Enhancing sampling in atomistic simulations of solid state materials for batteries: a focus on olivine NaFePO$_4$

Bruno Escribano · Ariel Lozano · Tijana Radivojević ·
Mario Fernández-Pendás · Javier Carrasco · Elena Akhmatskaya

Abstract The study of ion transport in electrochemically active materials for energy storage systems requires simulations on quantum- atomistic- and meso-scales. The methods accessing these scales not only have to be effective but also well compatible to provide a full description of the underlying processes. We propose to adapt the Generalized Shadow Hybrid Monte Carlo (GSHMC) method to atomistic simulation of ion intercalation electrode materials for batteries. The method has never been applied to simulations in solid state chemistry but it has been successfully used for simulation of biological macromolecules, demonstrating better performance and accuracy than can be achieved with the popular molecular dynamics (MD) method. It has been also extended to simulations on meso-scales, making it even more attractive for simulation of battery materials. We combine GSHMC with the dynamical Core-Shell model to incorporate polarizability into the simulation and apply the new Modified Adaptive Integration Approach, MAIA, which allows for a larger time step due to its excellent conservation properties. Also, we modify the GSHMC method, without losing its performance and accuracy, to reduce the negative effect of introducing a shell mass within a dynamical shell model. The proposed approach has been tested on olivine NaFePO$_4$, which is a promising cathode material for Na-ion batteries. The calculated Na-ion diffusion and structural properties have been compared with the available experimental data and with the results obtained using MD and the original GSHMC method. Based on these tests, we claim that the new technique is advantageous over MD and the conventional GSHMC and can be recommended for studies of other solid-state electrode and electrolyte materials whenever high accuracy and efficient sampling are critical for obtaining tractable simulation results.

B. Escribano · T. Radivojević · M. Fernández-Pendás
Basque Center for Applied Mathematics, Alameda de Mazarredo 14 (48009) Bilbao, Bizkaia, Spain
E-mail: bescribano@bcamath.org

A. Lozano
Basque Center for Applied Mathematics, Alameda de Mazarredo 14 (48009) Bilbao, Bizkaia, Spain
CIC EnergiGUNE, Albert Einstein 48 (01510) Miñano, Álava, Spain
E-mail: alozano@bcamath.org

J. Carrasco
CIC EnergiGUNE, Albert Einstein 48 (01510) Miñano, Álava, Spain

E. Akhmatskaya
Basque Center for Applied Mathematics, Alameda de Mazarredo 14 (48009) Bilbao, Bizkaia, Spain
IKERBASQUE, Basque Foundation for Science, E-48013 Bilbao, Spain

1 Introduction

The development of advanced materials for energy storage has grown into a topic of intense research due to their importance in powering portable devices, electric vehicles, and electrical grids collecting energy from renewable sources. During the last decade, Li-ion rechargeable batteries have become a gold standard in storing electrical energy [1–4]. However, in an ever-growing demand for better batteries, low cost and natural abundance of precursor materials are quickly emerging as

the basis for *beyond Li-ion* technology. In this context, the fifth most abundant element in the earth crust and the second lightest and smallest alkali metal after Li, Na, is currently considered as a natural candidate for the next generation of low cost batteries [5, 6].

Therefore, research on active materials for Na-based technologies is gathering momentum [7]. Atomic-scale computational approaches are becoming increasingly useful for these exploratory studies in order to avoid time-consuming trial and error approaches [8, 9]. Based on ab initio methods and modern information technology tools, these techniques enable the assessment of critical properties of interest such as phase stability, electronic structure, and ionic conductivity [10–17]. However, the computation of some of these properties can be very time consuming and, therefore, impracticable to tackle from a pure ab initio viewpoint. This is particularly true for solid state ionics and ion intercalation processes in electrolyte and electrode materials for batteries, where one should deal with the mobility of alkali ions.

Typically, ion diffusion occurs over long timescales and its statistically meaningful study usually requires to model systems containing thousands of atoms at least. Such simulations are not currently feasible with the use of ab initio methods. Classical interatomic potentials (force fields) are a practical solution to this problem in many cases, since such methods reduce the electronic degrees of freedom and thus allow for handling longer timescales and larger system sizes. Many studies of electrolyte and electrode materials based on interatomic potentials have been performed in the last two decades, mainly dealing with statical energy evaluations to determine ion diffusion paths and activation energies, defect chemistry, and stability of surfaces and nanostructures [10].

In spite of such success, classical interatomic potentials are often inefficient to properly account for rare events, specially when they are applied in molecular dynamics (MD) simulations. Ion diffusion processes are indeed rare events that involve ion hopping between adjacent sites and, sometimes, even colective ion transport. In order to properly simulate such phenomena in technologically relevant materials, very long simulations are normally required (see, e.g., Ref. [18]). To overcome this issue, different proposals for enhancing sampling efficiency in MD simulations of ion diffusion have been reported. For example, one can modify the particles momenta on the fly to stimulate the events of interest (ion particles jumps), but these methods do not preserve the desired distribution [19]. A similar approach is to rely on very high unphysical temperatures to force the observation of rare diffusion events [20, 21]. The so-called Generalized Shadow Hybrid Monte Carlo method

(GSHMC) is another promising technique, which has been proven to be successful when applied to the study of rare events in complex biological processes [22–24], but it has not been used for computing properties of solid crystalline systems yet.

In this work we investigate the effectiveness of enhanced sampling approaches in the simulation of various properties of olivine $NaFePO_4$ using GSHMC-based techniques. We focused on $NaFePO_4$ because this system is a promising candidate as a cathode material for Na-ion batteries [25]. It is the Na counterpart of $LiFePO_4$, which is used in many commercial Li-ion batteries nowadays [26]. In contrast to the Li case, $NaFePO_4$, forms a stable partially sodiated structure $Na_{2/3}FePO_4$ upon charge [25, 27] or chemical Na intercalation [13, 28]. The $NaFePO_4$ and $Na_{2/3}FePO_4$ systems offer us the opportunity to test the GSHMC sampling approach in a technologically relevant material and they are complex enough to analyze the performance of different sampling techniques. In addition, a force field specifically developed for olivine $NaFePO_4$ already exists [29].

The paper is organized as follows. In Section 2 we describe the force field used to model the bulk $NaFePO_4$. Then, in Section 3 we summarize the basics of the GSHMC method and explain the additional modifications that we have introduced to the original proposal. Section 4 compares the efficiency in terms of accuracy and performance of two variants of GSHMC and the standard MD method to account for structural and dynamical properties of the bulk $NaFePO_4$ and $Na_{2/3}FePO_4$. Finally, conclusions are presented in Section 5.

2 Computational model

The force field proposed for olivine $NaFePO_4$ by Whiteside et al. [29] follows the Born model, with the addition of shells to some ions. The shell-model is introduced to describe the ionic polarization as suggested by Dick and Overhauser [30]. In this model an ion is described using a central core with a charge X and a shell of a charge Y. These two charges are balanced so that the sum of $(X + Y)$ is the same as the valence state of the ion. A core and a shell are coupled together in a *core-shell unit* via a harmonic potential, which allows the shell to move with respect to the core, thus simulating a dielectric polarization.

The total potential energy is given by

$$U = V_C + V_{BH} + V_{CS}, \tag{1}$$

where V_C stands for the long-range Coulomb interactions, V_{BH} is a Buckingham potential that models short-range repulsions and van der Waals forces between atoms,

and V_{CS} is the interaction within each core-shell unit. In Eq. (1), the Coulomb interactions are computed between every pair of charged particles in the system but not within a core-shell unit. The short-range potential is considered solely between shells when core-shell units are involved and V_{CS} is computed for each core-shell unit.

The terms in Eq. (1) are explicitly given by

$$V_{\mathrm{C}}(r_{ij}) = \frac{1}{4\pi\epsilon_0} \sum_{i,j=1}^{N} \frac{q_i q_j}{r_{ij}},$$

where ϵ_0 is the vacuum permitivity, r_{ij} is the distance between particles i and j, q_i and q_j are their respective charges and N is the number of particles,

$$V_{\mathrm{BH}}(r_{ij}) = \sum_{i,j=1}^{N} A_{ij} \exp\left(-\frac{r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6},$$

where A_{ij}, ρ_{ij}, and C_{ij} are positive constants defining the shapes of the repulsive and the attractive terms of the potential, and

$$V_{\mathrm{CS}}(r_l) = \sum_{l=1}^{L} \frac{1}{2}\, k_l\, r_l^2,$$

where k_l is the spring constant for the l-th core-shell unit, r_l is the displacement between the shell center and its core, and L is the total number of shells.

In the work by Whiteside et al. [29] an extra three-body bonding term for the O-P-O angles in the PO$_4$ tetrahedral units was also included. It takes the form of a harmonic angle-bending potential given by

$$V_{\mathrm{Ang}}(\theta_k) = \sum_{k=1}^{K} \frac{1}{2} k_{\mathrm{ang}}(\theta_k - \theta_0)^2,$$

where k_{ang} is the spring constant, θ_0 is the equilibrium bond angle, θ_k is the current value of the bond k, and K is the total number of angle interactions.

For this study, we took from Ref. [29] the full set of parameters defining the force field for olivine NaFePO$_4$ (Table 1).

At this point, it must be mentioned an important issue regarding molecular dynamics simulations based on a Core-Shell potential model. In the original Core-Shell model, shell particles are massless and the model requires them to be always at their optimal positions with zero forces [30]. When atomic motions are considered during dynamical simulations the shells should respond instantaneously to the motions of the cores. Two main approaches are found in the literature to deal with the integration of equations of motion in this case: the so-called *shell relaxation* (CS-min) scheme [31] and the *adiabatic shells* (CS-adi) method [32].

Table 1 Force field parameters for olivine NaFePO$_4$ taken from Ref. [29].

BH			
Interaction	A (eV)	ρ (Å)	C (eV Å^6)
Na$^+$ - O$^{2-}_{\mathrm{sh}}$	629.757635	0.317034	0.0
Fe$^{2+}_{\mathrm{sh}}$- O$^{2-}_{\mathrm{sh}}$	1105.2409	0.3106	0.0
P^{5+} - O$^{2-}_{\mathrm{sh}}$	897.2648	0.3577	0.0
O$^{2-}_{\mathrm{sh}}$ - O$^{2-}_{\mathrm{sh}}$	22764.3	0.149	44.53

CS			
Species	Core charge	Shell charge	k (eV Å^{-2})
Fe^{2+}	-0.997	2.997	19.26
O^{2-}	0.96	-2.96	65.0

Ang		
bond	k_{ang} (eV rad^{-2})	θ_0 (deg)
O$^{2-}_{\mathrm{sh}}$-P^{5+}-O$^{2-}_{\mathrm{sh}}$	1.322626	109.47

The CS-min approach consists of three steps: i) to calculate the forces on all cores with the shells fully relaxed; ii) to update the core positions using the forces; iii) to relax the shells for the new core positions [31]. The last step involves the energy minimization in the multidimensional space of shell configurations which turns to be a very computationally demanding task. The CS-adi scheme was proposed as a faster alternative to the CS-min method. In the CS-adi approach, a small fraction x of the ion mass is put on the shell, whereas the remaining (1-x) fraction belongs to the core. Then, all the particles positions propagate following the conventional MD technique [32]. Having sufficiently small masses, the shells adiabatically follow the cores motion during the simulation. A proper choice of the mass distribution for the core-shell units is crucial for the accuracy of the method. Care has to be taken to ensure the negligible effect of an extra thermal energy, introduced by the relative motion between a core and its shell, on the kinetic energy of the simulated system. However, to date there is no a systematic way to assign mass values for shells.

In this study we choose to apply the adiabatic shell scheme due to its computational efficiency, and propose a novel approach for introducing a shell mass in the way that reduces its negative effect on the kinetic energy of the system.

3 Sampling

Our choice of the simulation technique for modeling olivine NaFePO$_4$ has been based on two requirements. We looked for an enhanced sampling method, which

can efficiently sample multi-dimensional space and detect the rare events, as well as be easily extended for simulations on meso-scales. Such properties are critical for effective study of ion transport in bulk and nanostructured materials.

The Generalized Shadow Hybrid Monte Carlo method or GSHMC by Akhmatskaya and Reich was originally developed for efficient atomistic simulation of complex systems [33] and then adjusted to simulation on meso-scales, without losing its capacity for exact sampling at the target temperature [23]. The method however has never been applied to solid state chemistry. In this study we investigate the performance of GSHMC in simulation of olivine NaFePO$_4$ and propose some modifications to the original algorithm aiming to improve its accuracy and sampling efficiency specifically in simulation of battery materials.

3.1 GSHMC: Generalized Shadow Hybrid Monte Carlo

The GSHMC method is a type of Markov Chain Monte Carlo, with better sampling performance than Monte Carlo or MD in molecular simulations and with a negligible computing overhead. GSHMC is especially appropriate when exploring configurational spaces of high dimensionality, finding global energy minima, and simulating rare events such as phase transitions. Its theoretical foundation has already been published elsewhere [23, 24, 33–35]. It has recently been implemented in an open-source MD package [36, 37] and applied to the study of proteins [22, 38]. In the following lines we present a brief summary of the method.

Essentially, GSHMC is a Hybrid Monte Carlo (HMC) method [39] that aims to achieve high efficiency by sampling with respect to modified energies (modified or shadow Hamiltonians). At the same time it preserves most of the dynamical information by applying a partial momentum update instead of fully resampling the momenta between molecular dynamics trajectories, as is the case of HMC.

Shadow Hamiltonians are asymptotic expansions of the true Hamiltonian in powers of the time step Δt. They are conserved better than true Hamiltonians by symplectic integrators such as the leapfrog / Verlet algorithm commonly used in molecular simulations [40]. Thus replacing Hamiltonians with shadow Hamiltonians in Metropolis tests leads to higher acceptance rates than those obtained in the HMC method. The computational cost required for the evaluation of shadow Hamiltonians is negligible compared to the force evaluation in an MD simulation. Efficient algorithms for computing modified energies can be found for example in Refs. [33, 41–43]. The GSHMC method employs the Lagrangian formulation of shadow Hamiltonians of an arbitrary order for the leapfrog integrator [33]. In the case of the 4th order of approximation it leads to the following shadow Hamiltonian:

$$\tilde{\mathcal{H}} = U + \frac{1}{2}\dot{\mathbf{x}}[M\dot{\mathbf{x}}] + \frac{\Delta t^2}{12}\dot{\mathbf{x}}[M\ddot{\mathbf{x}}] - \frac{\Delta t^2}{24}\ddot{\mathbf{x}}[M\ddot{\mathbf{x}}], \qquad (2)$$

where U is the potential energy, $\mathbf{x}$ is the positions vector, and M is the atomic mass matrix. The derivatives of the positions can be obtained using the finite difference approximation. The order of approximation of modified Hamiltonians used in the simulation also affects the acceptance rates. Higher approximation orders provide better acceptance rates, but they also require more time to compute.

The GSHMC method consists of a series of two alternating steps: first, one integrates a short MD trajectory at constant energy, and then performs a partial momentum update. Each of these steps can be accepted or rejected following the result of a Metropolis test where the acceptance probabilities are calculated using shadow Hamiltonians $\tilde{\mathcal{H}}$ instead of true Hamiltonians. The full algorithm can be summarized as follows:

- *Molecular dynamics (MD) step*:
  - Given vectors for positions $\mathbf{x}$ and momenta $\mathbf{p}$, temperature T and mass matrix M, integrate the Hamiltonian equations of the system:

    $$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{x}}, \quad \dot{\mathbf{x}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \qquad (3)$$

    with

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

    using a symplectic method $\Psi_{\Delta t}$ over L steps with time step Δt. This generates a new configuration $\Psi_{\mathcal{T}}(\mathbf{x}, \mathbf{p}) = (\mathbf{x}', \mathbf{p}')$, with $\mathcal{T} = L\Delta t$.
  - Accept or reject the new configuration $(\mathbf{x}', \mathbf{p}')$ by performing a Metropolis test with the probability

    $$\min\left\{1, \frac{\exp\left(-\beta\tilde{\mathcal{H}}(\mathbf{x}', \mathbf{p}')\right)}{\exp\left(-\beta\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p})\right)}\right\},$$

    where $\beta = 1/k_B T$ with k_B being the Boltzmann constant and $\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p})$ the shadow Hamiltonian.
    - If accepted: save $\mathbf{x}'$ and $\mathbf{p}'$ as the current positions and momenta $(\mathbf{x}, \mathbf{p})$.
    - If rejected: restore the initial $\mathbf{x}$ and $\mathbf{p}$ and negate the momenta to ensure the stationarity of the canonical distribution.
- *Partial momentum update (PMU) step*

– Generate a noise vector **u** from the Gaussian distribution $\mathcal{N}(0, \beta^{-1}M)$ as

$$\mathbf{u} = \beta^{-1/2}M^{1/2}\xi,$$

where $\xi = (\xi_1, \ldots, \xi_{3N})^T$, $\xi_i \sim \mathcal{N}(0,1)$, $i = 1, \ldots, 3N$ and N is the system size.

– For the current positions **x** update the momenta **p** using the partial momentum update procedure:

$$\begin{pmatrix} \mathbf{u}' \\ \mathbf{p}' \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix}, \qquad (4)$$

ϕ is a parameter taking values from $(0, \pi/2]$.

– Accept or reject the new momenta $\mathbf{p}'$ by performing a Metropolis test with the probability

$$\min\left\{ 1, \frac{\exp\left(-\beta[\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p}') + \frac{1}{2}(\mathbf{u}')^T M^{-1}\mathbf{u}']\right)}{\exp\left(-\beta[\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p}) + \frac{1}{2}\mathbf{u}^T M^{-1}\mathbf{u}]\right)} \right\}.$$

- If accepted: save $\mathbf{p}'$ as the current momentum **p**.
- If rejected: restore the initial **p**.

Repeat MD and PMU step for a desired number of iterations.

As the simulation is performed in a modified ensemble with respect to shadow Hamiltonians, reweighting has to be applied to calculations of statistical averages [33]. More specifically, given an observable $\Omega(\mathbf{x}, \mathbf{p})$ and its values Ω_i, $i = 1, \ldots, K$, along a sequence of states $(\mathbf{x}_i, \mathbf{p}_i)$, $i = 1, \ldots, K$, the averages $\langle \Omega \rangle$ are calculated as

$$\langle \Omega \rangle_K = \frac{\sum_{i=1}^K w_i \Omega_i}{\sum_{i=1}^K w_i}$$

with weight factors

$$w_i = \exp\left[-\beta\left(\mathcal{H}(\mathbf{x}_i, \mathbf{p}_i) - \tilde{\mathcal{H}}(\mathbf{x}_i, \mathbf{p}_i)\right)\right].$$

3.2 New features introduced in GSHMC for this study

3.2.1 Modified Adaptive Integration Approach (MAIA)

As was pointed above, the original GSHMC method uses the leapfrog integrator and the corresponding modified Hamiltonians of arbitrary accuracy [33]. The leapfrog integrator is a popular choice for molecular dynamics due to the favorable combination of properties such as the second order of accuracy, the reasonably long stability limit interval, the simplicity and computational efficiency. Recently, Radivojević et al. [44]

demonstrated that replacing the leapfrog integrator with the one-parameter 2-stage splitting adaptive integrator specially designed for shadow Hamiltonian Monte Carlo methods may significantly improve accuracy and sampling performance of GSHMC [42, 43]. The authors termed this scheme the Modified Adaptive Integration Approach or MAIA. The adaptive integrator is uniquely determined for a given simulated system and simulation time step, in such a way that the expected error in modified Hamiltonians, $\Delta\tilde{\mathcal{H}}$, is minimal. This immediately implies the best acceptance rates possible within a chosen setup since GSHMC does sample with respect to modified Hamiltonians and, therefore, the error $\Delta\tilde{\mathcal{H}}$ enters the Metropolis test.

In this study, we investigate the efficiency of the MAIA method in simulations of olivine NaFePO$_4$. We briefly summarize MAIA below.

A 2-stage one-parameter splitting integrator $\psi_{\Delta t}$ of a Hamiltonian system (3) with a Hamiltonian

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + U(\mathbf{x}) \equiv A + B \qquad (5)$$

is defined as a composition of solution g-flows of partial systems $X \in \{A, B\}, \Phi_g^X$, where $g = \{b\Delta t, \Delta t/2, (1-2b)\Delta t\}$, Δt is a time step and $0 < b < 1/2$ is a parameter of the family:

$$\psi_{\Delta t} = (\phi_{b\Delta t}^B \circ \phi_{\Delta t/2}^A \circ \phi_{(1/2-b)\Delta t}^B) \circ$$
$$(\phi_{(1/2-b)\Delta t}^B \circ \phi_{\Delta t/2}^A \circ \phi_{b\Delta t}^B) \equiv \Phi_{\Delta t/2}^1 \circ \Phi_{\Delta t/2}^2. \quad (6)$$

The maps $\Phi_{\Delta t/2}^1$ and $\Phi_{\Delta t/2}^2$ advance the solution over a first and a second halves step of length $\Delta t/2$ respectively, therefore the name 2-stage for this integrators family.

Such an integrator is symplectic as a composition of symplectic flows and reversible due to the palindromic structure of (6). A free parameter b fully describes a 2-stage integrator and can be chosen accordingly to some special requirements on the properties of an integrator. Several 2-stage splitting integrators with the parameters b fixed to some specific values are commonly used in molecular dynamics and/or Hybrid Monte Carlo methods [45, 46]. The most celebrated one is the Verlet/leapfrog integrator. Indeed, with $b = 1/4$ both maps in (6), $\Phi_{\Delta t/2}^1$ and $\Phi_{\Delta t/2}^2$, become a velocity Verlet (VV) algorithm with a time step of $\Delta t/2$:

$$\psi_{\Delta t} = (\phi_{\Delta t/4}^B \circ \phi_{\Delta t/2}^A \circ \phi_{\Delta t/4}^B) \circ (\phi_{\Delta t/4}^B \circ \phi_{\Delta t/2}^A \circ \phi_{\Delta t/4}^B)$$
$$= \Phi_{\Delta t/2}^1 \circ \Phi_{\Delta t/2}^2 \equiv \psi_{\Delta t/2}^{VV} \circ \psi_{\Delta t/2}^{VV}.$$

This suggests that in order to make a fair comparison in terms of computational efficiency between an arbitrary 2-stage scheme with the parameter $b \neq 1/4$

and the Verlet integrator in its usual formulation, a 2-stage integrator (6) should be run with a twice longer time step than Verlet, but for a twice shorter number of integration steps L, i.e. $\Delta t_{\text{2-stage}} = 2\Delta t_{\text{Verlet}}$ and $L_{\text{2-stage}} = L_{\text{Verlet}}/2$. Some specific choices of b in (6) lead to the 2-stage integrators, which are capable of outperforming Verlet in accuracy and efficiency with the appropriately selected time steps as it was demonstrated in Refs. [42, 43, 45, 46]. However, with an increasing time step the Verlet integrator shows better performance due to the longer stability limit interval (see for example Ref. [47]).

The MAIA approach [44] provides a rational choice of an integration parameter and identifies a unique value b^* of the parameter b (and thus a unique integrator) for a given simulated system and a chosen time step Δt as

$$b^* = \arg \min_{0 < b < \frac{1}{4}} \max_{0 < h < \bar{h}} \rho(h, b), \qquad (7)$$

where $\rho(h, b)$ is the upper bound for the expected value of the modified energy error $\Delta = \tilde{\mathcal{H}}(\psi_{Lh}(\mathbf{x}, \mathbf{p})) - \tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p})$ with respect to the modified density $\pi(\mathbf{x}, \mathbf{p}) \propto e^{-\beta\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p})}$, i.e. $\mathbb{E}_\pi(\Delta) \leq \rho(h, b)$. Here, as before, $\mathbf{x}$ and $\mathbf{p}$ are position and momentum respectively, ψ_{Lh} is a 2-stage integrator advancing the numerical solution over L steps, h is a dimensionless time step, and $\bar{h} = \sqrt{2}\omega\Delta t$ with ω being the highest frequency of the simulated system. Such a choice of b^* guarantees the best conservation of the modified Hamiltonians and thus the best acceptance of proposals in the GSHMC method. Depending on the values of the highest frequency of a simulated system ω, and a choice of a time step Δt, the adaptive integrator can either coincide with already known integrators with a fixed parameter, e.g. Verlet, the minimum-error integrator, ME [45], or BCSS [46] or be a new integrator, whose efficiency is the best under the chosen conditions.

The derivation of $\rho(h, b)$ is described in Ref. [44], whereas the formulae for modified Hamiltonians $\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p})$ of various orders of approximation corresponding to the multi-stage splitting integrators were obtained in Ref. [42, 43]. Here we only present the expressions we used in this study.

The 4th order modified Hamiltonian for 2-stage splitting integrators derived in terms of quantities available during a simulation reads as:

$$\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + U(\mathbf{x})$$
$$+ \Delta t^2 \left(\alpha \mathbf{p}^T M^{-1}\nabla\dot{U}(\mathbf{x}) + \beta\nabla U(\mathbf{x})^T M^{-1}\nabla U(\mathbf{x}) \right),$$

with

$$\alpha = \frac{6b^* - 1}{24},$$

$$\beta = \frac{6b^{*2} - 6b^* + 1}{12},$$

where $\nabla\dot{U}(\mathbf{x})$ is the numerical time derivative of the gradient of the potential $\nabla U(\mathbf{x})$ and b^* is a parameter of a system specific 2-stage integrator. The upper bound function $\rho(h, b)$ is calculated as [42–44]:

$$\rho(h, b) = \frac{(SB_h + C_h)^2}{2S(1 - A_h^2)},$$

$$S = \frac{1 + 2h^2\beta}{1 + 2h^2\alpha},$$

$$A_h = \frac{h^4 b(1 - 2b)}{4} - \frac{h^2}{2} + 1,$$

$$B_h = -\frac{h^3(1 - 2b)}{4} + h,$$

$$C_h = -\frac{h^5 b^2(1 - 2b)}{4} + h^3 b(1 - b) - h.$$

Importantly, finding the appropriate parameter b^* in (7) can be done at the pre-processing stage of the simulation. Therefore, the procedure does not introduce any computational overhead. Additionally, the method is available for constrained and unconstrained dynamics and it is thus applicable to a broad range of problems.

In Section 4 we compare performance of GSHMC achieved using two different integration schemes, the velocity Verlet and MAIA, for a range of time steps and lengths of MD trajectories. We find that using the MAIA integrators may improve performance of the original GSHMC method by a factor as high as 2.

3.2.2 Randomized Shell Mass Generalized Shadow Hybrid Monte Carlo (RSM-GSHMC)

One important drawback of the adiabatic dynamics core-shell approach is its potential negative effect on the simulated kinetic properties due to the introduction of a shell mass. Previous studies demonstrated that with the careful choice of a shell mass and a time step such an effect becomes negligible [31, 32]. There is not, however, a clear criterion for choosing these parameters, and finding the appropriate parameter values is a matter of trial and error.

In this paper, we propose to take an advantage of the flexibility of the GSHMC method in order to smooth the undesired effect of the shell mass on the kinetics of

a simulated system. The flexibility we refer to in this context is the possibility to vary on the fly in a rigorous manner the simulation parameters in GSHMC [42, 43]. This can be done before starting each new molecular dynamics trajectory, i.e. on each Monte Carlo step, by randomizing the simulation parameters around pre-assigned fixed values. These parameters can be selected independently from a chosen distribution. The randomization helps to avoid some bad combinations of fixed values that might lead to accuracy or performance degradation such as slow convergence and non-ergodicity. Based on this idea we introduced randomization of a shell mass in the GSHMC method.

We implemented the mass randomization as a part of the momentum update step. Before updating the momenta, we redistribute a fraction of the atomic mass between core and shell, keeping the total mass constant:

$$
\begin{aligned}
m_{c,i} &= m_{c,0} - \lambda_i r \\
m_{s,i} &= m_{s,0} + \lambda_i r
\end{aligned}
\tag{8}
$$

where $m_{c,i}$ and $m_{s,i}$ are the core and shell masses at Monte Carlo step i, $m_{c,0}$ and $m_{s,0}$ are their respective initial values, r is the amount of mass that we want to randomize, and λ_i is a random number generated from a uniform distribution $\mathcal{U}(0,1)$ at step i.

It is important to notice that $m_{s,0}$ has to be large enough to ensure the stability of the numerical integrator (its minimum value will depend on the time step used in the simulation). For a discussion about how to choose $m_{s,0}$ see Ref. [32]. On the other hand, r should not be bigger than $m_{c,0}/2$, as that could lead to a situation in which the shell is actually heavier than the core. Having these constraints is enough to rigorously implement the algorithm. However, the optimal choice of r remains empirical. Below we summarize the modified momentum update step in RSM-GSHMC.

– Given the mass matrix M, generate a randomized mass matrix M' by applying the randomization described in (8) to the core and shell particles.
– Generate a noise vector $\mathbf{u}$ from the Gaussian distribution as in the original GSHMC:

$$
\mathbf{u}' = \beta^{-1/2} M'^{1/2} \xi.
$$

– Adjust the current momenta $\mathbf{p}$ to the new masses:

$$
\mathbf{p}' = M'^{1/2} M^{-1/2} \mathbf{p}.
$$

– Update the candidate momenta $\mathbf{p}'$ using the partial momentum update procedure:

$$
\begin{pmatrix} \mathbf{u}'' \\ \mathbf{p}'' \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \mathbf{u}' \\ \mathbf{p}' \end{pmatrix}.
$$

– Accept or reject the new momenta $\mathbf{p}''$ by performing a Metropolis test with the probability

$$
\min \left\{ 1, \frac{\exp\left(-\beta[\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p}'') + \frac{1}{2}(\mathbf{u}'')^T M'^{-1} \mathbf{u}''] \right)}{\exp\left(-\beta[\tilde{\mathcal{H}}(\mathbf{x}, \mathbf{p}') + \frac{1}{2}(\mathbf{u}')^T M'^{-1} \mathbf{u}'] \right)} \right\}.
$$

– If accepted: save $\mathbf{p}''$ as the current momenta $\mathbf{p}$.
– If rejected: restore the initial $\mathbf{p}$.

3.2.3 Implementation - MultiHMC

The GSHMC method was implemented in the open source molecular dynamics software GROMACS [48], version 4.5.4. Details of this implementation can be found in Refs. [36] and [37]. GROMACS was chosen for its popularity, computational efficiency, and effective parallelization. The implementation of the GSHMC method was done in a self-contained manner, respecting the parallel scalability and introducing almost no computational overhead. The same software package was used for the implementation of 2-stage integrators [47] and, in particular, for the MAIA method [44]. We call the resulting package MultiHMC-GROMACS and it is available for public use under the GNU Lesser General Public License. All the classical atomistic simulations in this work were performed with the MultiHMC-GROMACS software.

Additionally, the randomized mass algorithm for the adiabatic core-shell model was implemented in the same software package as a single function call inside the partial momentum update step.

4 Numerical Experiments

In this section we present a series of numerical experiments performed to validate our computational model and to evaluate the performance of the proposed sampling approach. To this end we used four different atomistic simulation methods: MD (CS-min), MD (CS-adi), GSHMC and RSM-GSHMC. In addition, we compared the results with available experimental data and assessed the accuracy of the underlaying force field by performing some groundstate DFT calculations.

4.1 DFT calculations

The total energies were computed using the projected augmented wave (PAW) method [49, 50] within the PBE generalized gradient approximation (GGA) [51] as implemented in the VASP package version 5.3.3 [52]. The GGA+U approach, in which an effective Hubbard U-like term is added to exchange-correlation functional,

Table 2 NaFePO$_4$ surface energies (γ) for different terminations considered after relaxation, as determined by DFT calculations in the present work (γ_{DFT}) and with classical interatomic potentials (γ_{FF}) [29].

Surface	γ_{DFT} (J/m^2)	γ_{FF} (J/m^2)
(010)	0.51	0.52
(201)	0.59	0.63
(101)	0.60	0.74
(100)	0.67	0.68
(110)	0.70	0.54
(111)	0.97	0.68
(001)	1.15	0.90

was required to correctly account for the electronic correlation of iron $3d$ electrons [15]. We used a U value of 4.3 eV as suggested for NaFePO$_4$ in other works [28, 53, 54]. An energy cutoff of 600 eV and a proper k-point mesh were used to ensure that the total energies had converged within 5 meV per formula unit (f.u.). The geometry optimization was considered converged when forces on the atoms for each component became smaller than 0.02 eV/Å.

The surface energies of different possible terminations for NaFePO$_4$ were computed following the approach outlined by Wang et al. [17] for its lithium counterpart. In Table 2 the DFT results obtained for surface energies are shown in comparison with the ones reported using the interatomic potential chosen for this study [29]. The tested methods provide similar values of the surface energies and close trends in the surfaces stability order. The agreement is very good considering the fact that the classical model was obtained by fitting to bulk structural properties only. These results support our choice of the interatomic potential for atomistic simulations carried out in this study.

4.2 MD simulations

We considered two different systems: a fully sodiated NaFePO$_4$ and a partially sodiated Na$_{2/3}$FePO$_4$. The latter was chosen as the most stable compound reported by Saracibar et al. [28] for that composition, which corresponds to a stable ordered superstructure in the Na$_x$FePO$_4$ ($0 \leq x \leq 1$) phase diagram [53, 55, 56]. The unit cell of Na$_{2/3}$FePO$_4$ contains 12 f.u., i.e. 80 atoms.

For bulk NaFePO$_4$ we built a model system based on a ($6 \times 6 \times 6$) supercell containing 864 NaFePO$_4$ f.u. (10368 particles in total including the Fe and O shells). For Na$_{2/3}$FePO$_4$ we used a ($6 \times 3 \times 2$) supercell with 432 Na$_{2/3}$FePO$_4$ f.u. (5040 particles). In both cases the force field parameters were those presented in Section 2, with a cutoff of 12 Å for electrostatics and periodic boundary conditions applied in the three dimensions.

For the partially sodiated case, the extra charge in the system due to removing 1/3 of Na atoms was compensated by averaging the net charge on Fe atoms as was previously suggested in the similar study for LiFePO$_4$ [20].

All of the simulations were initially equilibrated with a 50 ps run using an NPT ensemble with a specified target temperature (T) and pressure $P = 1$ bar. We employed the Berendsen thermostat [57] and the Andersen barostat [58].

Na-ion diffusion events require the presence of vacant Na sites, thus only the partially sodiated system was used for the computation of diffusion coefficients. The production runs in these cases were performed in an NVT ensemble at temperatures between 10 K and 700 K using the Berendsen thermostat.

The velocity Verlet integrator was used for all MD simulations. The optimal choice of the time step in MD is discussed in Section 4.5.

4.3 GSHMC simulations

We tested two versions of the GSHMC method: the original approach and the RSM-GSHMC. We used the same MD setup as described above in the MD runs of the GSHMC methods with two exceptions. The MD trajectories were run in an NVE ensemble, thus no thermostats were involved, and the MAIA integrator was chosen instead of velocity Verlet for most of the tests.

The velocity Verlet integrator was coupled with the original GSHMC method in the parameters refining procedure in order to compare its performance with respect to the MAIA integrators. As in the case of MD simulations, the choice of the time step will be discussed in the following sections. The parameter ϕ in Eq. (4) was fixed to 0.2. The number of integration steps was 500 for MAIA and 1000 for velocity Verlet. The fourth order modified Hamiltonian was used in all tests.

4.4 Validation

First, we verified that the underlying force field used in this study, provides reliable results when employed for dynamical simulations. In addition, we wanted to check that the proposed simulation techniques with the chosen simulations settings are capable to accurately reproduce the properties of olivine NaFePO$_4$.

To this end, we calculated the lattice constants of the fully sodiated NaFePO$_4$ at $T = 300$ K and $P = 1$ bar, based on production runs of 0.5 ns using GSHMC, RSM-GSHMC, MD (CS-min) and MD (CS-adi). The results are shown in comparison with the experimental
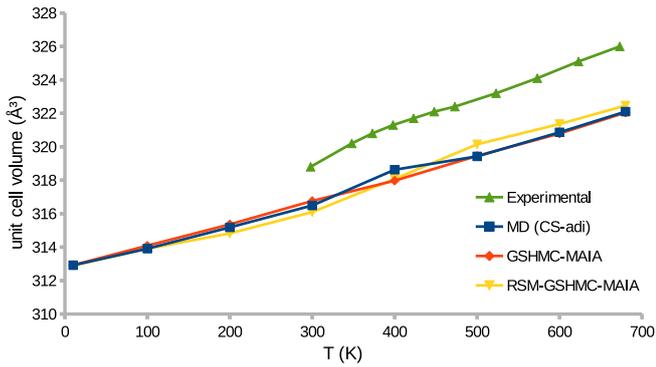
Fig. 1 Thermal expansion for olivine NaFePO$_4$ calculated using MD, GSHMC, RSM-GSHMC in an NPT ensemble with the Andersen barostat. Experimental values are taken from Ref. [14].

Fig. 2 Computational performance for all the considered methods at different temperatures.

data [14] and the DFT-based calculations in Table 3. We found that all the methods yield very similar lattice constants, with relative differences less than 2%.

We also considered the thermal expansion of bulk NaFePO$_4$ to evaluate the suitability of the force field. We computed the volume expansion of the unit cell as a function of temperature by performing simulations under an NPT ensemble. The pressure was maintained using the Andersen barostat. The target temperature was controlled by using the Berendsen thermostat in MD, while GSHMC keeps T constant by design. We considered temperatures between 10 K and 700 K, making sure that the box size and the potential energy were completely stabilized before measuring the volume. The resulting thermal expansion for NaFePO$_4$ is shown in Figure 1 along with the experimental results of Moreau et al. [14]. The three tested methods yield similar slopes and the small difference observed with respect to the experimental values is negligible (relative variations are less than 1%). Therefore, we can conclude that the model combined with the simulation methods under study properly accounts for the thermal expansion of olivine NaFePO$_4$.

As a final validation test we present in Table 4 the average values for potential (U) and kinetic energies (K), temperatures, and two structural parameters, the angles between the bonded O-P-O species (θ_{O-P-O}) and their corresponding P-O distances (d_{P-O}). As can be seen from Table 4, the randomized mass algorithm, RSM-GSHMC, provides the best agreement with the experimental data, which may imply the positive effect of the randomization of a shell mass on the overall accuracy of the core-shell adiabatic model.

In Figure 2 we show the computational performance measured in nanoseconds per day for the four methods considered. All simulations were run in parallel on 8 cores on the same computational server. In terms of
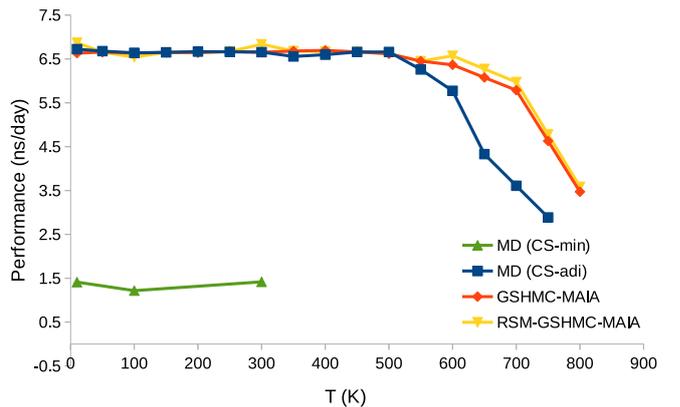
performance, the adiabatic core-shell approach offers a great advantage that outweighs any marginal loss of accuracy. The significantly lower performance observed with the MD (CS-min) scheme is due to the big overhead introduced by the search of optimal shell positions at each time step. The loss of performance registered at temperatures over 500 K is a consequence of using a smaller time step, which was found necessary to keep the simulations stable at such high temperatures. The GSHMC approaches always achieved a higher performance because their increased numerical stability allowed the use of longer time steps. For temperatures below 500 K the time steps were set to 1.15 fs for MD (CS-adi) and 2.3 fs for GSHMC methods whereas for higher temperatures they had to be reduced to 0.5 fs and 2 fs, respectively.

4.5 Accuracy and sampling performance

In order to include in our tests the calculation of Na self-diffusion coefficients we chose the partially sodiated Na$_{2/3}$FePO$_4$ as a benchmark system. The diffusion coefficients are notoriously difficult to determine from dynamical simulations because they require considerably long runs to reach convergence. In this work they were derived from the mean square displacement of Na-ions using the Einstein relation

$$\langle |\mathbf{x}_{Na}(t+\tau) - \mathbf{x}_{Na}(t)|^2 \rangle = 6D\tau, \tag{9}$$

where the term on the left side, which is the squared displacement of a Na-ion during an integrated interval τ, is proportional to the Na self-diffusion (or diffusion) coefficient (D) and τ.

In what follows, all reported properties are results of averaging over five different production runs of 2 ns each, unless stated otherwise.

Table 3 Computed lattice constants of olivine NaFePO$_4$ at 300 K using different approaches. Experimental values also shown are taken from Ref. [14].

Parameter (Å)	Exp.	DFT	MD (CS-min)	MD (CS-adi)	GSHMC	RSM-GSHMC
a	10.41	10.52	10.34	10.38	10.39	10.40
b	6.22	6.27	6.16	6.19	6.19	6.19
c	4.95	4.99	4.90	4.92	4.92	4.91

Table 4 Average values for temperature (T), potential energy (U), kinetic energy (K), O-P-O angles (θ_{O-P-O}) and P-O internuclear distances (d_{P-O}). Experimental values also shown are taken from Ref. [14].

Method	T (K)	U (kJ/mol)	K (kJ/mol)	$\theta_{O\text{-}P\text{-}O}$ (deg.)	$d_{P\text{-}O}$ (nm)
MD (CS-min)	297.02	-1.01x10^7	1.09x10^4	108.29	0.150
MD (CS-adi)	299.98	-1.03x10^7	1.63x10^4	108.26	0.151
GSHMC	296.40	-1.03x10^7	1.60x10^4	108.30	0.149
RSM-GSHMC	298.64	-1.03x10^7	1.61x10^4	108.32	0.156
Experiment	300.00	–	–	109.47	0.155
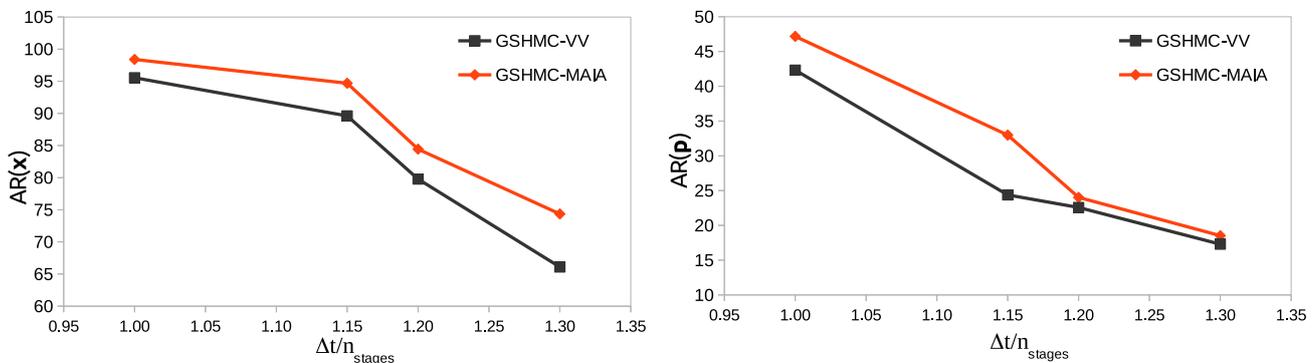
Fig. 3 Acceptance rates for positions (left) and momenta (right) for GSHMC simulations using MAIA and velocity Verlet integrators at $T = 300$ K.

The first step for optimizing the accuracy and performance of the novel approaches in prediction of various properties of the system of interest, is to find the best combination of numerical integrators and simulation parameters to be used. We begin with measuring the sampling efficiency of GSHMC in two different scenarios, namely when the method is combined with the new MAIA integrator or when it uses the standard velocity Verlet.

For these experiments we chose a time step in the MAIA integrator twice longer than the one in the velocity Verlet case presented in its usual, 1-stage formulation. However, since MAIA performs two force evaluations per each integration step in contrast to only one in velocity Verlet, a number of integration steps in MAIA has been chosen to be half as many as in velocity Verlet. Such a setup equalizes the computational effort required for each integrator and makes the comparison fair (see Section 3.2.1 for the detailed explanation). For simplicity, we use from now on an effective time step, defined as $\Delta t/n_{\text{stages}}$, where n_{stages} is equal to either 1 for velocity Verlet or to 2 for MAIA.

One can evaluate the influence of the tested integrators on the sampling performance by looking at the acceptance rates (AR) for the positions and momenta Metropolis tests. Ideally, the AR for positions in GSHMC should be close to 100%, to minimize the time spent on computing trajectories that are finally rejected. In Figure 3 we show the AR for positions and momenta when using the two integrators for a range of integration time steps (Δt). As follows from Figure 3, MAIA always leads to better acceptance rates, as for positions as for momenta, than can be achieved with velocity Verlet.

Another way to compare the sampling efficiency is to calculate integrated autocorrelation functions (IACF), defined as:

$$\text{IACF} = \sum_{l=0}^{K'} ACF(\tau_l), \qquad (10)$$

where $ACF(\tau_l)$, $l = 0, ..., K' < K$ is the standard autocorrelation function for the time series Ω_k of K samples,
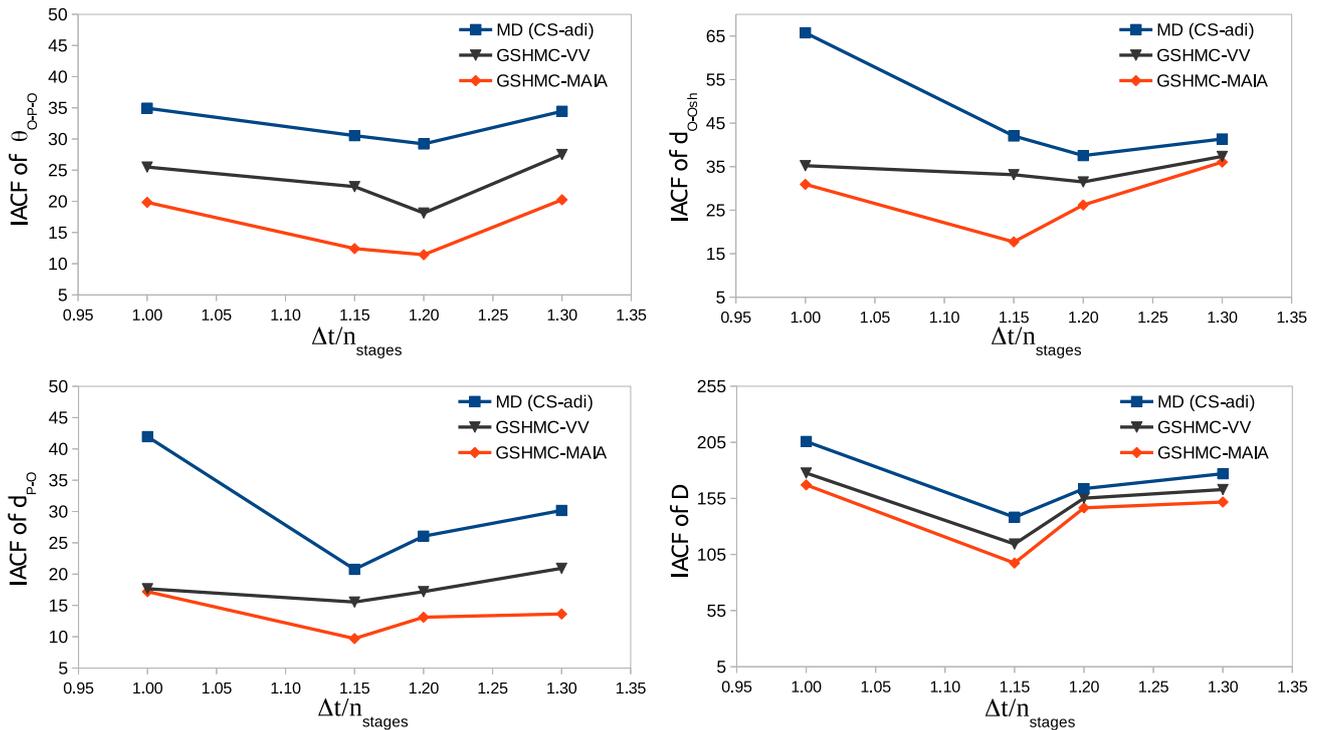
Fig. 4 Integrated autocorrelation functions for the diffusion coefficient and structural parameters at $T = 300$ K.

$k = 1, ..., K$, with the normalization

$$ACF(\tau_0) = ACF(0) = 1.$$

The IACF gives a quantitative measure of time required, on average, for generating a non-correlated sample, and thus lower values of IACF imply more efficient sampling.

In Figure 4 we present the IACF values for several properties of the system obtained with the MD and GSHMC methods for different effective time steps. The latter was combined with velocity Verlet (GSHMC-VV) and MAIA (GSHMC-MAIA). Clearly, the combination of GSHMC with MAIA always produces the lowest IACF values, which translates into a more efficient sampling. On the other hand, plotting IACF as a function of the effective time step, helps to reveal the influence of the time step on the overall performance and suggest a way to choose the optimal one. More specifically, we found that the best performance was observed for all the simulation methods at the effective time step of 1.15 fs and thus the rest of the tests were performed with this value. Also, since the GSHMC-MAIA combination provided the best sampling efficiency we proposed this setup for future studies.

Once we chose the proper settings for the MD and GSHMC methods, longer 4 ns simulations at constant volume and temperature (NVT) were performed. In Figure 5 we plot the relative IACF for the structural
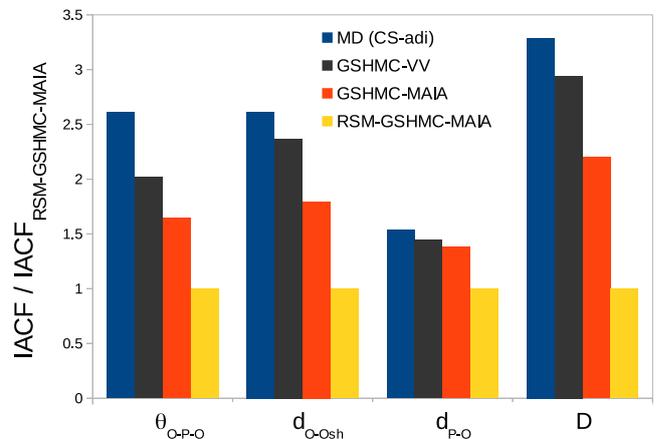
Fig. 5 Relative IACF with respect to RSM-GSHMC-MAIA for structural properties and diffusion coefficients obtained with the optimal simulation parameters at $T = 300$ K.

parameters and self-diffusion observed with MD (CS-adi), GSHMC-VV and GSHMC-MAIA at 300K with respect to the corresponding IACF values obtained with RSM-GSHMC-MAIA. Clearly, the best performance is obtained with the RSM-GSHMC-MAIA simulations for all simulated properties (lowest IACF values, up to 3.3 times better than in MD). This is a very promising result, especially for computing self-diffusion coefficients in solid bulk materials.
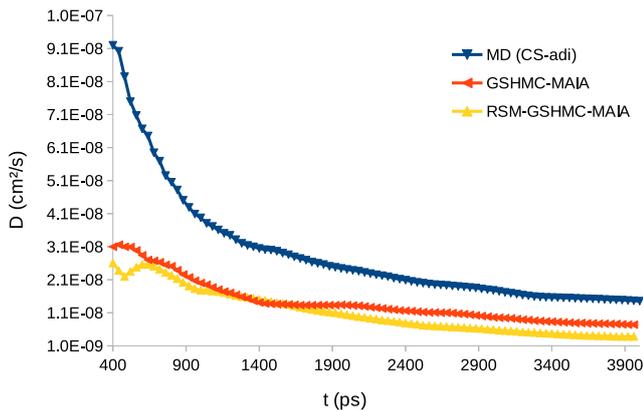
Fig. 6 Diffusion coefficient convergence at $T = 300$ K for MD, GSHMC and RSM-GSHMC methods.

The enhanced sampling of the Na-ion self-diffusion observed with RSM-GSHMC in Figure 5 implies shorter integration times required for obtaining the converged self-diffusion value. Figure 6 monitors the average self-diffusion obtained with MD, GSHMC and RSM-GSHMC at $T = 300$ K with increasing simulation time up to 4 ns. Though convergence is not fully achieved with any of the methods, the GSHMC-based, and especially RSM-GSHMC, demonstrate clear signs of convergence after 3 ns of simulation.

Next, we investigated the performance of MD, GSHMC and RSM-GSHMC in the range of temperatures by running a series of NVT simulations at temperatures between 10 K and 700 K. As before (see section 4.4), the integration time steps had to be reduced for temperatures greater than 500 K for all methods.

We introduced a variable X that measures the sampling performance by taking into account both the effective time step $\Delta t/n_{\text{stages}}$ and the IACF as:

$$X = \frac{\Delta t/n_{\text{stages}}}{\text{IACF}}. \tag{11}$$

In Figures 7 and 8 we present the performance of the methods for a range of temperatures and different quantities of interest in terms of relative X values with respect to the obtained with MD. We can see that the GSHMC methods with and without mass randomization offer a significant improvement over MD (up to 2.5 and 4.7 for GSHMC and RSM-GSHMC respectively).

As we noticed in Figure 5, the RSM-GSHMC method is particularly beneficial for calculating diffusion coefficients. This is apparent at all temperatures (see Figure 8). For other calculated properties RSM-GSHMC also demonstrates its superiority over MD and GSHMC at all temperatures though its performance differs less dramatically from the one offered by GSHMC. Yet another advantage of RSM-GSHMC over other tested methods
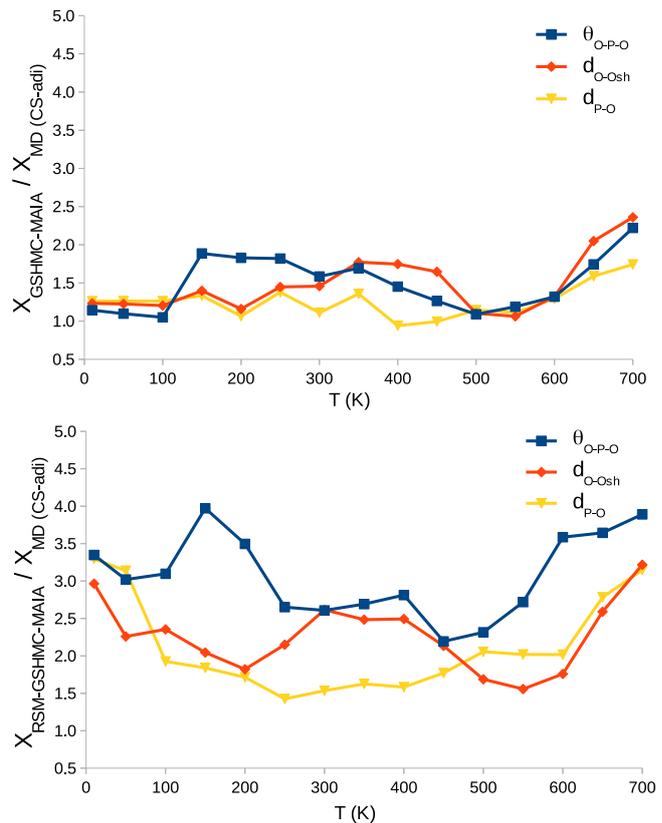
Fig. 7 Sampling performance (X) relative to MD at different temperatures achieved for several structural parameters when using the GSHMC (top) and RSM-GSHMC (bottom) methods.
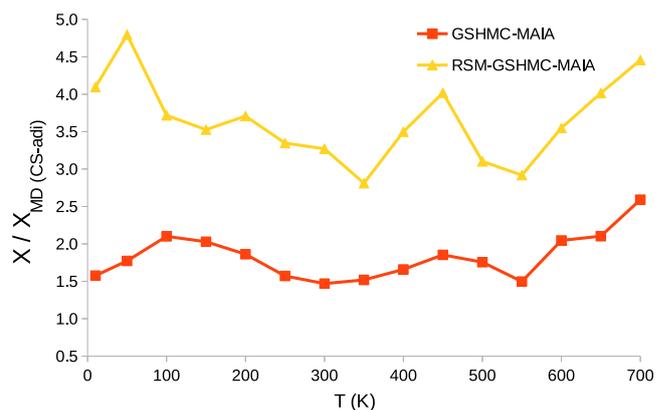
Fig. 8 Sampling performance (X) relative to MD at different temperatures achieved when computing the diffusion coefficients with GSHMC and RSM-GSHMC methods.

is that it can be further tuned by modifying the amount of randomized mass in Eq. (8) for each specific temperature.

5 Conclusions

We presented the new methodology for atomistic simulation of solid state materials for batteries, which offers a better accuracy and sampling efficiency than can be achieved with popular molecular dynamics (MD) approaches. The sampling in this method is performed with Generalized Shadow Hybrid Monte Carlo or GSHMC, which combines in a rigorous and effective manner molecular dynamics trajectories with Monte Carlo steps. The accuracy of the method is ensured by the new system specific adaptive integrators MAIA used in the MD step as well as by the modifications introduced in the adiabatic Core-Shell model for retaining the dynamics of a simulated system. Utilizing the adiabatic Core-Shell model or CS-adi in the new method instead of the Core-Shell relaxation scheme or CS-min, yields important performance gain saving up to 80% of the computational time. We have applied the method to the study of olivine NaFePO$_4$ systems and analyzed its accuracy and performance in comparison with available experimental data, the DFT computed properties, and the results obtained with other atomistic simulation methods (MD and conventional GSHMC). The accuracy of the method in the calculation of lattice constants and thermal expansion has been compared against DFT-based calculations and experimental data, obtaining reliable results for all properties. Moreover, the method demonstrates a better agreement with the experimental data than one can observe with other tested atomistic methods, namely MD (CS-min), MD (CS-adi) and the original GSHMC.

Introducing the novel MAIA integrator in our new methodology has also allowed for more efficient sampling when characterizing structural properties, such as average angles between atoms and bond lengths, as well as improving stability at higher temperatures.

Applying a randomization term to the shell mass improved not only the accuracy but also the sampling efficiency, especially when measuring diffusion coefficients. This modification of the GSHMC algorithm does not introduce significant overhead and is fully compatible with parallel implementations.

In summary, the proposed methodology can be viewed as an alternative to molecular dynamics for atomistic studies of solid-state battery materials whenever high accuracy and efficient sampling are critical for obtaining tractable simulation results.

References

1. Tarascon JM, Armand M (2001) Issues and challenges facing rechargeable lithium batteries. Nature 414:359–367, doi:10.1038/35104644
2. Etacheri V, Marom R, Elazari R, Salitra G, Aurbach D (2011) Challenges in the development of advanced Li-ion batteries: a review. Energy Environ Sci 4:3243–3262, doi:10.1039/C1EE01598B
3. Armand M, Tarascon JM (2008) Building better batteries. Nature 451:652–657, doi:10.1038/451652a
4. Goodenough JB, Park KS (2013) The Li-Ion Rechargeable Battery: A Perspective. J Am Chem Soc 135:1167–1176, doi:10.1021/ja3091438
5. Yabuuchi N, Kubota K, Dahbi M, Komaba S (2014) Research Development on Sodium-Ion Batteries. Chem Rev 114:11636–11682, doi:10.1021/cr500192f
6. Han MH, Gonzalo E, Singh G, Rojo T (2015) A comprehensive review of sodium layered oxides: powerful cathodes for Na-ion batteries. Energy Environ Sci 8:81–102, doi:10.1039/C4EE03192J
7. Choi JW, Aurbach D (2016) Promise and reality of post-lithium-ion batteries with high energy densities. Nat Rev Mater 1:16013, doi:10.1038/natrevmats.2016.13
8. Hautier G, Jain A, Ong SP (2012) From the computer to the laboratory: materials discovery and design using first-principles calculations. J Mater Sci 47:7317–7340, doi:10.1007/s10853-012-6424-0
9. Jain A, Shin Y, Persson KA (2016) Computational predictions of energy materials using density functional theory. Nat Rev Mater 1:15004, doi:10.1038/natrevmats.2015.4
10. Islam MS, Fisher CAJ (2014) Lithium and sodium battery cathode materials: computational insights into voltage, diffusion and nanostructural properties. Chem Soc Rev 43:185–204, doi:10.1039/C3CS60199D
11. Meng YS, Arroyo-de Dompablo ME (2009) First principles computational materials design for energy storage materials in lithium ion batteries. Energy Environ Sci 2:589–609, doi:10.1039/B901825E

12. Meng YS, Arroyo-de Dompablo ME (2013) Recent Advances in First Principles Computational Research of Cathode Materials for Lithium-Ion Batteries. Acc Chem Res 46:1171–1180, doi:10.1021/ar2002396

13. Ramzan M, Lebègue S, Larsson P, Ahuja R (2009) Structural, magnetic, and energetic properties of Na2FePO4F, Li2FePO4F, NaFePO4F, and LiFePO4F from ab initio calculations. J Appl Phys 106:043510, doi:10.1063/1.3202384

14. Moreau P, Guyomard D, Gaubicher J, Boucher F (2010) Structure and Stability of Sodium Intercalated Phases in Olivine FePO4. Chem Mater 22:4126–4128, doi:10.1021/cm101377h

15. Zhou F, Cococcioni M, Marianetti CA, Morgan D, Ceder G (2004) First-principles prediction of redox potentials in transition-metal compounds with LDA+U. Phys Rev B 70:235121, doi:10.1103/PhysRevB.70.235121

16. Dathar GKP, Sheppard D, Stevenson KJ, Henkelman G (2011) Calculations of Li-Ion Diffusion in Olivine Phosphates. Chem Mater 23:4032–4037, doi:10.1021/cm201604g

17. Wang L, Zhou F, Meng YS, Ceder G (2007) First-principles study of surface properties of LiFePO$_4$: Surface energy, structure, Wulff shape, and surface redox potential. Phys Rev B 76:165435, doi:10.1103/PhysRevB.76.165435

18. Burbano M, Carlier D, Boucher F, Morgan BJ, Salanne M (2016) Sparse Cyclic Excitations Explain the Low Ionic Conductivity of Stoichiometric Li$_7$La$_3$Zr$_2$O$_{12}$. Phys Rev Lett 116:135901, doi:10.1103/PhysRevLett.116.135901

19. Boulfelfel SE, Seifert G, Leoni S (2011) Atomistic investigation of Li+ diffusion pathways in the olivine LiFePO4 cathode material. J Mater Chem 21:16365–16372, doi:10.1039/C1JM10725A

20. Tealdi C, Spreafico C, Mustarelli P (2012) Lithium diffusion in Li1-xFePO4: the effect of cationic disorder. J Mater Chem 22:24870–24876, doi:10.1039/C2JM35585J

21. Yang J, Tse JS (2011) Li Ion Diffusion Mechanisms in LiFePO4: An ab Initio Molecular Dynamics Study. J Phys Chem A 115:13045–13049, doi:10.1021/jp205057d

22. Wee C, Samson M, Reich S, Akhmatskaya E (2008) 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, doi:10.1021/jp076712u

23. Akhmatskaya E, Reich S (2011) Meso-GSHMC: A stochastic algorithm for meso-scale constant temperature simulations. Procedia Computer Science 4:1353 – 1362, doi:10.1016/j.procs.2011.04.146

24. Escribano B, Akhmatskaya E, Reich S, Azpiroz J (2015) Multiple-time-stepping generalized hybrid Monte Carlo methods. J Comput Phys 280:1 – 20, doi:10.1016/j.jcp.2014.08.052

25. Galceran M, Saurel D, Acebedo B, Roddatis VV, Martin E, Rojo T, Casas-Cabanas M (2014) The mechanism of NaFePO4 (de)sodiation determined by in situ X-ray diffraction. Phys Chem Chem Phys 16:8837–8842, doi:10.1039/C4CP01089B

26. Yuan LX, Wang ZH, Zhang WX, Hu XL, Chen JT, Huang YH, Goodenough JB (2011) Development and challenges of LiFePO4 cathode material for lithium-ion batteries. Energy Environ Sci 4:269–284, doi:10.1039/C0EE00029A

27. Zhu Y, Xu Y, Liu Y, Luo C, Wang C (2013) Comparison of electrochemical performances of olivine NaFePO4 in sodium-ion batteries and olivine LiFePO4 in lithium-ion batteries. Nanoscale 5:780–787, doi:10.1039/C2NR32758A

28. Saracibar A, Carrasco J, Saurel D, Galceran M, Acebedo B, Anne H, Lepoitevin M, Rojo T, Casas Cabanas M (2016) Investigation of sodium insertion-extraction in olivine Na$_x$FePO4 $(0 \leq x \leq 1)$ using first-principles calculations. Phys Chem Chem Phys 18:13045–13051, doi:10.1039/C6CP00762G

29. Whiteside A, Fisher C, Parker S, Islam MS (2014) Particle shapes and surface structures of olivine NaFePO$_4$ in comparison to LiFePO$_4$. Phys Chem Chem Phys 16:21788–21794, doi:10.1039/C4CP02356K

30. Dick BG, Overhauser AW (1958) Theory of the Dielectric Constants of Alkali Halide Crystals. Phys Rev 112:90 – 103, doi:10.1103/PhysRev.112.90

31. Lindan P, Gillan M (1993) Shell-model molecular dynamics simulation of superionic conduction in CaF$_2$. J Phys: Condens Matter 5:1019, doi:10.1088/0953-8984/5/8/005

32. Mitchell P, Fincham D (1993) Shell model simulations by adiabatic dynamics. J Phys: Condens Matter 5:1031, doi:10.1088/0953-8984/5/8/006

33. Akhmatskaya E, Reich S (2008) GSHMC: An efficient method for molecular simulation. J Comput Phys 227:4934 – 4954, doi:10.1016/j.jcp.2008.01.023

34. Akhmatskaya E, Reich S (2012) New Hybrid Monte Carlo Methods for Efficient Sampling: from Physics to Biology and Statistics. Prog Nucl Sci Technol 2:447 – 462, doi:10.15669/pnst.2.447

35. Akhmatskaya E, Nobes R, Reich S (2011) Method, Apparatus and Computer Program for Molecular

Simulation. patent US007908129 4:1353 – 1362

36. Escribano B, Akhmatskaya E, Mujika J (2013) Combining stochastic and deterministic approaches within high efficiency molecular simulations. Cent Eur J Math 11:787–799, doi:10.2478/s11533-012-0164-x

37. Fernández-Pendás M, Escribano B, Radivojević T, Akhmatskaya E (2014) Constant pressure hybrid Monte Carlo simulations in GROMACS. J Mol Model 20, doi:10.1007/s00894-014-2487-y

38. Mujika JI, Escribano B, Akhmatskaya E, Ugalde JM, Lopez X (2012) Molecular Dynamics Simulations of Iron- and Aluminum-Loaded Serum Transferrin: Protonation of Tyr188 Is Necessary To Prompt Metal Release. Biochemistry 51:7017–7027, doi:10.1021/bi300584p

39. Duane S, Kennedy A, Pendleton B, Roweth D (1987) Hybrid Monte Carlo. Phys Lett B 195:216–222, doi:10.1016/0370-2693(87)91197-X

40. Hairer E, Lubich C, Wanner G (2002) Geometric Numerical Integration. Springer-Verlag, Berlin Heidelberg

41. Skeel R, Hardy D (2001) Practical construction of modified Hamiltonians. SIAM J Sci Comput 23:1172–1188, doi:10.1137/S106482750138318X

42. Radivojević T, Akhmatskaya E (2017) Mix & Match Hamiltonian Monte Carlo. Preprint

43. Radivojević T (2016) Enhancing Sampling in Computational Statistics Using Modified Hamiltonians. Ph.D. thesis, University of the Basque Country, doi:20.500.11824/323

44. Radivojević T, Fernández-Pendás M, Sanz Serna JM, Akhmatskaya E (2017) Adaptive splitting integrators for enhancing sampling efficiency of shadow Hamiltonian Monte Carlo methods in molecular simulation and statistics. Preprint

45. McLachlan RI (1995) On the numerical integration of ordinary differential equations by symmetric composition methods. SIAM J Sci Comput 16:151 – 168, doi:10.1137/0916010

46. Blanes S, Casas F, Sanz-Serna JM (2014) Numerical integrators for the Hybrid Monte Carlo method. SIAM J Sci Comput 36 (4):A1556 – A1580, doi:10.1137/130932740

47. Fernández-Pendás M, Akhmatskaya E, Sanz Serna J (2016) Adaptive multi-stage integrators for optimal energy conservation in molecular simulations. J Comput Phys 327:434 – 449, doi:10.1016/j.jcp.2016.09.035

48. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447, doi:10.1021/ct700301q

49. Blöchl PE (1994) Projector augmented-wave method. Phys Rev B 50:17953–17979, doi:10.1103/PhysRevB.50.17953

50. Kresse G, Joubert D (1999) From ultrasoft pseudopotentials to the projector augmented-wave method. Phys Rev B 59:1758–1775, doi:10.1103/PhysRevB.59.1758

51. Perdew JP, Burke K, Ernzerhof M (1996) Generalized Gradient Approximation Made Simple. Phys Rev Lett 77:3865–3868, doi:10.1103/PhysRevLett.77.3865

52. Kresse G, Furthmüller J (1996) Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Phys Rev B 54:11169–11186, doi:10.1103/PhysRevB.54.11169

53. Boucher F, Gaubicher J, Cuisinier M, Guyomard D, Moreau P (2014) Elucidation of the Na$_{2/3}$FePO$_4$ and Li$_{2/3}$FePO$_4$ Intermediate Superstructure Revealing a Pseudouniform Ordering in 2D. J Am Chem Soc 136:9144–9157, doi:10.1021/ja503622y

54. Lu J, Chung SC, Nishimura Si, Yamada A (2013) Phase Diagram of Olivine Na$_x$FePO$_4$ (0 < x < 1). Chem Mater 25:4557–4565, doi:10.1021/cm402617b

55. Galceran M, Roddatis V, Zúñiga FJ, Pérez-Mato JM, Acebedo B, Arenal R, Peral I, Rojo T, Casas-Cabanas M (2014) Na-Vacancy and Charge Ordering in Na$\approx_{2/3}$FePO$_4$. Chem Mater 26:3289–3294, doi:10.1021/cm501110v

56. Casas-Cabanas M, Roddatis VV, Saurel D, Kubiak P, Carretero-Gonzalez J, Palomares V, Serras P, Rojo T (2012) Crystal chemistry of Na insertion/deinsertion in FePO$_4$-NaFePO$_4$. J Mater Chem 22:17421–17423, doi:10.1039/C2JM33639A

57. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690, doi:10.1063/1.448118

58. Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. J Chem Phys 72:2384–2393, doi:10.1063/1.439486