering aspects of LIBs and NIBs (Sect. 2), as well as a short discussion of conventional simulation methods (Sects. 3.1 and 3.2) and force field development (Sect. 4). These sections can be selectively avoided by experienced practitioners.

We focus on the family of Modified Hamiltonian Monte Carlo (MHMC) methods recently introduced for atomistic simulations of complex systems [46, 68, 1, 2, 5, 32] and later adapted to multiscale [31] and mesoscale simulations [4]. In particular, we have found members of this family, the Generalized Shadow Hybrid Monte Carlo (GSHMC) method [2] and its latest extension Randomized Shell Mass GSHMC (RSM-GSHMC) [32], to be powerful tools for unveiling the structural, thermodynamic and transport properties of solid state battery materials. As application examples, we present our recent results in the modelling of ionic conduction in $NaFePO_4$ cathode material for NIBs and LLZO solid electrolyte material for LIBs

[32, 18, 19, 41]. Future work will focus on two key areas: probabilistic screening of advanced energy materials and multiscale simulation of composite materials.

Probabilistic Screening of Advanced Energy Materials

There is a considerable activity in the field of materials discovery using high-throughput computing combined with statistical screening approaches. However, incorporating macroscopic materials properties of practical importance, such as ionic conductivity, tensile moduli or phase transition temperatures into a materials design scheme in an efficient way remains a standing challenge.

Bayesian parameterization provides a principled way of combining prior information with data, within a solid decision theoretical framework. Past information and experimental data can readily be accounted for and form a prior distribution for future analysis. Moreover, when new observations become available, the previous posterior distribution can be used as a prior. More importantly, it provides interpretable answers from an experimental standpoint, such as "the optimal compound has a probability of 0.95 of falling in a 95 % credible interval".

Our in-house Bayesian parameterization technique, the Mix & Match Hybrid Monte Carlo (MMHMC) method [61], will be at the core of a novel framework that will combine incoming data from atomistic simulations and information from experiments, in order to provide the distribution of compositions that is most likely to maximize the expectation of a desired set of macroscopic properties (Fig. 15).

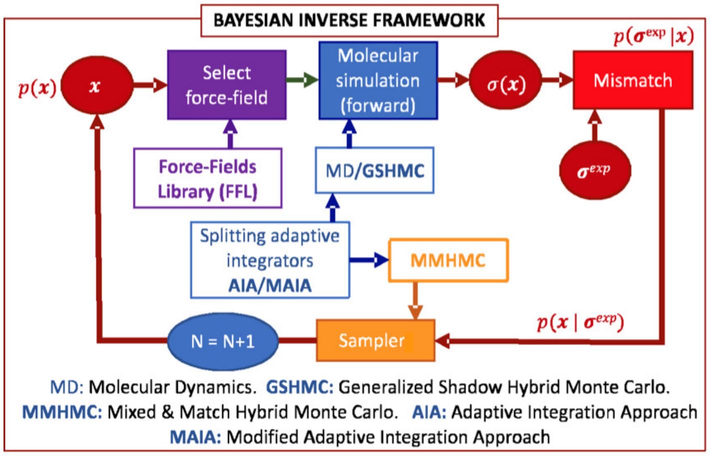
BAYESIAN INVERSE FRAMEWORK
$p(\sigma^{exp}|x)$
$p(x)$
x
Select force-field
Molecular simulation (forward)
$\sigma(x)$
Mismatch
Force-Fields Library (FFL)
MD/GSHMC
σ^{exp}
Splitting adaptive integrators AIA/MAIA
MMHMC
$p(x \mid \sigma^{exp})$
N = N+1
Sampler
MD: Molecular Dynamics. GSHMC: Generalized Shadow Hybrid Monte Carlo.
MMHMC: Mixed & Match Hybrid Monte Carlo. AIA: Adaptive Integration Approach
MAIA: Modified Adaptive Integration Approach

Fig. 15 One cycle in the proposed probabilistic materials screening approach. The number of iterations, *N*, the cycle is repeated, is set *a priori* by the user. The algorithms highlighted in bold were developed in our group and are discussed in the body of this manuscript

Multiscale Simulation of Composite Materials

In Sect. 2, we have briefly mentioned that one strategy to develop novel battery components with properties that no individual material has, is through the combination of several types of materials into segregated multiphase components called composites. Often, one of the phases is organic (e.g. a polymer) that provides mechanical flexibility, while the other is composed by inorganic particles (e.g. ionic crystals) providing good conductivity. The atomistic simulation of diffusional exchange between the various phases in composite materials is a relatively new subject. Despite the difficulty to experimentally examine the interface (and hence validate computational studies), simulations can provide very useful insights regarding the main barriers to particle exchange, and how the chemistry and environmental conditions affect them. Moreover, with advanced sampling methods such as GSHMC combined with splitting integrators, it is feasible to test many realizations of the polymer/crystal interface and systematically study the impact of contact defects. However, simulating an entire inorganic particle surrounded by polymer remains a great computational challenge that has been attempted only with very small nanoparticles. We are working on the coarse-grained simulation of such a system using our in-house simulation technique meso-GSHMC [4]. The transport coefficients and selectivities obtained this way can be incorporated in continuous or effective medium approaches, providing a multiscale modelling framework for the modelling of composite materials.

Acknowledgements We acknowledge the financial support by the Ministerio de Economia y Competitividad (MINECO) of the Spanish Government through grants MTM2016-76329-R (AEI/FEDER, EU) and BCAM Severo Ochoa accreditation SEV-2017-0718. This work was supported by the BERC 2018e2021 Program and by ELKARTEK Program, grants KK-2018/00054 and KK-2019/00068, funded by the Basque Government. The research is partially supported by the Basque Government under the grant "Artificial Intelligence in BCAM number EXP. 2019/00432". This work has been possible thanks to the support of the computing infrastructure of the i2BASQUE academic network and the technical and human support provided by IZO-SGI SGIker of UPV/EHU and European funding (ERDF and ESF).

References

1. Akhmatskaya, E., Reich, S.: The targeted shadowing hybrid Monte Carlo (TSHMC) method. In: Leimkuhler, B., Chipot, C., Elber, R., Laaksonen, A., Mark, A., Schlick, T., Schütte, C., Skeel, R., (eds.) New Algorithms for Macromolecular Simulation, pp. 141–153. Springer, Heidelberg (2006)
2. Akhmatskaya, E., Reich, S.: GSHMC: An efficient method for molecular simulation. J. Comp. Phys. **227**, 4934–4954 (2008)
3. Akhmatskaya, E., Bou-Rabee, N., Reich, S.: A comparison of generalized hybrid Monte Carlo methods with and without momentum flip. J. Comp. Phys. **228**, 2256–2265 (2009)
4. Akhmatskaya, E., Reich, S.: Meso-GSHMC: A stochastic algorithm for meso-scale constant temperature simulations. Proc. Comp. Sci. **4**, 1353–1362 (2011)
5. Akhmatskaya, E., Reich, S.: New Hybrid Monte Carlo methods for efficient sampling: from physics to biology and statistics. Prog. Nucl. Sci. Technol. **2**, 447–462 (2012)

6. Akhmatskaya, E., Fernández-Pendás, M., Radivojević, T., Sanz-Serna, J.M.: Adaptive Splitting Integrators for Enhancing Sampling Efficiency of Modified Hamiltonian Monte Carlo Methods in Molecular Simulation. Langmuir **33**, 11530–11542 (2017)
7. Amin, R., Balaya, P., Maier, J.: Anisotropy of Electronic and Ionic Transport in LiFePO4 Single Crystals. J. Power Sources. **10**, A13–A16 (2007)
8. Amores, M., Ashton, T.E., Baker, P.J.,Cussen, E.J., Corr, S.A.: Fast Microwave-Assisted Synthesis of Li-Stuffed Garnets and Insights into Li Diffusion from Muon Spin Spectroscopy. J. Mater. Chem. A **4**, 1729–1736 (2016)
9. Awaka, J., Kijima, N., Hayakawa, H., Akimoto, J.: Synthesis and structure analysis of tetragonal Li $_7$La $_3$Zr $_2$O $_{12}$ with the garnet-related type structure. J. of Solid State Chem. France **182**, 2046–2052 (2009)
10. Barsoukov, E., Macdonald, J.R.: Impedance Spectroscopy: Theory, Experiment, and Applications. Kluwer, Boston (2005)
11. Bauer, A., Song, J., Vail, S., Pan, W., Barker, J., Lu, Y.: The Scale-up and Commercialization of Nonaqueous Na-Ion Battery Technologies. Adv. Ener. Mat. **8**, 1702869 (2018)
12. Berendsen, H., van der Spoel, D., van Drunen, R.: GROMACS: A message-passing parallel molecular dynamics implementation. Comp. Phys. Comm. **91**, 43–56 (1995)
13. Bernuy-Lopez, C., Manalastas, W., Lopez del Amo, J.M., Aguadero, A., Aguesse, F., Kilner, J.A.: Atmosphere Controlled Processing of Ga-Substituted Garnets for High Li-Ion Conductivity Ceramics CHEM. Mat. **26**, 3610–3617 (2014)
14. Beskos, A., Pillai, N., Roberts, G., Sanz-Serna, J.M., Stuart, A.: Optimal tuning of the hybrid Monte Carlo algorithm. Bernoulli **19**, 1501–1534 (2013)
15. Biresselioglu, M.E., Kaplan, M.D., Yilmaz, B.K.: Electric mobility in Europe: A comprehensive review of motivators and barriers in decision making processes. Transp. Res. A: Policy Pract. **109**, 1–13 (2018)
16. Blanes, S., Casas, F., Sanz-Serna, J.M.: Numerical Integrators for the Hybrid Monte Carlo Method. SIAM J. Sci. Comput. **36**, A1556–A1580 (2014)
17. Blanes, S., Casas, F.: A Concise Introduction to Geometric Numerical Integration. Chapman and Hall (2016)
18. Bonilla, M. R., Lozano, A., Escribano, B., Carrasco, J., Akhmatskaya, E.: Revealing the Mechanism of Sodium Diffusion in Na_xFePO_4 Using an Improved Force Field. J. Phys. Chem. C. **122**, 8065–8075 (2018)
19. Bonilla, M. R., García Daza, F. A., Carrasco, J., Akhmatskaya, E.: Exploring Li-ion conductivity in cubic, tetragonal and mixed-phase Al-substituted $Li_7La_3Zr_2O_{12}$ using atomistic simulations and effective medium theory. Acta Materiala **175**, 426–435 (2019)
20. Borodin, O., Smith, G.D.: Na-ion mobility in layered Na2FePO4F and olivine Na[Fe,Mn]PO4. Macromolecules **39**, 1620–1629 (2006)
21. Botros, M., Djenadic, R., Clemens, O., Möller, M., Hahn, H.: Assisted Sintering of Fine-Grained $Li_{7-3x}La_3Zr_2Al_xO_{12}$ Solid Electrolyte and the Influence of the Microstructure on the Electrochemical Performance. J. Power Sources **309**, 108–115 (2016)
22. Brommer, P., Gähler, F.: Potfit: effective potentials from ab initio data. Model. Simul. Mater. Sci. Eng **15**, 295 (2007)
23. Brommer, P., Kiselev, A., Schopf, D., Beck, P.: Classical interaction potentials for diverse materials from ab initio data: a review of potfit. Model. Simul. Mater. Sci. Eng **23**, 074002 (2015)
24. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M.: CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. **4**, 187–217 (1983)
25. Crabtree, G., Kocs, E., Trahey, L.: The energy-storage frontier: Lithium-ion batteries and beyond. MRS Bulletin **40**,1067–1076 (2015)
26. Dick, B.G., Overhauser, A.W.: Theory of the Dielectric Constants of Alkali Halide Crystals. Phys. Rev. **112**, 90–103 (1958)
27. Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P.: A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem. **24**, 1999–2012 (2003)

28. Duane, S., Kennedy, A.D., Pendleton, B. J., Roweth, D.: Hybrid Monte Carlo. Phys. Lett. B **195**, 216–222 (1987)
29. Duin, A. C.T., Dasgupta, S., Lorant, F., Goddard, W.A.: ReaxFF: A Reactive Force Field for Hydrocarbons. J. Phys. Chem. A **41**, 9396–9409 (2001)
30. Escribano, B., Akhmatskaya, E., Mujika, J. I.: Combining stochastic and deterministic approaches within high efficiency molecular simulations. Centr. Eur. J. Math. **11**, 787–799 (2013)
31. Escribano, B., Akhmatskaya, E., Reich, S., Azpiroz, J. M.: Multiple-time-stepping generalized hybrid Monte Carlo methods. J. Comp. Phys. **280**, 1–20 (2015)
32. Escribano, B., Lozano, A., Radivojević, T., Fernández-Pendás, M., Carrasco, J., Akhmatskaya, E.: Enhancing sampling in atomistic simulations of solid-state materials for batteries: a focus on olivine $NaFePO_4$. Theo. Chem. Acc. **136** (2017)
33. Español, P., Warren, P.: Statistical mechanics of dissipative particle dynamics. Europhys. Lett. **30**, 191–196 (1995)
34. Famprikis, T., Canepa, P., Dawson, J.A., Islam, M.S.: Fundamentals of inorganic solid-state electrolytes for batteries Nat. Mater. **18**, 1278–1291 (2019)
35. Fernández-Pendás, M., Escribano, B., Radivojević, T., Akhmatskaya, E.: Constant pressure hybrid Monte Carlo simulations in GROMACS. J. Mol. Mod. **20**, 2487 (2014)
36. Fernández-Pendás, M., Akhmatskaya, E., Sanz-Serna, J.M.: Adaptive multi-stage integrators for optimal energy conservation in molecular simulations. J. Comp. Phys. **327**, 434–449 (2016)
37. Franco A.: Multiscale modelling and numerical simulation of rechargeable lithium ion batteries: concepts, methods and challenges. RSC Adv. **3**, 13027–13058 (2013)
38. Franco, A. A., Rucci, A., Brandell, D., Frayret, C., Gaberscek, M., Jankowski, P., Johansson, P.: Boosting rechargeable batteries R&D by multiscale modelling: Myth or reality?. Chem. Rev. **119**, 4569 (2019)
39. Frenkel, D., Smit, B.: Understanding Molecular Simulation: From Algorithms to Applications. 2nd edn. Academic Press (2002)
40. Fumi, F.G., Tosi, M.P.: Ionic sizes and born repulsive parameters in the NaCl-type alkali halides I. J. Phys. Chem. Solids **25**, 31–43 (1964)
41. García Daza, F. A., Bonilla, M. R., Llordés, A., Carrasco, J., Akhmatskaya, E.: Atomistic Insight into Ion Transport and Conductivity in Ga/Al-Substituted $Li_7La_3Zr_2O_{12}$ Solid Electrolytes. ACS Appl. Mater. Interfaces **11** (2019)
42. Hastings, W.K.: Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika **57**, 97–109 (1970)
43. Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E.: GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theo.Comp. **4**, 435–447 (2008)
44. Horowitz, A. M.: A generalized guided Monte Carlo algorithm. Phys. Lett. B **268**, 247–252 (1991)
45. Izaguirre, J. A., Reich, S., Skeel, R. D.: Longer time steps for molecular dynamics. J. Chem. Phys. **110**, 9853–9864 (1999)
46. Izaguirre, J. A., Hampton, S. S.: Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules. J. Comp. Phys. **200**, 581–604 (2004)
47. Jalem, R., Rushton, M.J.D., Manalastas, W., Nakayama, M., Kasuga, T., Kilner, J.A., Grimes, R.W.: Effects of Gallium Doping in Garnet-Type $Li_7La_3Zr_2O_{12}$ Solid Electrolyte. CHem Mat. **27**, 2821–2831 (2015)
48. Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J.: Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. **118**, 11225–11236 (1996)
49. Kennedy, A.D., Pendleton, B.: Cost of the generalised hybrid Monte Carlo algorithm for free field theory. Nuclear Physics B **607**, 456–510 (2001)
50. Kresse, G., Hafner, J.: Ab initio molecular dynamics for liquid metals. Phys. Rev. B **47**, 558 (1993)
51. Landauer, R.: Electrical Conductivity in Inhomogeneous Media J. AIP Conference Proceedings **40**, 2–45 (1978)

52. Leimkuhler, B., Reich, S.: Simulating Hamiltonian dynamics. In: Cambridge Monographs on Applied and Computational Mathematics Vol 14. University Press, Cambridge (2004)
53. Leimkuhler, B., Matthews, C.: Molecular Dynamics with Deterministic and Stochastic Numerical Methods. In: Interdisciplinary Applied Mathematics Vol 39. Springer (2015)
54. Li, M., Lu, J., Chen, C., Amine, K.: 30 Years of Lithium-Ion Batteries. Adv. Mater. **30**, 1800561 (2018)
55. Mitchell, P.J., Fincham, D.: Shell model simulations by adiabatic dynamics. J. Phys.: Condens. Matter **5**, 1031 (1993)
56. Murch, G.: The haven ratio in fast ionic conductors. Solid State Ion. **7**, 177–198 (1982)
57. Oganov, A.R., Kvashnin, A.G., Saleh, G.: Computational Materials Discovery: Dream or Reality? in Oganov, A.R., Saleh, G., Kvashnin, A.G. (eds) Computational Materials Discovery, pp. 1-14. The Royal Society of Chemistry, London (2019)
58. Park, M., Zhang, X., Chung, M., Less, G.B., Sastry, A.M.: A review of conduction phenomena in Li-ion batteries. J. Power Sources **195**, 794–7929 (2010)
59. Plummer, M., Best, N., Cowles, K., Vines, K.: CODA: Convergence Diagnosis and Output Analysis for MCMC. R News **6**, 7–1 (2006)
60. Radivojević, T., Fernández-Pendás, M., Sanz-Serna, J.M., Akhmatskaya, E.: Multi-stage splitting integrators for sampling with modified Hamiltonian Monte Carlo methods. J. Comp. Phys. **373**, 900–916 (2018)
61. Radivojević, T., Akhmatskaya, E.: Modified Hamiltonian Monte Carlo for Bayesian inference. Stat. and Comp. **30**, 377–404 (2020)
62. Ramaswamy. M., Venkataraman, T., Werner, W.: Fast Lithium Ion Conduction in Garnet-Type $Li_7La_3Zr_2O_{12}$. Angewandte Chemie. Edition: International **46**, 7778—7781 (2007)
63. Rettenwander, D., Redhammer, G., Preishuber-Pflugl, F., Cheng, L., Miara, L., Wagner, R., Welzl, A., Suard, E., Doeff, M.M., Wilkening, M., Fleig, J., Amthauer, G.: Structural and Electrochemical Consequences of Al and Ga Cosubstitution in $Li_7La_3Zr_2O_{12}$ Solid Electrolytes. Chem. Mat. **28**, 2384–2392 (1998)
64. Rettenwander, D., Geiger, C.A., Tribus, M., Tropper, P., Amthauer, G.: A Synthesis and Crystal Chemical Study of the Fast Ion Conductor $Li_7La_3Zr_2O_{12}$ with $x = 0.08$ to 0.84 in $Li_7La_3Zr_2O_{12}$ Solid Electrolytes. Chem. Mat. **53**, 6264–6269 (2014)
65. Sanz-Serna, J.M., Calvo, M.P.: Numerical Hamiltonian problems. Chapman and Hall (1994)
66. Saracibar, A., Carrasco, J., Saurel, D., Galceran, M., Acebedo, B., Anne, H., Lepoitevin, M., Rojo, T., Cabanas, M.C.: Investigation of sodium insertion-extraction in olivine Na_xFePO_4 $(0 \leq x \leq 1)$ using first-principles calculations. Phys. Chem. Chem. Phys. **18**, 13045–13051 (2016)
67. Schlick, T., Mandziuk, M., Skeel, R. D., Srinivas, K.: Nonlinear Resonance Artifacts inMolecular Dynamics Simulations. J. Comp. Phys. **140**, 1–29 (1998)
68. Sweet, C. R., Hampton, S. S., Skeel, R. D., Izaguirre, J. A.: A separable shadow Hamiltonian hybrid Monte Carlo method. J. Chem. Phys. **131**, 174106 (2009)
69. Tealdi, C., Spreafico, C., Mustarelli, P.: Lithium diffusion in $Li_{1-x}FePO_4$: the effect of cationic disorder. J. Mater. Chem. **22**, 24870–24876 (2012)
70. Tripathi, R. Wood, S. M., Islam, M. S. Nazar, L. F.: Na-ion mobility in layered Na_2FePO_4F and olivine Na[Fe,Mn]PO4. Energ. Environ. Sci. **6**, 2257–2264 (2013)
71. Tsai, C.L., Dashjav, E., Hammer, E.M., Finsterbusch, M., Tietz, F., Uhlenbruck, S., Buchkremer, H.P.: High Conductivity of Mixed Phase Al-Substituted $Li_7La_3Zr_2O_{12}$Nanoscale **35**, 25–32(2015)
72. Wang, J., Xueliang, S.: Olivine $LiFePO_4$: the remaining challenges for future energy storage. Energy Environ. Sci. **8**, 1110–1138 (2015)
73. Wee, C.L., Sansom, M.S.P., Reich, S., Akhmatskaya, E.: Improved Sampling for Simulations of Interfacial Membrane Proteins: Application of Generalized Shadow Hybrid Monte Carlo to a Peptide Toxin/Bilayer System. J. Phys. Chem. B **112**, 5710–5717 (2008)
74. Whiteside, A., Fisher, C.A.J., Parker, S.C., Islam, M.S.: Particle shapes and surface structures of olivine $NaFePO_4$ in comparison to $LiFePO_4$. Phys. Chem. Chem. Phys. **39**, 21788–21794 (2014)

75. Whittingham M.S.: Electrical Energy Storage and Intercalation Chemistry. Science **192**, 1126-1127 (1976)
76. Wu, J.F., Chen, E.Y., Yu, Y., Liu, L., Wu, Y., Pang, W.K., Peterson, V.K., Guo,: X. Gallium-Doped $Li_7La_3Zr_2O_{12}$ Garnet-Type Electrolytes with High Lithium-Ion Conductivity. ACS Appl. Mater. Interfaces **9**, 1542–1552 (2017)
77. Yabuuchi, N., Kubota, K., Dahbi, M., Komaba, S.: Research Development on Sodium-Ion Batteries. Chem. Rev. **114**, 11636–11682 (2014)
78. Yan, T., Burnham, C.J., Del Pópolo, M.G., Voth, G.A.: Molecular Dynamics Simulation of Ionic Liquids: The Effect of Electronic Polarizability. J. Phys. Chem. B. **108**, 11877–11881 (2004)
79. Wachter-Welzl, A., Kirowitz, J., Wagner, R., Smetaczek, S., Brunauer, G.C., Bonta, M., Rettenwander, D., Taibl, S., Limbeck, A., Amthauer, G., Fleig, J.: The Origin of Conductivity Variations in Al-Stabilized $Li_7La_3Zr_2O_{12}$ Ceramics. Solid State Ionics **319**, 203–208 (1978)
80. Zheng, F., Kotobuki, M., Song, S., Lai, Man, O., Lu, L.: Review on solid electrolytes for all-solid-state lithium-ion batteries. J. Power Sources. **389**, 198–213 (2018)
81. Zhu, Y., Xu, Y., Liu, Y., Luo, C., Wang, C.: Comparison of electrochemical performances of olivine $NaFePO_4$ in sodium-ion batteries and olivine $LiFePO_4$ in lithium-ion batteries. Nanoscale **5**, 780–787 (2013)