

A Foundation Model for Atomistic Materials Chemistry: MACE-MP-0

BCN: 4HUY8

Word Count: 4899

Cavendish Laboratory, Department of Physics, J J Thomson Avenue, Cambridge. CB3 0HE

Abstract

Machine learning potentials have revolutionised the field of atomistic simulations by providing the building blocks towards ab initio level accuracy at classical molecular dynamics (MD) efficiency. However, their lack of transferability has driven the development of “all-purpose” foundation models trained on large datasets, capable of accurately modelling a wide range of systems and materials. MACE-MP-0 is a promising recent foundation model, trained on data from the Materials Project, achieving great generalisability, efficiency and accuracy owing to the innovative MACE architecture. This study applies MACE-MP-0, along with its recent variants, to assess the suitability of a novel Ti_2CO_2 MXene/graphene anode material for sodium-ion batteries. Out of the models, the latest variant, MACE-OMAT-0, shows the closest agreement with the PBE reference data. This model produces stable NVT MD of sodium ion diffusion within the anode material for 100 ps at 300 K, where the results showed qualitative agreement with the PBE reference. Overall, this investigation provides a challenging test to benchmark the performance of future foundation models.

1. Introduction

Atomistic simulations are a vital tool for investigating the fundamental behaviour and properties of materials, biomolecules, electrochemistry and more. Historically, achieving both highly accurate and efficient atomistic simulations has been a grand challenge, as the improvement of one often results in a compromise of the other [1]. Ab initio techniques, such as Density Functional Theory (DFT), enable the accurate modelling of a wide variety of systems with the drawback of being computationally expensive [2]. This provides challenges for simulating large, complex systems [3] as well as long-timescale processes [4].

Classical Molecular Dynamics (MD) represents the other side of the spectrum, where accuracy is sacrificed for highly efficient, empirically fitted force fields and atoms move according to Newton’s Laws of Motion. These force fields often involve vast simplifications and assumptions regarding the functional form of the underlying potential energy surface (PES), which offer a dramatic speed up [5]. Furthermore, a major downside of classical MD is the inability to model bond breaking without specialised reactive force fields such as ReaxFF [6],

whereas bond breaking is inherently captured in electronic structure methods.

In recent years, a new approach for modelling potential energy surfaces has emerged: Machine Learning Potentials (MLPs). MLPs represent a significant leap towards the ultimate goal of ab initio level accuracy at classical force field efficiency [3]. MLPs are trained on a representative dataset of electronic structure calculations, where the model learns to minimise, for example, the root mean square error (RMSE) between its own energy and force predictions and the reference energies and forces [2]. Dataset generation is very costly due to the large number of required electronic structure calculations, however, this is a one-time cost as predictions are much cheaper.

However, MLPs are often tailored for specific systems and lack transferability [7], meaning each new system being studied requires the generation of costly DFT datasets. For this reason, recently, the focus has shifted to foundation models. These are models trained on vast, diverse datasets with the goal of being able to describe the potential energy surface of any configuration of any element or materials with ab initio level accuracy at a fraction of the cost [8]. These foundation models can be

applied “out-of-the-box”, or fine-tuned to enhance performance on more specific tasks. In recent years, there have been numerous efforts towards foundation models including CHGNet [9], M3GNet [10], GNoME [11] and more, trained on a variety of different datasets, many of which have been based on the Materials Project [8], and with many different architectures, including Graph Neural Networks, and recently equivariant transformers [12] (EquiformerV2 architecture [13]). However, many of these models struggle with their comprehensiveness when describing different phases and interactions of materials and so are not truly “foundation models” but are steps towards them. The recent MACE-MP-0 foundation model represents the most promising stride towards this goal, showing strong qualitative and sometimes quantitative agreement with PBE reference data for a wide range of materials, reactions and renewable energy applications [8].

A major challenge of renewable energy sources, such as solar or wind, is their intermittency [14]. Developing effective and affordable batteries for energy storage is vital to address this issue [15]. However, using Lithium-ion batteries (LIBs) is infeasible for large-scale energy storage due to Lithium’s diminishing supply and increasing cost [16]. Sodium-ion batteries (SIBs) are more promising, with high elemental abundance in the earth’s crust and high energy density [17]. The development of suitable anode materials is a major bottleneck, as common graphite and silicone anodes have too small interlayer spacings for easy reversible intercalation of the larger Na ion ($r = 1.02 \text{ \AA}$) compared to Li ($r = 0.71 \text{ \AA}$) [18].

Recently, 2D transition metal carbide MXenes have emerged as promising anode materials due to their high conductivity and ease of surface functionalisation [19]. MXenes are a family of 2D materials with general formula $M_{n+1}X_nT_m$, where M is a transition metal, X is carbon or nitrogen, and T are surface terminations [20]. Furthermore, creating a heterostructure by alternating MXene and graphene layers was shown to halve the energy barriers for Na-ion diffusion compared to stacking MXene layers, improving the charge/discharge efficiency [21].

This investigation will focus on the out-of-the-box application of the family of MACE foundation models for materials (MACE-MP-0, MACE-MPA-0 and MACE-OMAT-0) to model a next-generation graphene/Ti₂CO₂ MXene anode mate-

rial for sodium-ion batteries.

2. Methodology

The foundation model MACE-MP-0 is built using the MACE architecture, which is an equivariant Message Passing Neural Network (MPNN). MACE-MP-0 has gained popularity as a foundation model for not only achieving state-of-the-art benchmark performance, but also outperforming previous models in terms of training speed and generalisation to data outside of the training set [22].

Fundamentally, MLPs take an atomic configuration as an input and output its potential energy. The energy of an atom i can be written as the sum of different body-order contributions:

$$E_i = V_1(r_i) + \frac{1}{2} \sum_j V_2(r_{ij}) + \frac{1}{3!} \sum_{j,k} V_3(r_{ijk}) + \dots \quad (1)$$

where V_n represents an n body interaction potential and the sum is calculated over all distinct sets of interacting atoms [2]. However, even with an interaction cut-off, evaluating this sum to a sufficiently high order for the required accuracy is too computationally expensive [23].

The Atomic Cluster Expansion (ACE) is one method to tackle this issue, creating a complete basis which achieves linear scaling [23]. An important feature of the basis is physical symmetries, where scalars need to obey translational, rotational, reflectional, and permutational invariances, as well as for vectors and tensors to fulfil equivariance [24]. Firstly, the total potential energy E is broken down into site energies E_i [25], which ensures permutational invariance as the summation has arbitrary order:

$$E = \sum_i E_i. \quad (2)$$

The basis is made up of single-particle basis functions A_{iv} , where their products are used to create higher body-ordered terms, hence its linear scaling regardless of the expansion’s body-order [2]

$$E_i = \sum_v c_v^{(1)} A_{iv} + \sum_{v_1, v_2}^{v_1 \geq v_2} c_{v_1, v_2}^{(2)} A_{v_1} A_{v_2} + \sum_{v_1, v_2, v_3}^{v_1 \geq v_2 \geq v_3} c_{v_1, v_2, v_3}^{(3)} A_{v_1} A_{v_2} A_{v_3} + \dots \quad (3)$$

ACE achieves high computational and data efficiency when used in linear models, enabling fast and flexible MLPs [26]. MACE extends this, incorporating non-linearity through neural networks. MACE is a *non-linear* neural-network multilayer-ACE (MACE), where the ACE description is used in an equivariant MPNN, a type of graph neural network (GNN) [27]. Here, the atomic environment representations are no longer fixed, and are instead learned [2]. Graphs are naturally suited to describing molecules as they are made up of nodes and edges, where nodes are atoms and edges connect the nodes [2]. Considering a central atom i , only interactions with other atoms within its local environment are considered, defined by the cut-off radius hyperparameter [28]. Although this locality approximation is used in many MLPs, no assumptions regarding the functional form of the PES are made, where it is instead purely learned from the reference data [5]. Within this framework, the state of each node i in layer t of the MPNN is given by:

$$\sigma_i^{(t)} = (\mathbf{r}_i, z_i, \mathbf{h}_i^{(t)}) \quad (4)$$

where this represents an atom of element z_i at position \mathbf{r}_i with learnable features $\mathbf{h}_i^{(t)}$. When performing a forward pass in an MPNN, multiple message construction, update and readout steps are required. Firstly, a message is constructed by aggregating or pooling information from its neighbours in the local environment $\mathcal{N}(i)$:

$$\mathbf{m}_i^{(t)} \bigoplus_{j \in \mathcal{N}(i)} M_t(\sigma_i^{(t)}, \sigma_j^{(t)}) \quad (5)$$

where M_t is a learnable message function and the pooling operation, such as a sum, is permutation invariant as this is a required physical symmetry [22]. Next, the update step uses a learnable update function U_t to transform the message into new features:

$$\mathbf{h}_i^{(t+1)} = U_t(\sigma_i^{(t)}, \mathbf{m}_i^{(t)}). \quad (6)$$

After message construction and update steps have been repeated T times, the site energy E_i is obtained by applying the learnable readout functions \mathcal{R}_t to the node states $\sigma_i^{(t)}$:

$$E_i = \sum_{t=1}^T \mathcal{R}_t(\sigma_i^{(t)}). \quad (7)$$

In terms of the other required invariances, the MPNN only acts on interatomic distances, result-

ing in translational and rotational invariance [22]. However, in equivariant MPNNs, as used for the MACE architecture, the use of relative position vectors instead of distance scalars results in equivariance with respect to rotation of the internal features $\mathbf{h}_i^{(t)}$ [29]. Previous examples such as Tensor Field Networks [30] and NequIP [29] achieved superior performance upon incorporating equivariance but struggled with high computational cost due to using 4-6 message passing iterations [22].

Furthermore, multi-ACE frameworks have previously yielded high accuracy via increasing the depth of the neural network to 4-6 layers with 2-body messages or by using just one layer but with a higher body order expansion [22]. When using message passing, increasing the number of layers enables the sharing of information within a larger receptive field than the cutoff distance. This access to semi-local information helps MPNNs make accurate predictions but scales very poorly with the number of layers used [2].

MACE, however, combines these two approaches to use four-body messages, only needing two message-passing layers to reach high accuracy [22]. The use of higher body messages reduces the need for more message-passing layers whilst still capturing complex interactions, enabling accurate simulation of non-trivial systems. Therefore, the MACE architecture represents a balance between simplicity and flexibility, which has enabled it to generalise and scale so effectively when used as a foundation model [27]. Even though message passing allows MACE to incorporate semi-local information, locality is still a major limitation of MACE. This can have a large impact for systems where long-range dispersion or electrostatic interactions play a key role. However, it is possible to add D3 dispersion corrections to the foundation model by specifying this as an argument, however, a training set with a more robust functional than PBE would improve the model's reliability across a wider range of systems.

The MACE-MP-0 foundation model was created from training the MACE architecture on the MPTrj dataset. This is a dataset created from the Materials Project (MP) containing ~ 1.6 million relaxations and static calculations from $\sim 150,000$ bulk inorganic crystals, using the Perdew-Burke-Enzerhof (PBE) exchange-correlation functional, spanning 89 elements [9]. The foundation model is trained using these PBE energies, forces and

stresses, where 2,600 GPU hours were required to train the medium version of the model for 200 epochs on NVIDIA A100 GPUs[22].

Since the foundation model’s release, there have been new variations trained on different datasets such as MACE-MPA-0 and MACE-OMAT-0. MACE-MPA-0 is trained on the combined MPTrj and the sAlex dataset [12], consisting of \sim 12 million structures from \sim 3.4 million materials, and most recently MACE-OMAT-0 was trained on the even larger Open Materials 24 dataset consisting of \sim 100 million structures from \sim 3.2 million materials [12].

For MACE-MP-0, three different-sized models were trained: small, medium, and large, which increase in computational cost and accuracy when going from small to medium to large. The only difference between the models is the order of their message passing equivariance, L. The large model ($L=2$) provides a higher order, more descriptive representation of the system than the medium ($L=1$) or small ($L=0$) models, but at a higher computational cost [8].

3. Results Part 1: Graphene and Ti_2CO_2 MXene 2D Materials

Before combining the two 2D materials, it is essential to first evaluate the performance of MACE-MP-0, along with MACE-MPA and MACE-OMAT-0, on each material separately. These initial tests will provide insight into which foundation model is most suited to the system and help make hypotheses about performance on the more complex heterostructure. Firstly, $5 \times 5 \times 1$ graphene and $4 \times 4 \times 1$ Ti_2CO_2 MXene supercells were created as Atomic Simulation Environment (ASE) Atoms objects [34], with unit cell lattice parameters of 2.46 \AA and 3.03 \AA respectively. This setup will result in a lattice mismatch of less than 2% when the

| Bond | Optimised length (\AA) | Ref. length (\AA) |
|-------|-----------------------------------|------------------------------|
| Ti-C | 2.187 | 2.186 |
| Ti-O | 1.976 | 1.972 |
| Ti-Ti | 3.030 | 3.032 |

Table 1: Bond lengths of the optimised Ti_2CO_2 MXene structure compared to literature reference values [31]. All MACE models achieved the same optimised bond lengths to three decimal places.

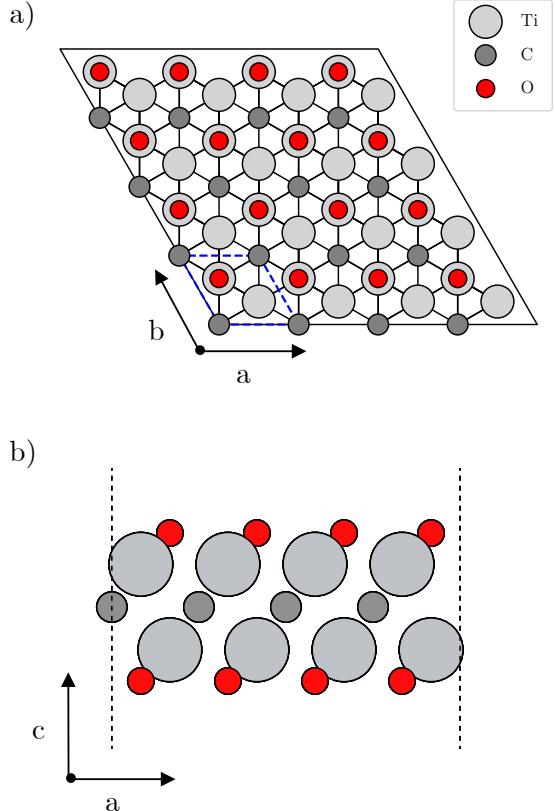


Figure 1: View of the Ti_2CO_2 MXene along the c-axis (a), with the unit cell shown in blue, and along the b-axis (b). Structures have been geometry optimised via BFGS with $F_{max} = 10^{-4} \text{ eV}/\text{\AA}$.

heterostructures are later created, as in the reference paper [35]. The structures were then geometry optimised using each variant of MACE. The Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimisation algorithm [36] was used until the maximum force acting on any atom did not exceed $10^{-4} \text{ eV}/\text{\AA}$ to ensure accurate relaxation [37]. Figure 1 shows the resulting relaxed MXene structure.

To validate the relaxed structures, basic structural properties can be calculated. The optimised thickness of the MXene was calculated as 4.46 \AA for all models, aligning closely to a D2 corrected PBE reference value of 4.45 \AA [38]. In both cases, the thickness was calculated as the average z-distance between the top and bottom oxygen layers. Focusing on bond lengths, Table 1 shows that the MACE-optimised bond lengths for MXene agree well with the reference values. As the Ti-O bond has a z component, its overestimation will have been a ma-

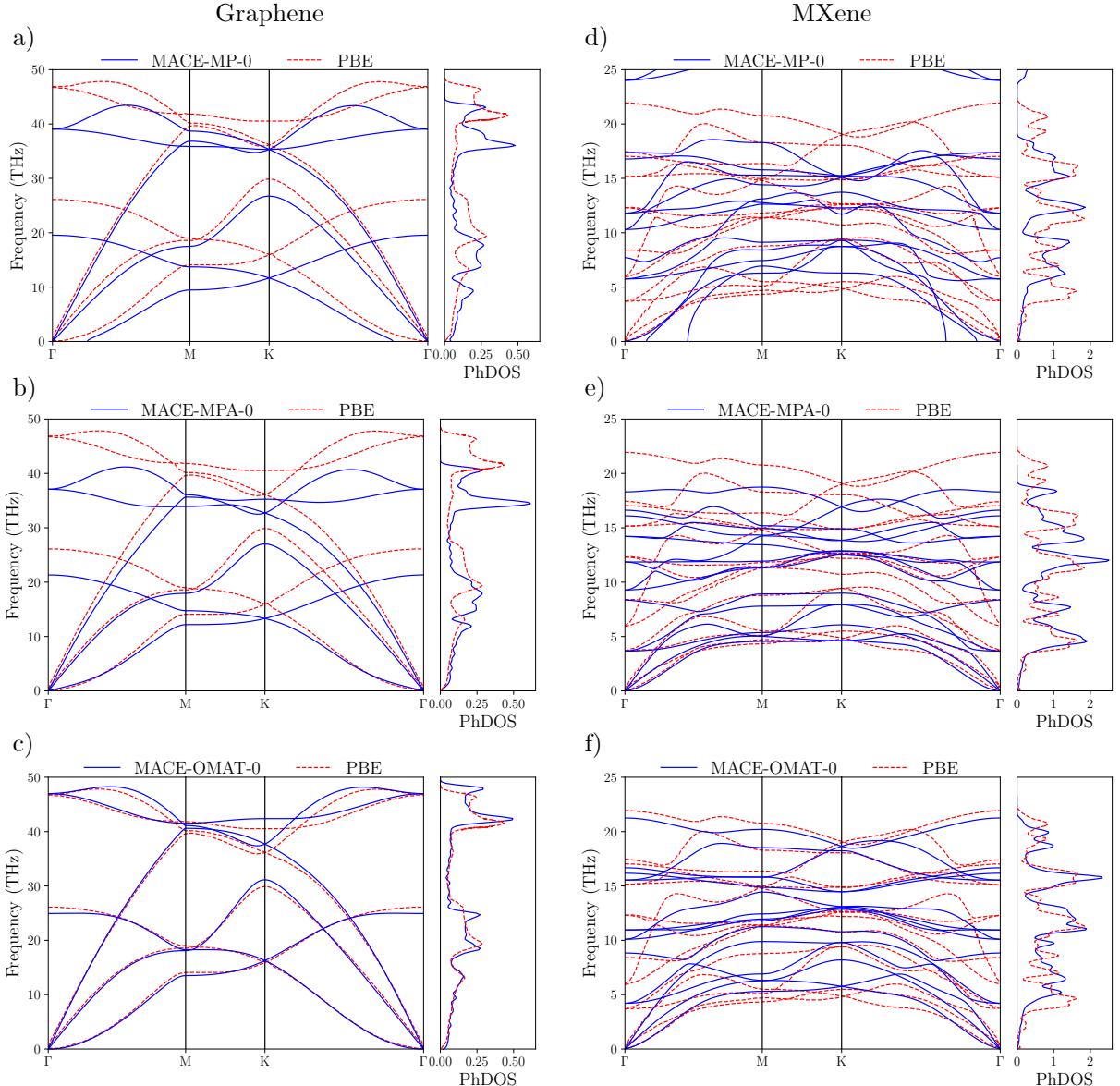


Figure 2: Comparison of the phonon dispersion and phonon density of states (PhDOS) calculated by PBE, MACE-MP-0, MACE-MPA-0 and MACE-OMAT-0 for graphene (a-c) [32] and Ti_2CO_2 MXene (d-e) [33]. Phonons and PhDOS were calculated on geometry optimised structures using PHONOPY, with $4 \times 4 \times 1$ and $5 \times 5 \times 1$ supercells used for MXene and graphene respectively. Phonon frequencies were computed along the high symmetry path Γ -M-K- Γ at 300 points. The phonon density of states was evaluated using a $20 \times 20 \times 1$ q-point mesh.

ajor contributor to the slightly larger than expected layer thickness.

Calculating the phonon dispersion is a more informative test, allowing conclusions on how well the MACE variants capture the local curvature of the PES and predict the material's stability [39]. Phonon dispersion and density of states were cal-

culated for graphene and Ti_2CO_2 MXene (Figure 2) via finite displacement using the PHONOPY library [40, 41]. For graphene, the improvement in performance when using MACE-OMAT-0 compared to MP-0 or MPA-0 variants is immediately obvious, especially when predicting higher frequency modes. The phonon density of states (Ph-

DOS) gives a simpler indication of performance, with MACE-OMAT-0 closely tracing the PBE reference. Comparing MP-0 and MPA-0, MP-0 predicts phonons closer to the reference for higher frequency modes, and MPA-0 performs better for lower frequency modes. Furthermore, MP-0 gives rise to imaginary frequencies around the Γ -point, which are not seen for MPA-0 or OMAT-0. This suggests that MP-0 fails to capture the stability of graphene's lattice, potentially due to insufficient training data on 2D materials. For classifying imaginary frequencies the same tolerance as Batatia et al. of -0.1 THz was used [8]. When benchmarking MACE-MP-0's phonon dispersion accuracy, Batatia et al. found that the highest frequency was often underestimated compared to PBE reference data [8]. This is true for MP-0 and MPA-0, however, OMAT-0 slightly overpredicts the highest frequency.

For Ti_2CO_2 MXene, the phonon dispersion is more complex due to having five atoms in its primitive cell instead of just two in graphene. MACE-MP-0 produces imaginary frequencies, a greatly overestimated maximum frequency and overall poor phonon dispersion prediction even qualitatively. MPA-0 offers a drastic improvement, with OMAT-0 improving even further, with both yielding no imaginary frequencies. Looking at the PhDOS, OMAT-0 appears to be qualitatively correct, capturing the overall shape of the curve. MPA-0 also performs well but lacks the performance of OMAT-0 at high frequencies. Clearly, the 2D Ti_2CO_2 MXene is far outside of MACE-MP-0's training set, but MPA-0 and OMAT-0 contain materials which more closely resemble its structure and elemental combination.

Overall, MACE-OMAT-0 outperforms the other MACE variants for phonon dispersion and PhDOS prediction, for both graphene and MXene. Therefore, it is the most promising MACE variant for modelling the complex heterostructure and will be the foundation model of choice for this investigation.

4. Results Part 2: Graphene/MXene Heterostructures

Modelling of the graphene/MXene heterostructure is a more complex task but will serve as a challenging test, with the potential of being a vital benchmark for MACE foundation models in the future. Firstly, the dominant attractive interactions between the two sheets will be dispersion

forces. As the foundation model is not trained on a functional with dispersion corrections, the empirical Grimme's D3 dispersion correction is applied via TorchDFTD3Calculator to aid in modelling these interactions [42]. This helps create realistic dynamics, as running MD without the attractive dispersion forces results in the two sheets separating. However, training or fine-tuning on a functional with dispersion corrections would result in much more accurate modelling of the heterostructure.

Here, three MXene/graphene heterostructures (Figures 3a-c) are investigated for their potential as anode materials: MXene/graphene (M/G), MXene/epoxylated graphene (M/GO), and MXene/hydroxylated graphene (M/GOH). The interlayer spacing between the graphene and MXene layers is a good indicator of the material's performance as an anode, with larger spacings linked to increased sodium ion storage capacity and diffusion [43]. As in the reference paper, the interlayer spacing is defined as the distance between the average z-coordinate of the carbon atoms in the graphene layer and the average z-coordinate of the layer of oxygen atoms facing graphene (indicated by the arrows in Figures 3a-c). To prevent interactions between periodic images, a 20 Å vacuum was added in the z direction. To inspect how the interlayer spacing varies with time, 25 ps NVT simulations were run in ASE on the geometry-optimised heterostructures, consisting of a 5 ps pre-equilibration and a 20 ps production run, as in the reference [35]. The Bussi thermostat, an extension of the Berendsen thermostat, was used as more consistent temperature control was observed for this system compared to alternative thermostats [44]. Stable MD at 300 K was observed and tested up to 100 ps using MACE-OMAT-0 applied to each heterostructure system.

The interlayer spacing in Figure 3d fluctuates around an average value, showing that it has reached equilibrium. Going from M/G to M/GO to M/GOH the average interlayer spacing increases, qualitatively agreeing with the PBE reference [35]. However, quantitatively, MACE-OMAT-0 overestimates this value by 0.85 Å, 0.34 Å and 0.25 Å for M/G, M/GO and M/GOH respectively. This is a significant distance, especially for M/G, and has the potential to impact observables dependent on interlayer spacing, such as Na ion diffusion. A likely explanation is that the D3 dispersion correction has undercorrected for dispersion forces, resulting in weaker than expected attraction between the two

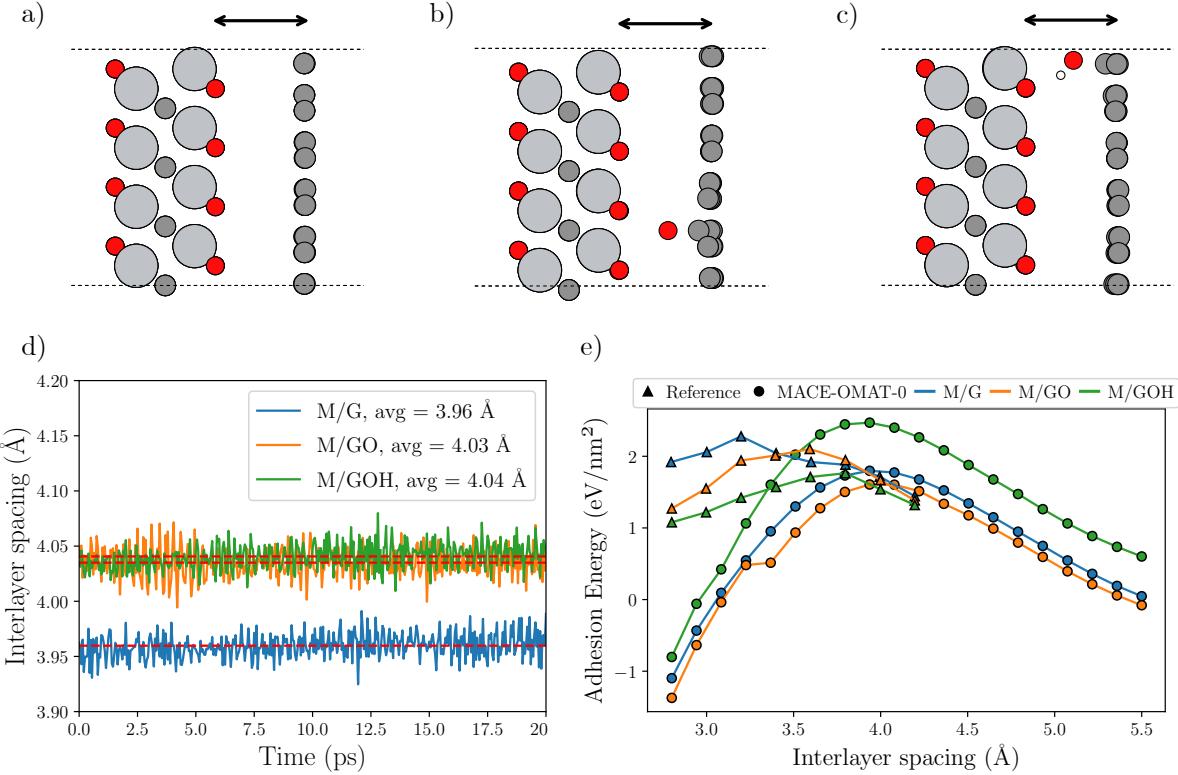


Figure 3: Comparison of interlayer spacing and adhesion energies between heterostructures. (a)-(c) shows the geometry optimised heterostructures, M/G (a), M/GO (b) and M/GOH (c). The interlayer distance measurement is indicated by the arrow between the average z coordinate of the graphene atoms and the average coordinate of the MXene oxygen layer facing graphene. (d) shows the fluctuation of interlayer spacing over a 20 ps NVT production run for the heterostructures, along with their equilibrium spacing in the legend. (e) shows two adhesion energy curves for the three different heterostructures, computed using MACE-OMAT-0 and the PBE reference. Maximum adhesion energies and their corresponding interlayer spacings: M/G ($1.79 \text{ eV}/\text{nm}^2$, 3.96 \AA), M/GO ($1.61 \text{ eV}/\text{nm}^2$, 3.99 \AA), M/GOH ($2.47 \text{ eV}/\text{nm}^2$, 3.89 \AA).

layers and therefore an equilibrium distance further apart. The addition of the -O and -OH groups incorporates dipole-dipole and hydrogen-bonding interactions, reducing the reliance of the interlayer spacing on the accuracy of the dispersion forces, resulting in lower deviation from the PBE reference.

To quantify the interlayer interactions, the adhesion energy was calculated as the difference in energy between the separate structures and the heterostructure, per unit interfacial area A [35]:

$$E_{\text{adhesion}} = \frac{(E_{\text{MXene}} + E_{\text{G}}) - E_{\text{G/MXene}}}{A}. \quad (8)$$

E_{G} , E_{MXene} and $E_{\text{G/MXene}}$ are the MACE-OMAT-0 calculated energies for graphene (or fractionalised graphene), MXene and the heterostructure respec-

tively. E_{adhesion} is defined such that a positive value indicates that the heterostructure is more stable than the separated sheets. The adhesion energy is a balance of attractive dispersion forces (all heterostructures), dipole-dipole interactions (M/GO, M/GOH), and hydrogen bonding (M/GOH) with Pauli and electrostatic repulsion. These are all interactions which MACE has been shown to struggle with, however, as the system is relatively small, most of the atoms will be within the 12 \AA receptive field, making this a good initial test system.

To calculate the adhesion energy for a particular interlayer spacing d , ASE's FixedPlane constraint is applied to the graphene and MXene layers in order to fix d during geometry optimisation. The energy of the optimised structure $E_{\text{G/MXene}}$ was then cal-

culated with MACE-OMAT-0. This was repeated for 20 d values for each heterostructure to produce a smooth curve (Figure 3e).

Qualitatively, OMAT-0 predicts the same trend as the PBE reference for M/G and M/GO, where M/GO has a lower maximum adhesion energy, located at a larger interlayer spacing than M/G. In the case of M/GOH, the trend seen in the reference is broken, where OMAT-0 predicts the adhesion energy to be stronger, and to hold layers at a shorter distance than for M/G and M/GO. This discrepancy arises from OMAT-0's inability to accurately capture the hydrogen bonding interactions between the -OH attached to graphene and MXene's oxygens. However, the interlayer spacing at the peak adhesion energy shows the closest quantitative agreement with the reference. This is likely not too meaningful due to the qualitative disagreement. In theory, the interlayer spacing at which the peak adhesion energy is seen should be the heterostructure's equilibrium spacing (Figure 3d) as it represents the most energetically favourable distance for the materials to be apart. In the PBE reference, the interlayer spacing at the maximum adhesion energy is lower than its equilibrium spacing. MACE-OMAT-0 reproduces these findings, which is promising but suggests a higher level of theory than PBE is required to thoroughly model this system.

Furthermore, it is important to question the thoroughness of the reference data. The reference adhesion energy is a sharply peaked curve, where taking the maximum value will result in an unreliable estimate of the interlayer spacing due to the sparsity of data points. Previous studies, such as Polfus et al. in 2021 [45], on the adhesion energy of Van der Waals (layered) materials such as graphene, H-BN and MoS₂ result in curves more closely resembling the smooth MACE-OMAT-0 curves in Figure 3e than the reference data curves containing a sharper peak. Overall, the main conclusion from this plot is that it confirms the presence of van der Waals forces between the graphene and MXene layers. This provides confidence that MD simulations will yield at least qualitatively correct physics, where fine-tuning of the foundation model would be required to achieve quantitative accuracy.

5. Results Part 3: Sodium Ion diffusion Within the Heterostructures

Since MACE-OMAT-0 overall performs qualitatively well for these heterostructures, particularly M/G and M/GO, the next step is to investigate sodium ion diffusion between the two sheets. The same simulation procedure used in section 4 was first applied to one sodium ion intercalated between the graphene and MXene sheets and then multiple ions for a realistic battery scenario. To generate more robust observables with less noise, the simulation time from the reference was extended from 25 ps to 100 ps. This takes advantage of the fact MLPs are less computationally expensive than ab initio methods, as 100 ps may not have been feasible for the researchers who conducted the original investigation, depending on their resource and time constraints.

Figures 4a-c show the mean squared displacement (MSD) of the Na ion in the three heterostructures over the last 95 ps, calculated via a moving average window. Figures 4d-f show the path of the Na ion during the simulation. For M/G, the MSD and diffusion map show that the Na ion remains confined in a gap between three oxygens for the entire duration, as observed in the reference study. This behaviour is likely due to the small interlayer spacing for M/G resulting in a large activation barrier for diffusion, as seen in studies of other anode materials [46]. Furthermore, being surrounded by three electronegative oxygen atoms is a stable configuration for the positive Na ion, increasing the barrier further. Potentially, simulating at higher temperatures would provide the thermal energy for the Na ion to overcome this energy barrier and diffuse, however, significantly higher temperatures ($T > 370$ K) are not practical for sodium ion battery applications [47]. Therefore, M/G is likely not a suitable anode material.

M/GO has a larger interlayer spacing due to the additional oxygen atom in between the layers, substantially lowering the diffusion activation barrier. As a result, the Na ion exhibits a discrete site hopping mechanism, where the ion spends significant time in each stabilised site with quick transitions or hops between them. This is shown visually via the path of the Na ion (Figure 4e) and the different levels seen in the MSD (Figure 4b). The reference shows the same behaviour, however, they observe more frequent hopping and no backwards hopping, which is observed here by the decrease in

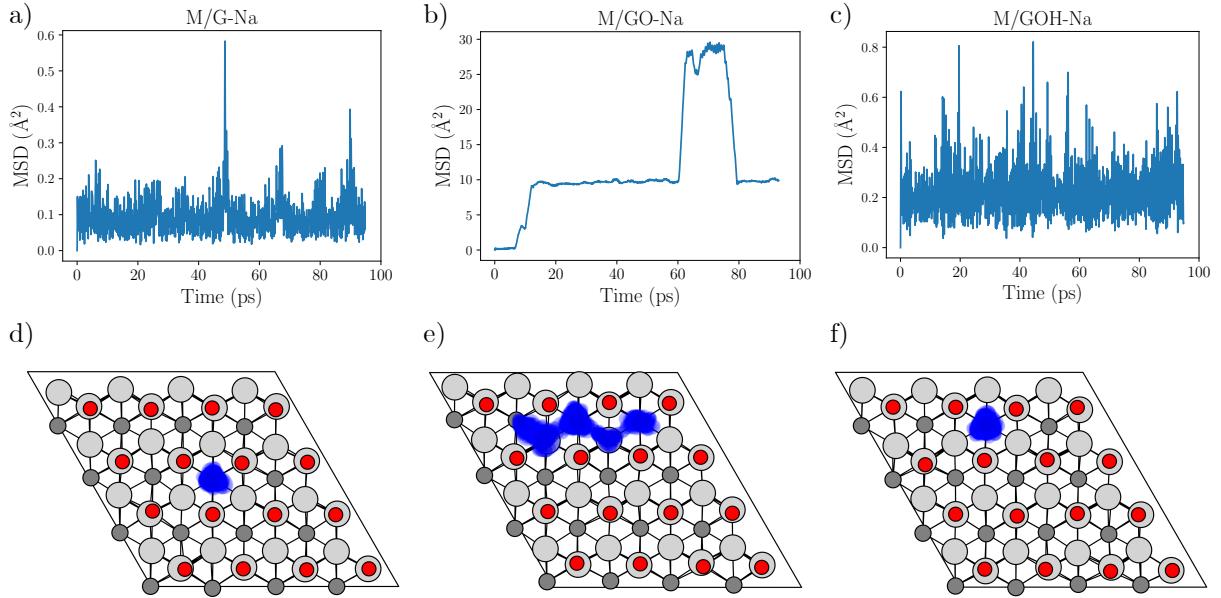


Figure 4: Diffusion of a single sodium ion within the heterostructures. (a-c) show the mean squared displacement (MSD) of the Na ion in each structure over the last 95 ps of the 100 ps simulation, calculated with a moving window average. (d-f) show the corresponding path of the Na ion, with the graphene layer removed for visibility. This is the position of the sodium ion at every simulation frame in the last 95 ps overlaid.

the MSD at 80 ps. This difference is likely due to OMAT-0 overestimating the activation barrier for hopping, making hopping events more rare. However, MACE-OMAT-0 performs impressively well considering how far this task is outside its training set. The observed diffusion shows potential for M/GO as an anode material.

Diffusion in M/GOH is very similar to M/G, where the Na ion stays fixed at one site, whereas in the reference it exhibits the same site hopping mechanism as in M/GO. However, more variation in the MSD is seen than in M/G which is qualitatively correct. OMAT-0’s underprediction of the interlayer spacing has resulted in an increased diffusion energy barrier, constraining the Na ion to a single site.

To further inspect the interlayer spacing during Na diffusion, Figures 5a, 5d, and 5g show a snapshot of Na intercalation in the heterostructure and Figures 5c, 5f and 5i show an interlayer spacing heatmap. The heatmaps show the average interlayer spacing over the last 95 ps, which is calculated by finding the spacing for each section of an 8 x 8 grid within the supercell. For all heterostructures, distortion of the graphene layer around the Na ion and its path is seen, resulting in increased interlayer

spacing. This effect is known as pillaring, and for a single Na ion, only the nearby graphene atoms are pillared [35]. In M/G, the rest of the graphene compensates for this upwards distortion by distorting slightly downwards. This results in graphene encapsulating the Na ion, hindering its movement. In M/GO, greater upwards and downwards distortion is seen, as the presence of the Na ion increases the interlayer spacing whilst the attraction between GO’s oxygen and the electron-deficient Ti atoms of MXene decreases the spacing around the oxygen [48]. The interlayer spacing ranges from 3.76 Å around the oxygen to 4.33 Å around Na. This larger interlayer spacing aids Na diffusion, reducing its energy barrier.

In M/GOH, the effect is similarly pronounced as the -OH group forms hydrogen bonds with the top layer of MXene oxygens. The interlayer spacing around the Na ion is just 0.02 Å less than for M/GO, but the minimum spacing is 0.09 Å lower due to the strong hydrogen bonding. Given that this distortion is relatively similar to M/GO but no diffusion is seen, this shows that the Na ion is likely very close to overcoming the diffusion energy barrier to exhibit the expected diffusion behaviour. In this case, a small quantitative error has resulted in a

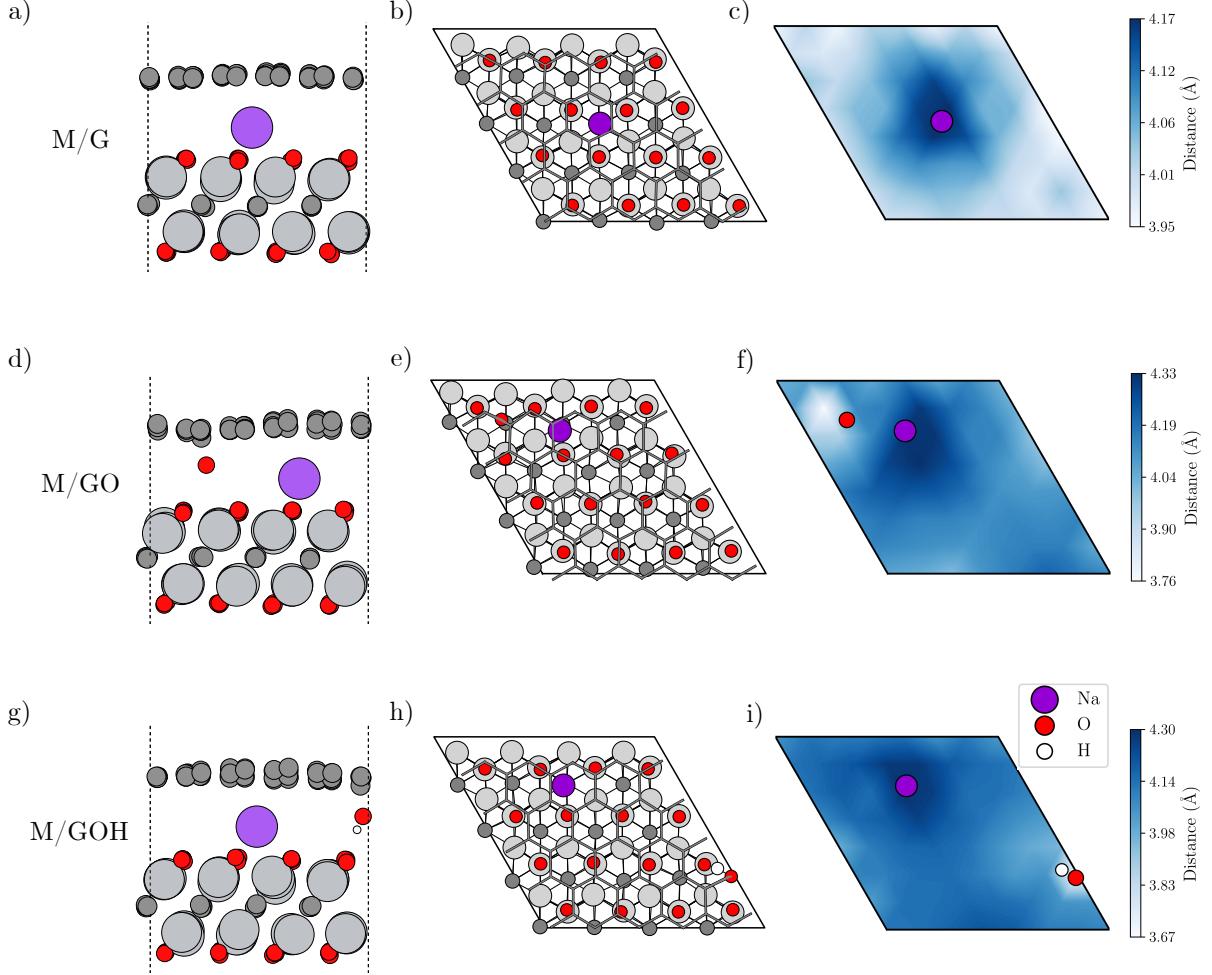


Figure 5: The diffusion of one intercalated sodium ion diffusion and its effect on the interlayer spacing of M/G (a-c), M/GO (d-f) and M/GOH (g-i) heterostructures. The first column (a, d, g) shows a view of the last simulation frame along the b-axis, and the second column (b, e, h) shows the c-axis view. The last column (c, f, i) shows the interlayer spacing heatmaps of each structure, calculated by averaging the interlayer spacing over the last 95 ps of the simulation.

qualitative error.

To simulate a more realistic battery situation, eight Na ions were intercalated into the heterostructure to achieve half monolayer (ML) coverage, or half as many Na ions as oxygens in Mxene's top layer. Here, more substantial pillarizing of the graphene layer occurs, where the Na ions are able to effectively support the graphene layer, shifting it upwards in the z-direction. The layer shifts more uniformly, observing less variation in interlayer spacing compared to with one Na ion. This increased spacing promotes Na ion diffusion, however, there are now more Na ions present and their

repulsive interactions can interfere with the diffusion process. Therefore diffusion will be a balance of these two aspects. Figures 6c, 6f, 6i show the average MSD of the Na ions in each heterostructure. All the heterostructures show the discrete site hopping mechanism seen before with M/GO and a single Na ion. M/G and M/GOH show increased diffusion compared to a single Na ion, where the increase in interlayer spacing allows the ions to overcome the diffusion energy barrier [46]. M/GO shows decreased diffusion, suggesting the increased number of Na ions could be hindering diffusion and that a decrease in the ML ratio of Na ions may be

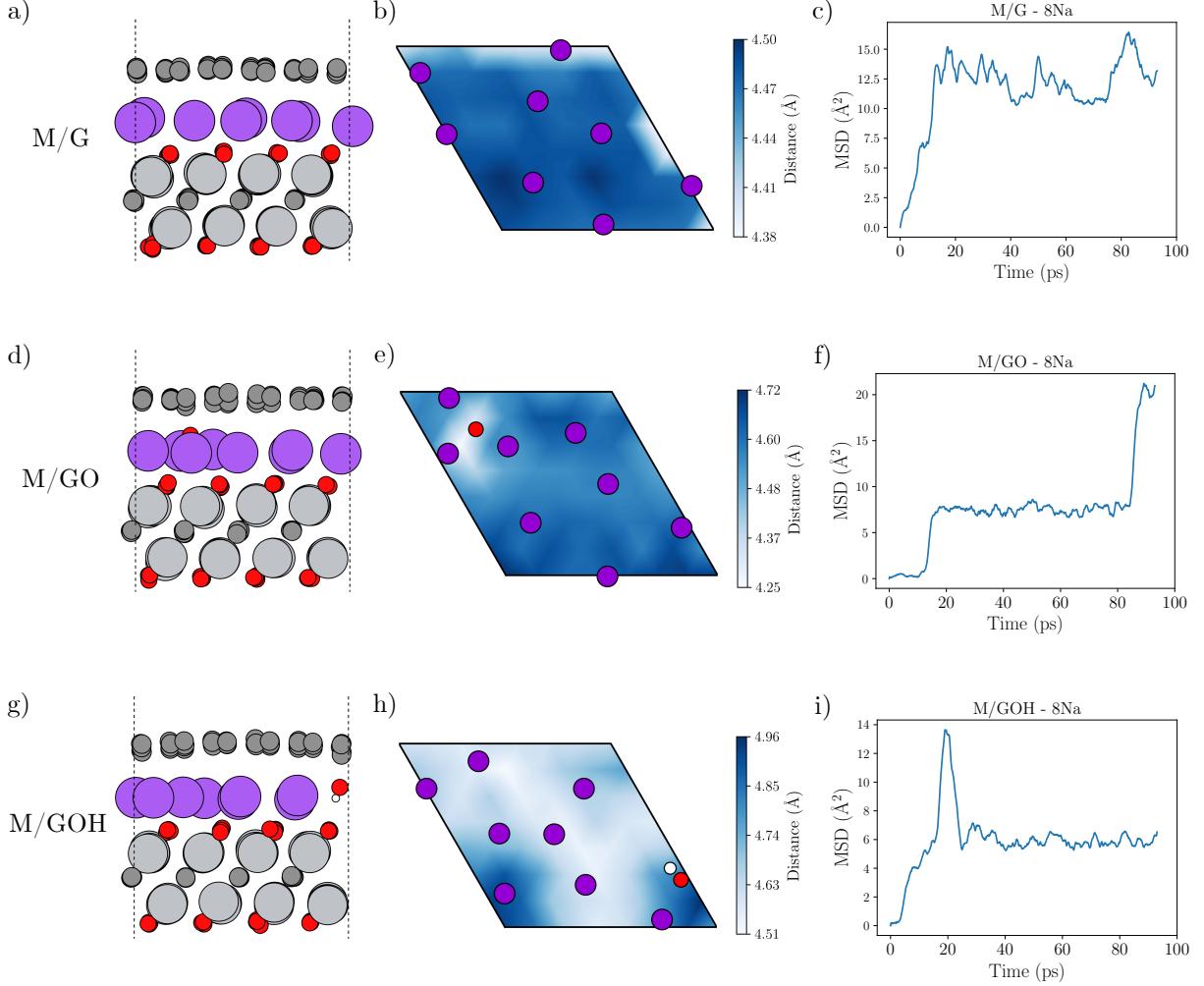


Figure 6: The diffusion of 8 sodium ions intercalated within heterostructures M/G (a-c), M/GO (d-f) and M/GOH (g-i), and their effect on the interlayer spacing of the heterostructures. The first column (a, d, g) shows a view of the last simulation frame along the b-axis. The second column (b, e, h) shows the interlayer spacing heatmaps of each structure, averaged over the last 95 ps of the simulation (post-equilibration). The last column (c, f, i) shows the average MSD of the 8 Na ions, determined using tracking the position of each ion, taking a moving window average over the last 95 ps and then calculating the average MSD.

optimal. The hopping mechanism with M/GO is still well defined, suggesting that the Na ions hop in a concerted fashion.

Based on these MSD results, M/GO exhibits the greatest potential for a sodium ion battery anode material due to the highest extent of diffusion and consistent performance with a varying number of Na ions. This conclusion agrees with the conclusion made from the reference investigation using ab initio MD at the PBE level of theory.

6. Conclusions

In summary, MACE-OMAT-0 performed remarkably well at modelling sodium ion diffusion through a stack of two different 2D materials, overall achieving high qualitative agreement with the PBE reference. Essentially, the *same qualitative conclusions* as the ab initio-based reference paper have been made whilst saving substantial computational costs. In terms of the other MACE variants, judging on the phonon dispersion prediction performance, it is likely that MACE-MPA would have

also performed well. Although phonon dispersion is a good test, further tests with MACE-MPA-0 would be required to validate this hypothesis. Furthermore, MACE-MP-0 was unable to form physical predictions of the MXene and therefore it is unlikely to be suitable for the heterostructure.

This investigation was a challenging task for numerous reasons. Firstly, dispersion interactions play a crucial role in all the heterostructures, and the MACE models are only trained on PBE data with the option for an empirical dispersion correction afterwards. Electrostatic, dipole-dipole and hydrogen bonding interactions also became significant for some heterostructures, which are not taken into account in the training datasets either. Furthermore, the training set lacks examples on 2D materials. Essentially, this investigation has tested MACE-OMAT-0 not on only a challenging material, but also on challenging intermolecular interactions. Despite this, the qualitative performance has remained remarkably accurate.

The major weak point has been modelling hydrogen bonding, which is a common issue in the literature [8]. To make progress on hydrogen bonding or to reach quantitative accuracy, fine-tuning of the foundation model would be required using dispersion-corrected functionals such as revPBE-D3 or hybrid functionals such as revPBE0-D3. It has been shown that less than 100 additional training examples would be required to reach quantitative accuracy when investigating sublimation enthalpies [49], which is much less computationally expensive than training a foundation model with a new functional. Furthermore, further work should concentrate on understanding Na ion diffusion energy barriers, and how they can be decreased. The nudged elastic band (NEB) method could be used to study these barriers by looking at the transition state between two stabilised sites [50]. This will bring us one step closer in designing the best sodium-ion battery anode possible.

Acknowledgements

I would like to thank Dr Christoph Schran for his supervision throughout the project and Isaac Parker for helpful discussions on performing MD simulations.

Supplementary Information

Simulation and analysis scripts can be found at:
https://github.com/joehart2001/WA_MACE-MP-0

References

- [1] R. Sundararaman, D. Vigil-Fowler, K. Schwarz, Improving the accuracy of atomistic simulations of the electrochemical interface, *Chemical reviews* 122 (12) (2022) 10651–10674.
- [2] F. L. Thiemann, N. O’neill, V. Kapil, A. Michaelides, C. Schran, Introduction to machine learning potentials for atomistic simulations, *Journal of Physics: Condensed Matter* 37 (7) (2024) 073002.
- [3] R. Martin-Barrios, E. Navas-Conyedo, X. Zhang, Y. Chen, J. Gulín-González, An overview about neural networks potentials in molecular dynamics simulation, *International Journal of Quantum Chemistry* 124 (11) (2024) e27389.
- [4] S. Mukherjee, M. Pinheiro Jr, B. Demoulin, M. Barbatti, Simulations of molecular photodynamics in long timescales, *Philosophical Transactions of the Royal Society A* 380 (2223) (2022) 20200382.
- [5] V. L. Deringer, M. A. Caro, G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, *Advanced Materials* 31 (46) (2019) 1902765.
- [6] A. C. Van Duin, S. Dasgupta, F. Lorant, W. A. Goddard, Reaxff: a reactive force field for hydrocarbons, *The Journal of Physical Chemistry A* 105 (41) (2001) 9396–9409.
- [7] P. Rowe, V. L. Deringer, P. Gasparotto, G. Csányi, A. Michaelides, An accurate and transferable machine learning potential for carbon, *The Journal of Chemical Physics* 153 (3) (2020).
- [8] I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, et al., A foundation model for atomistic materials chemistry, *arXiv preprint arXiv:2401.00096* (2023).
- [9] B. Deng, P. Zhong, K. Jun, J. Riebesell, K. Han, C. J. Bartel, G. Ceder, Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling, *Nature Machine Intelligence* 5 (9) (2023) 1031–1041.
- [10] C. Chen, S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, *Nature Computational Science* 2 (11) (2022) 718–728.
- [11] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, E. D. Cubuk, Scaling deep learning for materials discovery, *Nature* 624 (7990) (2023) 80–85.
- [12] L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick, Z. W. Ulissi, Open materials 2024 (omat24) inorganic materials dataset and models, *arXiv preprint arXiv:2410.12771* (2024).
- [13] Y.-L. Liao, B. Wood, A. Das, T. Smidt, EquiformerV2: Improved equivariant transformer for scaling to higher-degree representations, *arXiv preprint arXiv:2306.12059* (2023).
- [14] S. Aflaki, S. Netessine, Strategic investment in renewable energy sources: The effect of supply intermit-

- tency, Manufacturing & Service Operations Management 19 (3) (2017) 489–507.
- [15] Z. Zhu, T. Jiang, M. Ali, Y. Meng, Y. Jin, Y. Cui, W. Chen, Rechargeable batteries for grid scale energy storage, Chemical Reviews 122 (22) (2022) 16610–16751.
- [16] H. Zhang, Y. Gao, X. Liu, L. Zhou, J. Li, Y. Xiao, J. Peng, J. Wang, S.-L. Chou, Long-cycle-life cathode materials for sodium-ion batteries toward large-scale energy storage systems, Advanced Energy Materials 13 (23) (2023) 2300149.
- [17] J. Chen, G. Adit, L. Li, Y. Zhang, D. H. Chua, P. S. Lee, Optimization strategies toward functional sodium-ion batteries, Energy & Environmental Materials 6 (4) (2023) e12633.
- [18] S. Qiao, Q. Zhou, M. Ma, H. K. Liu, S. X. Dou, S. Chong, Advanced anode materials for rechargeable sodium-ion batteries, ACS nano 17 (12) (2023) 11220–11252.
- [19] P. Ma, D. Fang, Y. Liu, Y. Shang, Y. Shi, H. Y. Yang, Mxene-based materials for electrochemical sodium-ion storage, Advanced Science 8 (11) (2021) 2003185.
- [20] S.-u. Chae, S. Yi, J. Yoon, J. C. Hyun, S. Doo, S. Lee, J. Lee, S. J. Kim, Y. S. Yun, J.-H. Lee, et al., Highly defective $\text{Ti}_3\text{C}_{2}\text{Tx}$ -mxene-based fiber membrane anode for lithium metal batteries, Energy Storage Materials 52 (2022) 76–84.
- [21] I. Demiroglu, F. M. Peeters, O. Gülsen, D. Çakır, C. Sevik, Alkali metal intercalation in mxene/graphene heterostructures: a new platform for ion battery applications, The journal of physical chemistry letters 10 (4) (2019) 727–734.
- [22] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, G. Csányi, Mace: Higher order equivariant message passing neural networks for fast and accurate force fields, Advances in neural information processing systems 35 (2022) 11423–11436.
- [23] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, Physical Review B 99 (1) (2019) 014104.
- [24] Y. Lysogorskiy, A. Bochkarev, M. Mrovec, R. Drautz, Active learning strategies for atomic cluster expansion models, Physical Review Materials 7 (4) (2023) 043801.
- [25] G. Dusson, M. Bachmayr, G. Csányi, R. Drautz, S. Etter, C. van Der Oord, C. Ortner, Atomic cluster expansion: Completeness, efficiency and stability, Journal of Computational Physics 454 (2022) 110946.
- [26] W. C. Witt, C. van der Oord, E. Gelžinytė, T. Järvinen, A. Ross, J. P. Darby, C. H. Ho, W. J. Baldwin, M. Sachs, J. Kermode, et al., Acepotentials.jl: A julia implementation of the atomic cluster expansion, The Journal of Chemical Physics 159 (16) (2023).
- [27] N. Bernstein, From gap to ace to mace, arXiv preprint arXiv:2410.06354 (2024).
- [28] D. P. Kovács, J. H. Moore, N. J. Browning, I. Batatia, J. T. Horton, V. Kapil, W. C. Witt, I.-B. Magdáu, D. J. Cole, G. Csányi, Mace-off23: Transferable machine learning force fields for organic molecules, arXiv preprint arXiv:2312.15211 (2023).
- [29] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, B. Kozinsky, E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, Nature communications 13 (1) (2022) 2453.
- [30] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, P. Riley, Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds, arXiv preprint arXiv:1802.08219 (2018).
- [31] Y. Wang, J. Yao, Q. Zhou, D. Su, W. Ju, L. Wang, Transition-metal doped Ti_2Co_2 as gas sensor toward NH_3 : A dft study, Chemical Physics 586 (2024) 112376.
- [32] X. Yang, D. Han, H. Fan, M. Wang, M. Du, X. Wang, First-principles calculations of phonon behaviors in graphether: a comparative study with graphene, Physical Chemistry Chemical Physics 23 (1) (2021) 123–130.
- [33] S.-H. Yin, X.-H. Li, R.-Z. Zhang, H.-L. Cui, First-principles calculations of the phonon, elastic, and thermoelectric properties of a Ti_2Co_2 monolayer, ACS omega 8 (50) (2023) 48549–48566.
- [34] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbørg, J. Kubal, K. Kaasbørg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Petersen, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, The atomic simulation environment—a python library for working with atoms, Journal of Physics: Condensed Matter 29 (27) (2017) 273002.
- [35] X. Qiu, Y. Zheng, H. Li, K. Qu, R. Li, Tuning na-ion diffusion in mxene/graphene oxide heterostructures: An ab initio molecular dynamics study, Langmuir 40 (25) (2024) 13155–13166.
- [36] J. D. Head, M. C. Zerner, A. broyden—fletcher—goldfarb—shanno optimization procedure for molecular geometries, Chemical physics letters 122 (3) (1985) 264–270.
- [37] J. Kermode, C. Ortner, N. Hine, M. Probert, L. Mones, Preconditioned geometry optimisers for the castep and onetep codes, Chem. Phys 144 (2016) 164109.
- [38] A. K. Mandia, N. A. Koshi, B. Muralidharan, S.-C. Lee, S. Bhattacharjee, Electrical and magneto-transport in the 2d semiconducting mxene Ti_2Co_2 , Journal of Materials Chemistry C 10 (23) (2022) 9062–9072.
- [39] I. Pallikara, P. Kayastha, J. M. Skelton, L. D. Whalley, The physical significance of imaginary phonon modes in crystals, Electronic Structure 4 (3) (2022) 033002.
- [40] A. Togo, L. Chaput, T. Tadano, I. Tanaka, Implementation strategies in phonopy and phono3py, J. Phys. Condens. Matter 35 (35) (2023) 353001. doi:10.1088/1361-648X/acd831.
- [41] A. Togo, First-principles phonon calculations with phonopy and phono3py, J. Phys. Soc. Jpn. 92 (1) (2023) 012001. doi:10.7566/JPSJ.92.012001.
- [42] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu, The Journal of chemical physics 132 (15) (2010).
- [43] B. Chen, S. Sui, F. He, C. He, H.-M. Cheng, S.-Z. Qiao, W. Hu, N. Zhao, Interfacial engineering of transition metal dichalcogenide/carbon heterostructures for electrochemical energy applications, Chemical Society Reviews 52 (22) (2023) 7802–7847.
- [44] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, The Journal of chemical physics 126 (1) (2007).
- [45] J. M. Polfus, M. B. Muñiz, A. Ali, D. A. Barragan-Yani,

- P. E. Vullum, M. F. Sunding, T. Taniguchi, K. Watanabe, B. D. Belle, Temperature-dependent adhesion in van der waals heterostructures, *Advanced Materials Interfaces* 8 (20) (2021) 2100838.
- [46] J. Ma, C. Yang, X. Ma, S. Liu, J. Yang, L. Xu, J. Gao, R. Quhe, X. Sun, J. Yang, et al., Improvement of alkali metal ion batteries via interlayer engineering of anodes: from graphite to graphene, *Nanoscale* 13 (29) (2021) 12521–12533.
- [47] Z. Li, Y. Zhang, J. Zhang, Y. Cao, J. Chen, H. Liu, Y. Wang, Sodium-ion battery with a wide operation-temperature range from- 70 to 100 c, *Angewandte Chemie International Edition* 61 (13) (2022) e202116930.
- [48] K. P. Marquez, K. M. D. Sisican, R. P. Ibabao, R. A. J. Malenab, M. A. N. Judicpa, L. Henderson, J. Zhang, K. A. S. Usman, J. M. Razal, Understanding the chemical degradation of $\text{Ti}_3\text{C}_2\text{Tx}$ mxene dispersions: A chronological analysis, *Small Science* 4 (10) (2024) 2400150.
- [49] H. Kaur, F. Della Pia, I. Batatia, X. R. Advincula, B. X. Shi, J. Lan, G. Csányi, A. Michaelides, V. Kapil, Data-efficient fine-tuning of foundational models for first-principles quality sublimation enthalpies, *Faraday Discussions* 256 (2025) 120–138.
- [50] L. Shi, T. Zhao, A. Xu, J. Xu, Ab initio prediction of a silicene and graphene heterostructure as an anode material for li-and na-ion batteries, *Journal of Materials Chemistry A* 4 (42) (2016) 16377–16382.