Journal of Physics: Energy

An accurate machine learning calculator for the lithium-graphite system

Mohammad Babar[1] , **Holden L Parks**[1] , **Gregory Houchins**[2]
and Venkatasubramanian Viswanathan[1,2]

[1] Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America
[2] Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America

E-mail: venkvis@cmu.edu

Abstract

Machine-learning potentials are accelerating the development of energy materials, especially in identifying phase diagrams and other thermodynamic properties. In this work, we present a neural network potential based on atom-centered symmetry function descriptors to model the energetics of lithium intercalation into graphite. The potential was trained on a dataset of over 9000 diverse lithium–graphite configurations that varied in applied stress and strain, lithium concentration, lithium–carbon and lithium–lithium bond distances, and stacking order to ensure wide sampling of the potential atomic configurations during intercalation. We calculated the energies of these structures using density functional theory (DFT) through the Bayesian error estimation functional with van der Waals correlation exchange-correlation functional, which can accurately describe the van der Waals interactions that are crucial to determining the thermodynamics of this phase space. Bayesian optimization, as implemented in Dragonfly, was used to select optimal set of symmetry function parameters, ultimately resulting in a potential with a prediction error of 8.24 meV atom^{-1} on unseen test data. The potential can predict energies, structural properties, and elastic constants at an accuracy comparable to other DFT exchange-correlation functionals at a fraction of the computational cost. The accuracy of the potential is also comparable to similar machine-learned potentials describing other systems. We calculate the open circuit voltage with the calculator and find good agreement with experiment, especially in the regime $x \geq 0.3$, for x in Li_xC_6. This study further illustrates the power of machine learning potentials, which promises to revolutionize design and optimization of battery materials.

1. Introduction

Graphite-based anodes have been crucial to the success of lithium-ion battery technology [1, 2]. Despite many alternate material options, graphite remains the most common anode material due to its low intercalation potential, high gravimetric capacity, cycling stability, reversible capacity, and low cost [2]. Due to its importance in the context of lithium-ion batteries, lithium intercalation into graphite has been studied extensively through both experiments [3–6] and theoretical calculations [7–11].

As graphite is a layered material, lithium intercalation is understood to occur in 'stages' [12, 13], where a stage n compound contains $n-1$ empty layers between lithium-containing layers. For example, a fully-charged graphite anode (LiC_6) is a stage 1 compound having no empty layers, whereas after some lithium de-intercalation it transitions to a stage 2 compound having alternating lithiated and empty layers. While the structures of highly-lithiated phases are well-accepted, several factors make it difficult to determine structures for lower-lithium phases ($0 < x < 0.3$ in Li_xC_6) [5]. Previous studies have reported different stable stacking orders [14] and phases depending on the temperature [4] and current rate [15, 16] during synthesis. In addition, low-lithium phases are highly disordered [9, 17]. Elucidating the intercalation mechanisms is

essential for understanding typical lithium-ion battery operation as well as the occurrence of lithium plating on the anode surface at low temperatures or high current rates [9].

Previous theoretical efforts into understanding the lithium-graphite phase diagram have generally fallen into two groups. The first approach uses density functional theory (DFT) to predict the intercalation energy of different phases [10, 11, 18]. The DFT predictions are accurate and can explicitly account for van der Waals effects [10], but they are too expensive to perform for more than a handful of structures. Therefore, in constructing voltage and phase diagrams from these calculations, assumptions must be made about the phases present at each lithium concentration. The second approach relies on a cheap and fast energy calculator to search over many configurations to determine the convex hull of minimum-energy phases [7–9]. While an empirical potential can be used [19], the calculator is typically a cluster expansion model in which the system is modeled as a lattice containing lithium occupation sites, i.e. [8], $E_{\text{struct}} = \sum_i V\sigma_i + \sum_{i,j} V_{i,j}\sigma_i\sigma_j$, where σ_i is the occupation of the site i, V is the energy associated with Li occupation, and $V_{i,j}$ models the interactions between two Li atoms occupying sites i and j in the lattice model. This approach is much cheaper computationally than DFT, and it is possible to evaluate a large number of structures to determine the minimum-energy phase at each lithium concentration. However, this approach is limited by the accuracy of the cluster expansion model, which typically only models Li occupation and Li–Li interactions within some cutoff radius. It is not straightforward to extract other thermodynamic properties such as vibrational contributions and structural properties from cluster expansion models.

In recent years, machine learning (ML) approaches based on DFT have been used for materials discovery in the battery space. For example, candidate screening for solid electrolytes for lithium-metal batteries was performed by Sendek *et al* using likelihood-based logistic regression [20] and by Ahmad *et al* using crystal graph convolutional neural networks [21]. Joshi *et al* employed a variety of ML techniques (deep neural networks, support vector machines, kernel ridge regression) to learn electrode voltages in metal-ion batteries using features from the Materials Project database [22]. Machine-learned atomistic potentials have also been used to model a variety of battery materials [18, 23, 24]. ML potentials have grown increasingly popular as they have approached the accuracy of *ab initio* methods at a fraction of the cost [25]. These potentials have been used to great effect in a variety of systems, including molecules [26, 27] and periodic systems with one [28–31] or multiple species [32, 33]. A strength of ML potentials are their flexibility, as there are many featurizations and model frameworks that can be used depending on the amount of data available or the system being studied [34]. Because studies of battery materials require many evaluations of large supercells, ML approaches will likely play a larger role in the future due to their accuracy and low computational cost [35].

In this work, we present an atomistic neural network potential with atom-centered symmetry function descriptors to model lithium intercalation into graphite. The potential was trained on DFT data generated using the Bayesian error estimation functional with van der Waals correlation (BEEF-vdW) exchange-correlation functional [36], which explicitly accounts for van der Waals interactions and accurately describes covalent and ionic interactions. To ensure optimal hyper-parameters when training the potential, hyper-parameter optimization was performed using Bayesian optimization through the DRAGONFLY package [37]. We first use the model to predict structural and elastic properties to confirm its accuracy in comparison to a handful of generalized-gradient-approximation (GGA)-level [38] DFT exchange-correlation functionals. We then use the calculator to predict the open circuit voltage as a function of lithium fraction by simulating the lithium-vacancy ordering using grand canonical Monte Carlo simulations. Due to the speed and low cost of the ML potential, we are able to evaluate the effect of different stacking orders at low lithium concentrations to accurately determine the convex hull. We find that the calculator accurately reproduces the experimental open circuit voltage, with the an error less than 0.1 V for moderate-to-high lithium concentration, $x \geq 0.3$. The calculator predicts a mix of ABC- and AB-stacked structures for low x structures, in agreement with previous studies that found low-lithium structures to have variable stacking. Previous studies using machine-learned potentials for the lithium-graphite system evaluated the diffusivity of lithium in graphite [24, 39] and the structural properties [39]. To our knowledge, this is the first application of a machine-learned atomistic potential to determine the open-circuit voltage of the lithium-graphite system.

2. Methods

2.1. Training data
A set of DFT calculations were used to the train the neural network potential. The dataset consisted of a total of 9189 structures, which were separated into training, validation, and test data in an 8-1-1 ratio. The structures were generated using the following methods to ensure diverse chemical environments in terms of strain and stress, lithium concentration, Li-C and Li-Li distances, supercell size, and defects.

- 1290 structures were generated by populating the graphitic structures made available by Wen and Tadmore [30] with between 1 and 8 Li atoms. The original structures are comprised of graphite with varied inter-layer and intra-layer lattice constants, structures with point defects, and structures from *ab initio* molecular dynamics calculations.
- 733 equilibrium lithium-graphite structures made available by Pande and Viswanathan [9] with Li inside AA hexagonal sites in different stages and compositions ($0 \leq x \leq 1$).
- 6456 structures of varying supercell sizes, up to a maximum size of $3 \times 3 \times 5$, using a 6 carbon atom unit cell. The structures varied in layer separation (between 2.5 and 5 Å), in-plane compression and tension ($\pm 10\%$ of the BEEF-vdW lattice constant), and stacking (AA or AB). Li atoms were inserted at random positions in the layer while ensuring all Li-C distances were at least 0.5 Å. We performed relaxations on a subset of these structures using a BFGS optimization [40] until the magnitude of the force on any atom was at most 0.05 eV Å^{-2}. Of the 6456 structures, 4448 contained between 1 and 20 Li. Of these 4448 structures, 3043 are intermediate structures from a BFGS relaxation. The remaining 1648 structures have between 25 and 45 Li and are intermediate structures from a BFGS relaxation.
- 1070 AA-stacked $4 \times 4 \times 2$ supercells using a 6 atom unit cell with between 1 and 48 Li at hexagonal centers. The structures were compressed and stressed in the in-plane and out-of-plane directions and were included specifically so that the potential could learn the energy-volume relationship that plays a significant role in the grand canonical Monte Carlo simulations in section 2.5.

The data were generated using the real-space projector-augmented wave method [41, 42] as implemented in the GPAW DFT code [43, 44]. We use the BEEF-vdW exchange-correlation functional [36], which is chosen due to its accurate description of covalent, ionic, and van der Waals forces [36]. We used a real-space grid spacing of 0.16 Å and a (12, 12, 4) Monkhorst–Pack k-point grid for a conventional graphite unit cell. For larger supercells, the number of k-points are reduced to keep this same k-point density. For each DFT calculation, electron densities were converged to 10^{-4} electrons and energies were converged to 5×10^{-4} eV electron^{-1}.

2.2. Model architecture

In this paper, we use the atom-based featurization scheme suggested by Behler and Parrinello [45, 46] as implemented in the Atomistic Machine-learning Package (AMP) [47]. In this scheme, the total energy of a material composed of n atoms is decomposed into the sum of energetic contributions from individual atoms,

$$E(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) = \sum_{i}^{n} E(\mathbf{r}_i), \tag{1}$$

where $\mathbf{r}_i$ denotes the Cartesian coordinates of atom i, for $1 \leq i \leq n$. To calculate the contribution of atom i to the total energy, the Cartesian coordinates are transformed using 'symmetry functions.' This process converts the coordinates into a transitionally- and rotationally-invariant representation. The transformed coordinates are then used as inputs to a fully-connected feed forward neural network to calculate the atomic energy used in the sum in (1). The weights of the network are the same for a given chemical element but vary between chemical elements, which guarantees the total energy is not affected by an interchange of atoms [45]. The symmetry functions in the Behler–Parrinello scheme for atom i with neighbors indexed by j and k are given by

$$G_i^2 = \sum_{j \neq i} e^{-\eta(r_{ij}-r_s)^2/r_c^2} f_c(r_{ij}), \tag{2}$$

and

$$G_i^4 = 2^{1-\zeta} \sum_{j \neq i} \sum_{k \neq j} (1 + \lambda \cos\theta_{ijk})^\zeta e^{-\eta(r_{ij}^2 + r_{jk}^2 + r_{ik}^2)/r_c^2} \times f_c(r_{ij}) f_c(r_{jk}) f_c(r_{ik}), \tag{3}$$

where

$$f(r_{ij}) = \begin{cases} 0.5 \left[\cos\left(\frac{\pi r_{ij}}{r_c}\right) \right] & r_{ij} \leq r_c, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

and where r_{ij} is the distance between atoms i and j; r_c is the cutoff radius; θ_{ijk} is the angle formed by r_{ij} and r_{ik}; and η, ζ, and λ are hyper-parameters affecting the shapes of G_i^2 and G_i^4. We discuss a method for optimizing these hyper-parameters for our system in section 2.3. The first symmetry function, G_i^2, models the

Table 1. Summary of the parameters used to construct the G_i^2 and G_i^4 symmetry functions.

Symmetry function	η	r_s	r_c (Å)	λ	ζ
G^2	1.2, 1.3, 1.9, 10.8, 11.7, 17.1	0	6.5		
G^4	1.2, 1.3, 1.9, 10.8, 11.7, 17.1		6.5	$\pm 1, \pm 3, \pm 6$	4

interactions of atoms i and j, while the second symmetry function, G_i^4, models all three-body interactions of atom i with neighbors j and k. Both symmetry functions do not model interactions over distances greater than r_c.

To balance between computational efficiency in training the model and an accurate description of the varying chemical environments of the system, we use a combined total of 48 G_i^2 and G_i^4 symmetry functions. Table 1 summarizes the parameters used to generate the G_i^2 and G_i^4 symmetry functions. Each chemical element was modeled with a neural network consisting of three hidden layers with 30 nodes per layer, in addition to the input and output layers.

After the featurization of the atomic positions using the symmetry functions, the potential is trained to minimize the error on the training data by minimizing the loss function [47]

$$\mathcal{L} = \sum_{\alpha=1}^{M} \left(\frac{\hat{E}_\alpha - E_\alpha}{n_\alpha} \right)^2, \tag{5}$$

where M is the total number of atomic configurations (or energy data points) in the training data, $\hat{E}_\alpha$ and E_α are the predicted and reference energies of configuration α, and n_α is the number of atoms in atomic configuration α. Using the hyper-parameters determined in section 2.3, we trained the model for 1000 iterations. The root mean square error (RMSE) for the training and validation sets as a function of training epoch for the neural network potential is shown in figure 1. The calculator with the best performance on the validation data occurred after 820 training epochs with a performance of 8.41 meV atom^{-1} on the validation set and 8.24 meV atom^{-1} on the test set. A parity plot of the energy predictions of the potential is shown in figure S1 in the supplemental material [48].[3]

We performed an additional intuitive test of the final ML calculator on structures characterized using two handcrafted descriptors: the average of shortest lithium-carbon bond length for all lithium atoms (R_1) and the integral of radial distribution function (RDF) of lithium atoms weighted by inverse Li–Li distances (R_2). We generated 500 $3 \times 3 \times 2$ AA graphite structures and populated them with lithium atoms at random positions in the middle of the two layers. When all lithium atoms are in the same z-plane, the descriptor R_1 identifies the average position of lithium atoms with respect to hexagonal center sites of AA graphene layers. The descriptor R_2 is calculated by dividing the RDF, which measures the density of lithium atoms within a particular radius, by the radius of separation and integrating this quantity over all lithium atoms. In effect, a higher R_2 value indicates smaller average separation of lithium atoms, and vice versa. The variation of energies of the structure calculated using the ML model and the equations for R_1 and R_2 are shown in section S2. We observe lower energies for the structures with higher and lower values of R_1 and R_2, respectively, implying a greater stability for structures with lithium atoms closer to the hexagonal centers and further from each other on average. Moreover, the energies vary with respect to R_2 in an approximate straight line behaviour, correctly signifying a coulombic interaction amongst lithium atoms in graphite.

2.3. Hyper-parameter optimization

In order to achieve optimal performance of the neural network potential with respect to the (unseen) test data, the best values for the hyper-parameters $\eta, \zeta, \lambda, r_c$ and r_s must be determined. We assume atom-centric behaviour of the symmetry functions for both Li and C, thus keeping $r_s = 0$. The cutoff radius r_c is also left at its default value, 6.5 Å, because adjacent graphene layers are well within this range and negligible long-range interaction is assumed with the next-nearest layers. We identify η to be the dominant parameter for controlling the radial resolution for both two-body and three-body symmetry functions. For the three-body

---

[3]See the supplemental material for the performance of the AMP potential on test data, the performance of the potential on structures differing in the R_1 and R_2 descriptors described in section 2.2, a discussion of training the network on forces and the effect of structural relaxation on the energy, a short tutorial on accessing the potential and the training data used in this work, and a calculation of vibrational contributions to the free energy.
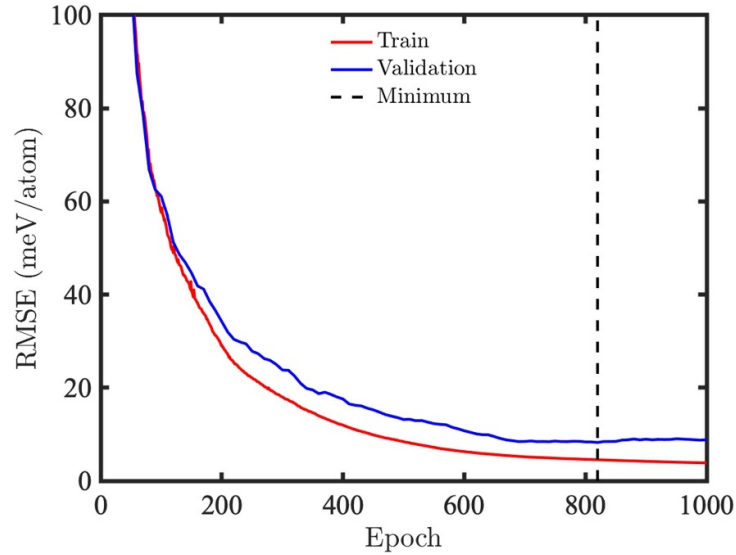
Figure 1. The root mean squared error (RMSE) on the training and validation datasets as a function of training epoch. The best calculator occurs at epoch 820 with an RMSE of 8.41 meV atom^{-1} on the validation data.

symmetry functions, we modulate λ within integer values (± 1 to ± 10) to obtain different locations of cosine maxima and keep ζ constant, as it has a similar effect on the function as η [46].

While these limits on the hyper-parameters reduces the number of potential combinations, an extensive search for the best combination is computationally infeasible. We therefore employ the Bayesian optimization package DRAGONFLY [37] to determine the best set of hyper-parameters. DRAGONFLY uses an adaptive sampling method to change its acquisition function, as the performance of an acquisition function may vary depending on the optimization task. The acquisition functions included Gaussian process upper confidence bound [48, 49], Thompson sampling [50], expected improvement [51], and top-two expected improvement [52]. We ran the optimization routine for 100 acquisition runs. For each run, the potential was trained on the energy to an accuracy of 5 meV per atom. The potential achieving this accuracy was then passed to DRAGONFLY, which determined the set of hyper-parameters for the next run based on the performance of the potential on the validation set and the chosen acquisition function. This process is shown in figure 2 with labels for the acquisition function for each evaluation. The best model found using this algorithm, summarized in table 1, had an RMSE error of 8.24 meV atom^{-1} on the test set. The optimized values for η differ greatly, as half of the symmetry functions have $\eta < 2.0$ and the other half have $\eta > 10.0$. This leads to an identification of long- and short-range interactions by the low and high magnitude η symmetry functions, respectively.

2.4. Structural properties
As a first test of the accuracy of the AMP model, we calculated structural and elastic properties of graphite and several lithium-graphite structures using the AMP model and a handful of GGA-level exchange-correlation functionals. The lattice constants of AA- and AB-stacked graphite, a, d_{AA}, and d_{AB}, where a is the in-plane lattice constant and $d_{AA/AB}$ is the inter-layer separation, were determined by fitting a two-dimensional fourth-order polynomial in a and d_{AA} for AA graphite and in a and d_{AB} for AB graphite [53]. To fit this polynomial, for each compound we calculated the energy of 36 structures having in-plane lattice constants between 0.8 and 1.2 times the experimental value of $a = 2.461$ Å and out-of-plane lattice constants between 0.8 and 1.4 times the experimental value ($d_{AB} = 3.353$ Å [54] and $d_{AA} = 3.44$ Å [55]). The values of the lattice constants correspond to the minimum energy of the fitted polynomial. The lattice constants of the lithium-graphite phases $Li_{0.13}C_6$, $Li_{0.33}C_6$, $Li_{0.5}C_6$, and LiC_6 were also predicted using this method.

The elastic constants for AB-stacked graphite were calculated using the following energy derivatives with respect to the lattice constants [54]:

$$C_{11} + C_{12} = \frac{1}{c_0 \sqrt{3}} \cdot \frac{\partial^2 E}{\partial a^2}, \tag{6}$$

$$C_{13} = \frac{1}{a_0 \sqrt{3}} \cdot \frac{\partial^2 E}{\partial a \partial c}, \tag{7}$$
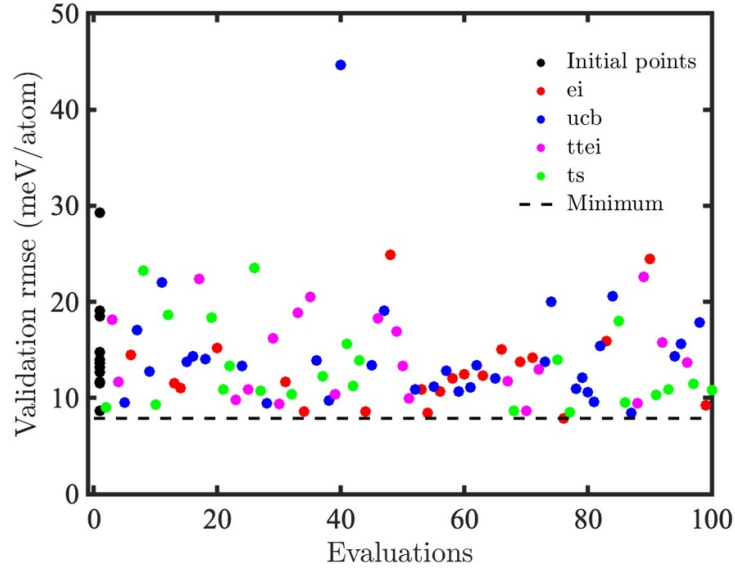
Figure 2. DRAGONFLY evaluations given a budget of 100 evaluations using the expected improvement (ei), upper confidence bound (ucb), top-two expected improvement (ttei), and Thompson sampling (ts) acquisition functions. The performance of the best calculator in the DRAGONFLY evaluations, 8.41 meV atom^{-1}, is denoted with the dotted line.

$$C_{33} = \frac{2c_0}{a_0^2\sqrt{3}} \cdot \frac{\partial^2 E}{\partial c^2}, \tag{8}$$

$$C^t = \frac{(C_{11}+C_{12})+2C_{33}-4C_{13}}{6}, \tag{9}$$

$$B_0 = \frac{C_{33}(C_{11}+C_{12})-2C_{13}^2}{6C^t}, \tag{10}$$

where C^t is the tetragonal shear modulus, B_0 is the equilibrium bulk modulus, and a_0 and $c_0 = 2d_{AB}$ are the equilibrium lattice constants. The derivatives were calculated using a central finite difference of the energies, which required three energy evaluations in (6) and (8) and four energy evaluations in (7). We used a displacement size of 0.01 times the equilibrium lattice constant for each finite difference formula.

2.5. Grand canonical Monte Carlo

We employed a grand canonical Monte Carlo (GCMC) simulation to calculate the open circuit voltage as a function of Li insertion by identifying stable lithium-vacancy ordering at each composition. A fully-lithiated AA-stacked graphite conventional cell with 6 carbon atom basis was used as a reference structure for the simulation input and to identify intercalation sites. We fixed the supercell size as $3 \times 3 \times 5$ ($12.816 \times 12.816 \times 16.75$ Å^3 rhombohedral cell) times the conventional cell to keep a balance between accuracy and computational cost. The vacant sites were identified from the positions of lithium in the fully-lithiated input structure. Subsequently, all Li were removed and the lattice constants of the pristine structure (*a* and *c*) were optimized using an equation of state (EOS) fit. For this optimization, we evaluated internal energies at 11 inter-layer separations between 3.5 and 4.5 Å and 11 strain values between 90% and 110% of the in-plane lattice constant. The optimized structure was then fed into the GCMC simulation, which uses the ML potential to estimate energy.

To account for the potentially disordered stacking in the dilute regime [5], low-lithium structures were tested for alternate stacking orders by having empty layers shifted into AB or AC stacking. The simulation was run for 200 sweeps. In every sweep, one Li is added or removed at a random vacant site with equal probability, and alternate stackings of empty layers are tested to get the new state. The addition/removal is accepted based on the formation energy of the new state E_2 relative to the previous state energy E_1, given by the probability

$$P_{\text{insert}} = \min\left[1, \frac{V}{\Lambda^3(n+1)}\exp\left(-\beta[E_2-E_1-\mu_{Li}]\right)\right], \tag{11}$$

$$P_{\text{remove}} = \min\left[1, \frac{n\Lambda^3}{V}\exp\left(-\beta[E_2-E_1-\mu_{Li}]\right)\right], \tag{12}$$

where Λ is the thermal de Broglie wavelength, μ_{Li} is the chemical potential of lithium, n is the number of Li in the material, $\beta = 1/k_B T$ ($T = 300$ K), and V is the volume available for Li intercalation given by $V = V_{cell} - \frac{4}{3} n \pi r_{Li}^3$ [56], where $r_{Li} = 182$ pm [57] is the van der Waals radius of Li. Every sweep is accompanied by $N/2$ internal swaps of Li to identify stable Li ordering on the convex hull, where N is the number of carbon atoms in the supercell. The cell parameters for the ordered structure are optimized using an EOS fit to account for changes in lattice constants after each successful addition or removal of Li. Near the end of the simulation, the number of internal swaps is increased by increasing the number of sweeps and decreasing the acceptance probability. This is done to ensure sufficient sampling of the high-lithiation structures, which have converging energies and less flexible ordering (i.e. low configurational entropy).

We use the same methodology as Pande and Viswanathan [9] to calculate the chemical potential of Li in reference to a material with well known formation energy [58]. The core criterion for selection of the reference compound is to match the oxidation state for the system under study. Formation energy of lithium can be corrected with reference to LiC$_6$ as $\mu_{Li} = E_{LiC_6}^{DFT} - E_{C_6}^{DFT} - \Delta H_f^{exp}$. We insert the DFT references from Pande and Viswanathan's calculations [9] and experimental value of formation enthalpy $\Delta H_f^{exp}(LiC_6) = -0.114$ eV from Reynier *et al* [6].

2.6. Entropic contribution

We consider vibrational and configurational entropy in determining the Gibbs free energy of lithium-graphite system, $\Delta G(T) = \Delta H(T) - T \Delta S(T)$, in order to account for finite-temperature effects on the open circuit voltage. Since pressure work is negligible in comparison to internal energy change [59], the enthalpy of formation is approximated with 0 K single-point (DFT or AMP) energies and vibrational effects.

To account for vibrational effects on the Gibbs free energy as a function of lithiation, we repeated the phonon calculations from Pande and Viswanathan [9] using the BEEF-vdW exchange-correlation functional. Following the procedure outlined by Pande and Viswanathan, we performed density functional perturbation calculations to calculate the finite-temperature vibrational enthalpy and entropy. We then fit a fourth-order polynomial expression to the Gibbs free energy values and used this polynomial to calculate the vibrational contribution to the free energy for each phase. See section S5 for details about this calculation.

Configurational entropy can be obtained using the following expression,

$$S_{conf} = k_B \ln \Omega, \tag{13}$$

where Ω is the total number of lithium-ordering microstates of equal energy. There are $N/2$ possible intercalation lattice points for a supercell comprising of N carbon atoms. Hence for dilute phases, the subsequent lithium atoms can occupy a position anywhere, excluding the positions adjacent to other Li due to strong interactions between them [9]. As a result, we identify 7 in-plane (6 adjacent and one occupied site) and 2 out of plane adjacent sites around a lithium atom in a layered hexagonal symmetry lattice of graphite that are not available for further intercalation. At higher concentrations, the number of sites adjacent to lithium is large enough to effectively leave only $N/6$ sites out of the $N/2$ total sites as favorable for lithium occupation. As a result, the number of orderings can be obtained by choosing $Nx/6$ sites out of $N/6$ positions for lithium to intercalate at a given composition (x in Li$_x$C$_6$). Choosing a threshold of $x = 0.2$ between low- and high-lithium concentration, the number of possible configurations for each case is given by

$$\Omega = \begin{cases} \dfrac{\prod\limits_{i=0}^{i=n-1} (N/2 - ir)}{n!}, & \text{if } \frac{n}{N/6} < 0.2, \\ \binom{N/6}{n} & \text{if } \frac{n}{N/6} > 0.2, \end{cases} \tag{14}$$

where $r = 9$ is the number of unfavourable positions around a Li atom and n is the number of Li atoms inserted, so $x = \frac{n}{N/6}$. In section 3.2 we examine the effect of including vibrational free energy and configurational entropy on the convex hull of stable phases and the open circuit voltage.
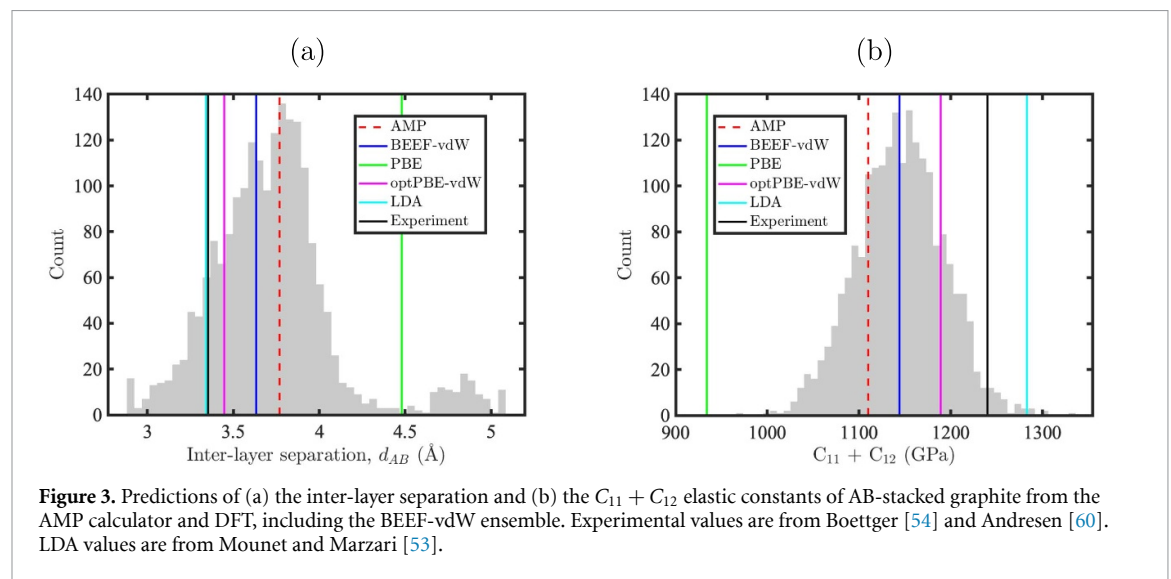
3. Results and discussion

3.1. Structural and thermodynamic properties

The structural and elastic properties of AA- and AB-stacked graphite are summarized in table 2. We compare the AMP predictions to DFT predictions at the GGA-level using the BEEF-vdW, Perdew–Burke–Ernzerhof (PBE) [62], and optPBE-vdW [63] exchange-correlation functionals, and to LDA values from Mounet and Marzari [53]. The spreads reported for the BEEF-vdW values correspond to the standard deviations of the ensemble predictions [36]. In general, we see good performance of the AMP potential in comparison to both DFT predictions and experimental values. The most accurate DFT predictions in comparison to experiment

Table 2. Elastic and structural properties of graphite predicted using the AMP potential and a handful of DFT exchange-correlation functionals. BEEF-vdW spreads correspond to one standard deviation of the ensemble predictions. Properties include the in-plane lattice parameter, a_0; the layer separation of AB- and AA-stacked graphite, d_{AB} and d_{AA}; the elastic constants $C_{11} + C_{12}$, C_{13}, C_{33}; the tetragonal shear modulus, C_t; and the bulk modulus, B_0.

	AMP	BEEF-vdW	PBE	optPBE-vdW	LDA [53]	Experiment
a_0 (Å)	2.461	2.464 ± 0.031	2.463	2.470	2.439	2.461 ± 0.003 [54]
d_{AB} (Å)	3.769	3.633 ± 0.382	4.48	3.447	3.341	3.353 ± 0.001 [54]
d_{AA} (Å)	3.858	3.856 ± 0.375	5.109	3.702		3.44 [55]
$C_{11} + C_{22}$ (GPa)	1110	1144 ± 47	934	1189	1283	1240 ± 40 [60]
C_{13} (GPa)	2.3	-0.9 ± 3.5	-0.5	-5.5	-2.8	15 ± 5 [60]
C_{33} (GPa)	7.5	27 ± 39	-0.7	36	29	36.5 ± 1 [60]
C_t (GPa)	186	200 ± 18	158	215	225	208.8 [61]
B_0 (GPa)	7.5	25.9 ± 36.3	-0.7	35.9	27.8	35.8 [61]

Figure 3. Predictions of (a) the inter-layer separation and (b) the $C_{11} + C_{12}$ elastic constants of AB-stacked graphite from the AMP calculator and DFT, including the BEEF-vdW ensemble. Experimental values are from Boettger [54] and Andresen [60]. LDA values are from Mounet and Marzari [53].

are from the vdW exchange-functionals and LDA, which performs well despite not modeling vdW interactions [53]. The AMP model tends to over-predict lattice constants and under-predict elastic constants in comparison to experiment. This trend is visible in figures 3(a) and (b), which plot the DFT and AMP predictions for the d_{AB} lattice constant and $C_{11} + C_{12}$ elastic constants. However, all AMP predictions fall within one ensemble standard deviation of the BEEF-vdW value, which indicates that the predictions of the AMP calculator are in line with most predictions of the GGA-level exchange-correlation functionals in the ensemble.

Results for structural properties of the $Li_{0.13}C_6$, $Li_{0.33}C_6$, $Li_{0.5}C_6$, and LiC_6 lithium-graphite phases, corresponding to stage 4, stage 3, stage 2, and stage 1 structures, are shown in table 3. Predictions of the inter-layer separation are also plotted figure 4 as a function of lithiation. The AMP calculator and DFT exchange-correlation functionals all predict similar in-plane lattice constants and that the in-plane lattice expands as more Li is inserted. The calculators differ more in their predictions of inter-layer separation. The disagreement is greatest for pristine AB graphite, with the PBE d value lying 2.36σ away from the BEEF-vdW value, where σ is the standard deviation of the BEEF-vdW ensemble. The DFT and AMP values tend to agree more as the lithiation increases, which is also reflected in the ensemble through smaller standard deviations. The predictions for inter-layer separation become more accurate in comparison with experiment for high-lithium phases. As lithiation increases, van der Waals interactions are screened by the inserted lithium atoms. As a result, short-range Li-C and Li-Li interactions dominate the lattice. Such interactions are easier for GGA-level DFT and the AMP calculator to model, resulting in greater prediction accuracy and a smaller ensemble σ for structures with more lithium.

3.2. Convex hull and open-circuit voltage

The GCMC routine was used to determine the convex hull of stable lithium-graphite phases and the corresponding OCV. The curvature of the convex hull and its minima give an indication of the stability of phases and lithium saturation fraction, respectively. The OCV for a given composition x is the chemical potential of a Li in Li_xC_6 in reference to bulk lithium [6]. As such, the convex hull and OCV curves are

Table 3. The structural properties of Li_xC_6 for $x = 0.13, 0.33, 0.5, 1$ predicted by the AMP calculator and DFT. The BEEF-vdW spreads correspond to one standard deviation of the ensemble predictions. The calculators tend to agree at higher lithium concentrations, with the widest spread in predictions occurring for pristine graphite.

|		AMP	BEEF-vdW	PBE	optPBE-vdW
$Li_{0.13}C_6$	a (Å)	2.464	2.452 ± 0.031	2.454	2.473
	d (Å)	3.866	3.630 ± 0.250	4.221	3.518
$Li_{0.33}C_6$	a (Å)	2.470	2.472 ± 0.031	2.474	2.481
	d (Å)	3.864	3.738 ± 0.183	3.880	3.654
$Li_{0.5}C_6$	a (Å)	2.474	2.477 ± 0.031	2.479	2.485
	d (Å)	3.874	3.746 ± 0.138	3.844	3.696
LiC_6	a (Å)	2.489	2.494 ± 0.032	2.495	2.501
	d (Å)	3.767	3.736 ± 0.090	3.766	3.726
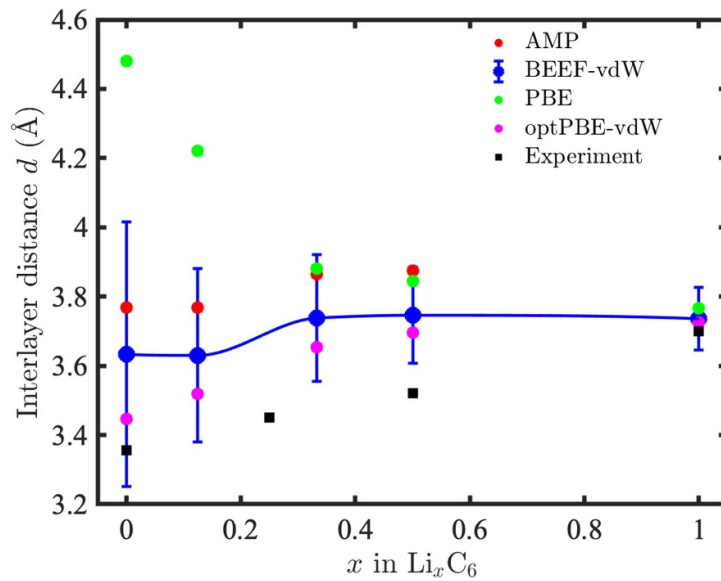
Figure 4. The graphite inter-layer separation as a function of lithium fraction, x in Li_xC_6. Experimental values are from Dahn *et al* [64]. Error bars correspond to one standard deviation of the ensemble predictions.

representative of the combined effects of various thermodynamic variables on lithium intercalation in graphite. The AMP potential is not trained to predict forces, so structural relaxations are not feasible as part of the GCMC routine. However, the effects of not including these relaxations are negligible [66].[4]

Following completion of the GCMC simulation, we mapped the thermodynamic vacancy orderings using convex hull analysis [23]. Extra lithium atoms were inserted by running the simulation at -3 V overpotential to capture the super-saturation regime. In addition to the internal energy from the AMP calculator, we accounted for the vibrational free energy and configurational entropy of each structure on the hull to get the Gibbs free energy, as Reynier *et al* [6] previously identified these effects as significant for dilute lithium phases.

Figure 5 shows the convex hull resulting from the GCMC simulation. For comparison, the convex hull with and without entropic effects is shown in comparison with the result from Pande and Viswanathan [9]. While the total allowed Li vacancies are half the number of carbon atoms in a graphite block, the GCMC simulation saturated at nearly 1 Li atom per six carbon atoms. This result matches the experimentally-observed full loading at (Li_1C_6) [4], which occurs due to strong interactions between adjacent Li atoms [9]. We term the phases where $x > 1$ as 'super-saturated' since an extra voltage is required push lithium ions in the supercell and form these phases. The effects of super-saturation are evident from the sharp increase in formation energies after $x = 1$ which imply a negative OCV and unstable interactions. There is also a steep decrease in the energies indicating increasing stability in the initial phases. The convex

---

[4]To evaluate the effect of structural relaxations on the energies, we used DFT to relax the most stable structure for every value of x output by the GCMC simulation to a maximum force of 0.1 eV Å^{-1}. Figure S3 shows the DFT-calculated maximum force on any atom for each structure before relaxation. The maximum force on any atom over the range $0 \leq x \leq 1$ is less than 0.7 eV Å^{-1} over the entire range of structures. Figure S4 shows the difference in energies of each structure before and after the relaxation, with the largest difference being just 0.15%.
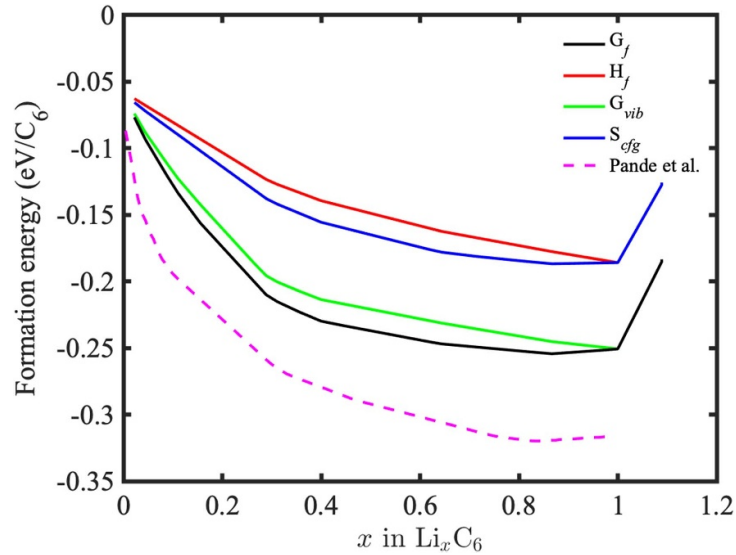
Figure 5. The convex hull of stable structures calculated using the machine learning potential with vibrational and entropic effects included (G_f). Additional curves show the effect of including only vibrational free energies (G_{vib}), only configurational entropy (S_{cfg}), or neither (H_f) in the determination of the hull. At $x = 1$, the slope of each curve becomes positive, signifying that intercalation is no longer energetically favourable. The configurational entropy reduces to zero at the endpoints ($x = 0, 1$), indicating fully ordered structures. Results from the cluster expansion simulation from Pande *et al* [9] are shown for comparison.

hull obtained from cluster expansion [9] shows an even steeper decrease in formation energies in the dilute phases that causes a near vertical drop in the corresponding OCV.

Straight lines in the convex hull signify the stable phase regions. The corresponding phase domain is identified by the compositions at the end points of the straight line. The OCV is calculated as the derivative of the total Gibbs free energies between the stable phase compositions. Transition from one composition to another during intercalation is described by the reaction

$$Li_{x_1}C_6 + (x_2 - x_1)(Li^+ + e^-) \leftrightharpoons Li_{x_2}C_6. \tag{15}$$
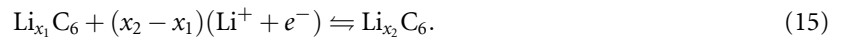
The average voltage can be obtained from the Gibbs free energies of the pristine ($x = 0$) and fully-intercalated ($x = 1$) structures as

$$V_{avg} = -\frac{G_{LiC_6} - G_{C_6} - G_{Li(s)}}{e}. \tag{16}$$

The open circuit voltage at a given composition ($x = x_1$) is simply related to the difference in energy of stable orderings at $x = x_1$ and $x = x_2$,

$$V = -\frac{G_{Li_{x_2}C_6} - G_{Li_{x_1}C_6} - (x_2 - x_1)G_{Li(s)}}{e(x_2 - x_1)}. \tag{17}$$

Here, $e = -1.602 \times 10^{-19}$ C is the charge of one electron.

Figure 6 shows the open circuit voltage calculated using the AMP potential as a function of composition, x in Li_xC_6, with and without entropic effects. We also show results from the cluster expansion analysis by Pande and Viswanathan [9] and experimental results from Dahn *et al* [66] and Reynier *et al* [6], which were obtained from equilibrated half cells cycled at a slow C/100 current rate. The AMP calculator shows good agreement with the cluster expansion result from Pande and Viswanathan [9], with differences of less than 0.1 V for phases with $x \geq 0.0833$. The OCV is much higher for $x < 0.0833$, which is likely due to their different treatment of dilute lithium phases. The cluster expansion model is low-dimensional, so it does not correctly describe long-range interactions, such as dispersion forces or lithium screening of long-range forces. The formation enthalpies of stage 3 and 4 compounds in the cluster expansion result were also approximated from stage 2 compounds, whereas we model structures explicitly and account for different possible stacking orders. Still, our voltages lie well within the BEEF-vdW ensemble deviations of the cluster expansion OCV shown in figure 6 of their paper [9]. The AMP OCV is also in good agreement with the experimental results for $0.3 \leq x \leq 1.0$, as the difference is less than 0.1 V over this span. For most phases with $x < 0.3$, the AMP result differs from either experimental result by 0.3–0.5 V. We identify kinetic effects and
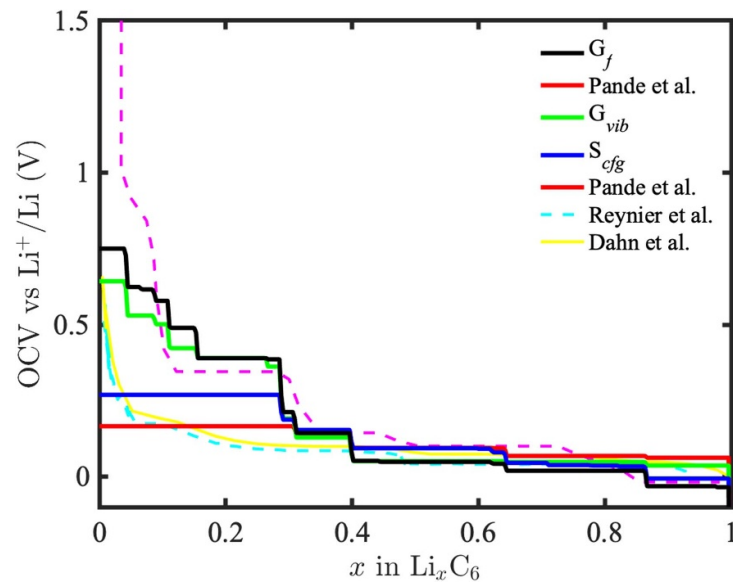
Figure 6. Open-circuit voltage (OCV) curves obtained from the machine learning model with vibrational and entropic effects included (G_f). Additional curves show the effect on the OCV of including only vibrational effects (G_{vib}), only configurational entropy (S_{cfg}), or neither (H_f) in the convex hull analysis. For comparison we show experimental results from Reynier *et al* [65] and Dahn *et al* [66] and the cluster expansion result from Pande and Viswanathan [9]. All voltages drop sharply at the dilute lithium limit showing fluctuating disordered phases and reach a uniform phase after $x = 0.4$ indicating mixed stage 1 compounds until full loading.

supercell size effects as the most likely factors causing this discrepancy. For example, Smith *et al* [67] showed that the disagreement in OCV between experiment and their analytical model in the $x < 0.3$ region can be corrected by treating the system as a solid solution, which allowed effects like concentration gradients, diffusivity, and other transport properties to be taken into account. As the GCMC model is a purely thermodynamic model, these kinetic effects cannot be modeled here. In addition, in the region $x < 0.03$, Dahn [4] observed a dilute 1 L solid-solution structure rather than a stage 4 or higher structure. This region coincides with the highest drop in the experimental OCV and a deviation of up to 0.5 V from the AMP OCV. The GCMC model, on the other hand, assumes uniform staging as lithium fraction increases due to a limited supercell size, and therefore cannot model the 1 L structure. The supercell size also limits the phases that can be observed for $x < 0.1$, as it limits the resolution of the lithium fraction x (in this case, the resolution is $1/45 \approx 0.022$). This is particularly an issue for probing very dilute lithium cases ($x < 0.03$), a regime where there is a significant decrease in experimental OCV with increasing x. Further complicating comparison with experiment for low x are kinetic and other experimental effects related to solid electrolyte interphase (SEI) formation in the half-cell measurements. The stacking order is variable in experiments [68], and magnitude of the current rate has been shown to affect the formation of the initial phases [69]. The formation of the solid-electrolyte interphase for low-lithium phases, $x < 0.03$, is known to affect the experimental OCV [6, 9].

The effects of including vibrational and configurational entropy in the Gibbs free energies can be seen in figures 5 and 6. Including both effects decrease the formation energies, which can be seen in the lowered hulls in figure 5. As a result, we observe higher voltages compared to experiment at dilute lithium phases ($x < 0.2$). For low x, the energetic contributions from the vibrational free energy is greater than contributions from entropy [9]. The configurational entropy increases at low x then decreases as x approaches 1. These contributions cause the OCV, S_{cfg}, to be higher at low x (compared to H_f) and lower at high x. Perfect ordering for pristine and fully-intercalated stages reduces the configurational entropy to zero, so the S_{cfg} and H_f convex hull curves intersect at $x = 0$ and $x = 1$.

The stable phases predicted by the AMP calculator are reported in table 4 along with the stage, stacking order, and inter-planar lattice constant for each phase. Stable stages were identified by the flat regions in the OCV curve. The results show that the inter-planar lattice constant increases monotonically with x, indicating lattice expansion of the graphite host to accommodate lithium. As indicated by Didier [5], mixed stages in the dilute limit have empty graphite layers characterized by AB and ABC stacking. The differential stacking order allows for different stages to occur at separate intervals of the intercalation period. Table 4 shows four types of phases and mixed 1 and 2 stage phases before and after $x = 0.1$.

Table 4. The phases, x for Li_xC_6, predicted by the AMP calculator including vibrational and entropic effects. The stage, stacking order, and in-plane lattice constant (a) for each phase are also reported.

x	Stage	Stacking	a (Å)
0.022	5	ABC	2.466
0.044	1,4	ABC	2.466
0.066	2,3	ABC	2.467
0.088	2,3	AB	2.468
0.111	1,2	AB	2.469
0.133	1,2	AA	2.469
0.156	1,2	AA	2.470
0.177	1,2	AA	2.471
0.291	1,2	AA	2.474
0.312	1	AA	2.475
0.402	1	AA	2.478
0.645	1	AA	2.485
0.864	1	AA	2.491
1.000	1	AA	2.496

4. Conclusion

In this work, we have presented an neural network potential using the Behlar–Parrinello atom-centered symmetry function featurization scheme to model the lithium-graphite system. The potential was trained on energies calculated using DFT with the BEEF-vdW exchange-correlation functional so that the potential could accurately model van der Waals interactions in large supercells. The potential was used to predict structural and elastic properties and tested against values from experiment and different DFT exchange-correlation functionals. The performance of the potential on unseen test data (an error of 8.24 meV atom^{-1}) and in predicting structural and elastic properties is comparable to the performance of similar machine-learned potentials on other systems [34]. This level of accuracy is also similar to other DFT exchange-correlation functionals at the GGA level and within the ensemble error estimation of BEEF-vdW.

We then predicted the convex hull of stable phases and the corresponding open circuit voltage using this calculator through a GCMC simulation. Using approximations for the configurational entropy and previously reported vibrational free energies from DFT-based phonon calculations, we were able to obtain a convex hull of stable phases and a good agreement with the experimental OCV curve within 0.1 V for $x > 0.3$. The AMP calculator performs better as compared to cluster expansion method in predicting the OCV along with structural properties. This study represents an another important proof point that machine learning potentials can accelerate the design and optimization of battery materials. Future developments will focus on incorporating force and stress calculation and predicting phonon properties.

The potential is available online for download via Github[5] as an AMP potential [47], which can be used as a calculator in the atomic simulation environment [70]. The training data is available for download as an ASE database. More details about accessing the potential and training data are presented in section S4.

ORCID iDs

Mohammad Babar ● https://orcid.org/0000-0001-6779-3859
Holden L Parks ● https://orcid.org/0000-0002-9530-0764
Gregory Houchins ● https://orcid.org/0000-0002-6858-9267
Venkatasubramanian Viswanathan ● https://orcid.org/0000-0003-1060-5495

[5]See https://github.com/BattModels/atomistic-potentials.

References

[1] W A van Schalkwijk and B Scrosati eds 2002 *Advances in Lithium-Ion Batteries* (New York: Springer)
[2] Asenbauer J, Eisenmann T, Kuenzel M, Kazzazi A, Chen Z and Bresser D 2020 *Sustain. Energy Fuels* **41** 5387-416
[3] Woo K C, Mertwoy H, Fischer J E, Kamitakahara W A and Robinson D S 1983 *Phys. Rev. B* **27** 7831–4
[4] Dahn J R 1991 *Phys. Rev. B* **44** 9170–7
[5] Didier C, Pang W K, Guo Z, Schmid S and Peterson V K 2020 *Chem. Mater.* **32** 2518–31
[6] Reynier Y, Yazami R and Fultz B 2003 *J. Power Sources* **119–121** 850–5
[7] Ceder G and der Ven A V 1999 *Electrochim. Acta* **45** 131–50
[8] Persson K, Hinuma Y, Meng Y S, der Ven A V and Ceder G 2010 *Phys. Rev. B* **82** 125416
[9] Pande V and Viswanathan V 2018 *Phys. Rev. Mater.* **2** 125401
[10] Lenchuk O, Adelhelm P and Mollenhauer D 2019 *J. Comput. Chem.* **40** 2400–12
[11] Thinius S, Islam M M, Heitjans P and Bredow T 2014 *J. Phys. Chem.* **118** 2273–80
[12] Hofmann U and Rüdorff W 1938 *Trans. Faraday Soc.* **34** 1017–21
[13] Kita Y, Watanabe N and Fujii Y 1979 *J. Am. Chem. Soc.* **101** 3832–41
[14] Imai Y and Watanabe A 2007 *J. Alloys Compd.* **439** 258–67
[15] Taminato S *et al* 2016 *Sci. Rep.* **6** 28843
[16] Zinth V, von Lüders C, Hofmann M, Hattendorff J, Buchberger I, Erhard S, Rebelo-Kornmeier J, Jossen A and Gilles R 2014 *J. Power Sources* **271** 152–9
[17] Zheng T and Dahn J R 1996 *Phys. Rev. B* **53** 3061–71
[18] Huang J X, Csányi G, Zhao J B, Cheng J and Deringer V L 2019 *J. Mater. Chem. A* **7** 19070–80
[19] Raju M, Ganesh P, Kent P R C and van Duin A C T 2015 *J. Chem. Theory Comput.* **11** 2156–66
[20] Sendek A D, Cubuk E D, Antoniuk E R, Cheon G, Cui Y and Reed E J 2018 *Chem. Mater.* **31** 342–52
[21] Ahmad Z, Xie T, Maheshwari C, Grossman J C and Viswanathan V 2018 *ACS Cent. Sci.* **4** 996–1006
[22] Joshi R P, Eickholt J, Li L, Fornari M, Barone V and Peralta J E 2019 *ACS Appl. Mater. Interfaces* **11** 18494–503
[23] Houchins G and Viswanathan V 2020 *J. Chem. Phys.* **153** 054124
[24] Fujikake S, Deringer V L, Lee T H, Krynski M, Elliott S R and Csányi G 2018 *J. Chem. Phys.* **148** 241714
[25] Nguyen T T, Székely E, Imbalzano G, Behler J, Csányi G, Ceriotti M, Götz A W and Paesani F 2018 *J. Chem. Phys.* **148** 241725
[26] Rupp M, Ramakrishnan R and von Lilienfeld O A 2015 *J. Phys. Chem. Lett.* **6** 3309–13
[27] Hansen K, Biegler F, Ramakrishnan R, Pronobis W, von Lilienfeld O A, Müller K R and Tkatchenko A 2015 *J. Phys. Chem. Lett.* **6** 2326–31
[28] Deringer V L and Csányi G 2017 *Phys. Rev. B* **95** 094203
[29] Rowe P, Csányi G, Alfè D and Michaelides A 2018 *Phys. Rev. B* **97** 054303
[30] Wen M and Tadmor E B 2019 *Phys. Rev. B* **100** 195419
[31] Bartók A P, Kermode J, Bernstein N and Csányi G 2018 *Phys. Rev.* **8** 041048
[32] Hajinazar S, Shao J and Kolmogorov A N 2017 *Phys. Rev. B* **95** 014114
[33] Artrith N, Urban A and Ceder G 2017 *Phys. Rev. B* **96** 014112
[34] Zuo Y *et al* 2020 *J. Phys. Chem. A* **124** 731–45
[35] Deringer V 2020 *J. Phys.: Energy* **2** 041003
[36] Wellendorff J, Lundgaard K T, Møgelhøj A, Petzold V, Landis D D, Nørskov J K, Bligaard T and Jacobsen K W 2012 *Phys. Rev. B* **85** 235149
[37] Kandasamy K, Vysyaraju K R, Neiswanger W, Paria B, Collins C R, Schneider J, Poczos B and Xing E P 2020 *J. Mach. Learn. Res.* **21** 1–27
[38] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1993 *Phys. Rev. B* **46** 6671–87
[39] Panosetti C, Annies S B, Grosu C, Seidlmayer S and Scheurer C 2019 Dftb modelling of lithium intercalated graphite with machine-learned repulsive potential (arXiv:1904.13351)
[40] Fletcher R 1981 *Practical Methods of Optimization* (Chichester, England, New York: J. Wiley)
[41] Blöchl P E 1994 *Phys. Rev. B* **50** 17953–79
[42] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758–75
[43] Enkovaara J *et al* 2010 *J. Phys.: Condens. Matter.* **22** 253202
[44] Mortensen J J, Hansen L B and Jacobsen K W 2005 *Phys. Rev. B* **71** 035109
[45] Behler J and Parrinello M 2007 *Phys. Rev. Lett.* **98** 146401
[46] Behler J 2011 *J. Chem. Phys.* **134** 074106
[47] Khorshidi A and Peterson A A 2016 *Comput. Phys. Commun.* **207** 310–24
[48] Auer P 2002 Using confidence bounds for exploitation-exploration trade-offs *J. Mach. Learn. Res.* **3** 397–422
[49] Srinivas N, Krause A, Kakade S M and Seeger M 2009 Gaussian process optimization in the bandit setting: no regret and experimental design (arXiv:0912.3995)
[50] Thompson W R 1933 *Biometrika* **25** 285–94
[51] Jones D R, Schonlau M and Welch W J 1998 *J. Glob. Optim.* **13** 455–92
[52] Qin C, Klabjan D and Russo D 2017 Improving the expected improvement algorithm (arXiv:1705.10033)
[53] Mounet N and Marzari N 2005 *Phys. Rev. B* **71** 205214
[54] Boettger J C 1997 *Phys. Rev. B* **55** 11202–11
[55] Lee J K *et al* 2016 *Sci. Rep.* **6** 39624
[56] Boutin A, Tavitian B and Fuchs A H *et al* 1997 *Faraday Discuss.* **106** 307–23
[57] Bondi A V 1964 *J. Phys. Chem.* **68** 441–51
[58] Christensen R, Hansen H A and Vegge T 2015 *Catal. Sci. Technol.* **5** 4946–9
[59] Ceder G, Aydinol M and Kohan A 1997 *Comput. Mater. Sci.* **8** 161–9
[60] Andresen J M 2002 *Energy Fuels* **16** 218–218
[61] Blakslee O L, Proctor D G, Seldin E J, Spence G B and Weng T 1970 *J. Phys. D: Appl. Phys.* **41** 3373–82
[62] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865–8
[63] Klimeš J, Bowler D R and Michaelides A 2009 *J. Phys.: Condens. Matter.* **22** 022201
[64] Dahn J R, Fong R and Spoon M J 1990 *Phys. Rev. B* **42** 6424–32

[65] Reynier Y F, Yazami R and Fultz B 2004 *J. Electrochem. Soc.* **151** A422
[66] Dahn J R, Zheng T, Liu Y and Xue J S 1995 *Science* **270** 590–3
[67] Smith R B, Khoo E and Bazant M Z 2017 *J. Phys. Chem.* **121** 12505–23
[68] Kirczenow G 1985 *Phys. Rev. Lett.* **55** 2810–13
[69] Safran S 1987 Stage ordering in intercalation compounds *Solid State Physics* (Amsterdam: Elsevier) pp 183–246
[70] Larsen A H *et al* 2017 *J. Phys.: Condens. Matter.* **29** 273002