Implanted neural network potentials: Application to Li-Si alloys

Berk Onat, Ekin D. Cubuk, Brad D. Malone, and Efthimios Kaxiras

Phys. Rev. B **97**, 094106 — Published 20 March 2018

DOI: 10.1103/PhysRevB.97.094106

Implanted Neural Network Potentials: application to Li-Si Alloys

Berk Onat, Ekin D. Cubuk, Brad D. Malone, Efthimios Kaxiras

Department of Physics and School of Engineering and Applied Sciences,

Harvard University, Cambridge MA 02138

(Dated: March 6, 2018)

Abstract

Modeling the behavior of materials composed of elements with different bonding and electronic structure character for large spatial and temporal scales and over a large compositional range, is a challenging problem. A case in point are amorphous alloys of Si, a prototypical covalent material, and Li, a prototypical metal, which are being considered as anodes for high-energy-density batteries. To address this challenge, we develop a methodology based on neural networks, that extends the conventional training approach to incorporate pre-trained parts that capture the character of different components, into the overall network; we refer to this model as the "implanted neural network" method. We show that this approach works well for the Si-Li amorphous alloys for a wide range of compositions, giving good results for key quantities like the diffusion coefficients. The method is readily generalizable to more complicated situations that involve two or more different elements.

I. INTRODUCTION

The modeling of materials with atomistic detail is subject to severe limitations: one can model either systems that are very small (in spatial and temporal scales) with reliable *ab initio* quantum mechanical methods, or simulate large systems for long time-scales but with increasing loss of accuracy as the size increases. The latter is the result of using approximate classical interatomic potentials to capture the quantum mechanical interactions between electrons and ions that are responsible for bonding of atoms in a solid. The challenge in developing reasonable potentials is exacerbated in situations where the nature of bonding is complex and defies classification in the usual categories like covalent, metallic, ionic, etc. Often, materials that are of great practical interest exhibit this kind of complex bonding arrangement between the constituent atoms, that can even vary as the material is cycled through its phases during the application.

A case in point is lithium alloys that show promise for high-energy density batteries, which are highly desirable in portable electronics and transportation[1,2]. Among the candidate electrode materials for high-energy density batteries, Si has received extensive attention as a promising anode material because of its high capacity for lithium absorption (4200 mAh g$^-$1), which compares very favorably to the common anode material graphite (372 mAh g$^-$1)[3]. However, upon Li insertion, a Si electrode can expand up to 400% in volume and the anode fractures. This greatly reduces the number of recharging cycles before the battery fails[4,5]. Understanding the nature and the details of these structural transformations is crucial to designing a reliable electrode. It also poses a great challenge to modeling, with the obvious need for capturing length scales from the atomistic structure and time scales related to atomic diffusion to those related to large deformation and fracture. The complex nature of the Si-Li interaction in Li_xSi, with $x = 0 - 4.25$, which evolves from pure covalent at $x = 0$ to pure metallic at $x > 2.5$, makes this challenge even more daunting. This is confirmed by the recent work of Xia and Carter[6], who found that methods like kinetic-energy-functional approaches show limited ability to predict alloy formation energies for the amorphous Si-Li system, which indicates the limitations in transferability of these functionals for such complex physical systems.

In this work we present a new method for meeting these challenges using neural networks to represent the atomic interactions in complex environments. The use of neural

networks to model materials was pioneered by Behler and Parrinello[7]. We found that all the variants of an atom-based neural network (ANN) that we were able to come up with could not adequately cope with the complexities of a system like the Li-Si alloy; based on this observation, we introduced sub-networks for datasets that are designed for species of atoms in their pure- and multi-element environments (bulk Si, bulk Li and the interaction between these two species, respectively), as well as sub-networks for specific phases including those of amorphous Si and Li-Si alloys. We then trained these networks in hierarchical building stages, that is, successively more complicated situations; consequently, the combined network is trained in a wide range of structures and can capture the physics of the system over the entire compositional range of interest. Recently, a similar approach consisting of hierarchical construction of NNs, based on different elements and their combinations in a material, has been introduced, under the name of "stratified neural network" (SNN) potentials[8]. In contrast to the SNN approach, which uses *element-oriented* NNs and their combinations in the hierarchical building of the final NN potential, in our method we employ *dataset-oriented* NNs that are used to transfer the pre-trained (through data) NNs to any level of the hierarchical construction of the final NN potential. We call this new method the "implanted" neural network (INN) model, since the parts of the network that correspond to a specific dataset of interest are not re-trained but maintain their identity in the combined network, while connections to the rest of the network are allowed to adjust at later stages of training. The present INN method substantially reduces the prediction errors and, more importantly, helps avoid overfitting, two major sources of problematic behavior in NN potential construction.

We emphasize that our approach is qualitatively different from other types of approaches in which weight-sharing is used to take advantage of the inherent symmetries of the problem, such as translational symmetry in images, as implemented in convolutional neural networks[9,10]. Instead, the freezing of weights in the INN achieves adaptability of the network to a wide range of structures and compositions with little added complexity of the model, while the inherent symmetries (translational, rotational and permutational symmetries of the atomic structures) are captured by the Behler-Parrinello type of symmetry functions[7].

We describe both the construction of the INN model and its application to simulate the diffusion of Li and Si atoms in the alloy at various compositions, which is of central importance in understanding the mechanisms of lithiation/de-lithiation. We find that our

INN model leads to significant improvements both in reducing the errors when the network is applied to validation data, and in reproducing experimental results for the complicated process of diffusion in the amorphous Li-Si alloy.

II. MODEL CONSTRUCTION AND VALIDATION

To provide a direct comparison between the present model and potentials obtained with the conventional ANN method[11,12], we constructed two different Li-Si potentials one with the INN method and the other with the ANN method. Since the SNN method relies on the ANN approach for the construction of pure element NNs, it is subject to the same limitations as ANN at the level of pure-element network construction. Therefore, hereafter we compare our INN method with the ANN approach, which has limitations in predictions even for pure Si. In general, the development of the neural network (NN) potential models have two components: the database to be used for the training and validation of the potential, the procedure for building the architecture of the NN, and the execution of the training and validation. We describe those next, separately.

A. Database for training and validation

Our database consists of four subsets: (i) the Si crystal (c-Si), (ii) the Li crystal (c-Li), (iii) the combined amorphous Si (a-Si) and c-Si bulk structures, and (iv) the amorphous Li-Si (a-LiSi) at various concentrations. For the pure element data sets, we included the cubic-diamond (cd) and tetrahedrally-bonded amorphous structures of Si and the fcc, bcc and R9 crystal phases of Li. In addition, we included lattice distortions (contraction, expansion and shearing) for the cd Si and bcc Li structures. For the a-Li-Si dataset, we used the a-Li_xSi alloy structures that were reported in our earlier work[13], with $0 \leq x \leq 4.25$ in increments $\Delta x = 0.125$. These structures were generated using density functional theory (DFT) calculations to obtain the energy optimized geometry at zero temperature. We also included structures generated by molecular dynamics (MD) simulations for c-Si, a-Si, bcc Li and a-Li-Si, at temperatures ranging from 300 K to 800 K at 100 K intervals; these extend the range to configurations away from equilibrium. At each temperature the sampling is performed with NVT runs after the structures are thermalized in NPT simulations for 1 ps

to determine their zero-pressure configurations at the desired temperature. The geometry optimizations and MD simulations were performed using DFT calculations with the SIESTA code[14]. To reduce the number of similar structures in the database, we thinned the training data set by analyzing the cross-correlation of atomic positions between consecutive steps in MD simulations and relaxation calculations. Specifically, we selected one out of every 100 steps in the MD simulations and one out of every 1000 steps (including the last step) in the relaxations; this procedure left 9×10^3 structures in the database, from the original 48×10^3 structures.

B. Construction of NNs in ANN and INN Potentials

The total energy in NN potentials is given by $E = \sum_i E_i$, where, the sum is over the number of atoms in the system and the atomic energy E_i of atom i is defined for each type of atom. For the NN in our potential, we employ the successful model architecture that has been used in many previous studies[7,11,15–20]. The NN architecture represents the contribution of atomic energies with respect to input descriptors extracted from the Cartesian coordinates of the neighboring atoms in the local chemical environment of each atom. To extract the respective structural characteristics for the local environment of each atom, we used radial and angular symmetry functions[11] that are given by

$$G_i^1 = \sum_{j \neq i} e^{\eta R_{ij}^2} f_c\left(R_{ij}\right) \tag{1}$$

$$G_i^2 = \sum_{j,k \neq i} 2^{1-\xi} \left(1 + \lambda \cos\theta_{ijk}\right)^{\xi} e^{\zeta\left(R_{ij}^2 + R_{ik}^2 + R_{jk}^2\right)} \times f_c\left(R_{ij}\right) f_c\left(R_{ik}\right) f_c\left(R_{jk}\right) \tag{2}$$

with the sumations running over the neighbors j and k of atom i, $\theta_{ijk} = \cos^{-1}\left(\mathbf{R}_{ij}\mathbf{R}_{jk}/R_{ij}R_{jk}\right)$ and $R_{ij} = |\mathbf{R}_{ij}|$; η, ζ, λ, ξ are the parameters that define the symmetry functions, R_{ij} is the distance between atom i and atom j, and $f_c(R)$ is the cutoff function defined as

$$f_c(R) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{\pi R}{R_c}\right) + 1\right] & \text{for } R_{ij} \leq R_c \\ 0 & \text{for } R_{ij} > R_c, \end{cases} \tag{3}$$

with $R_c = 6.0$ Å the cutoff radius (see Supplemental Material[21] for the parameter values). For simplicity in notation, we define the following labels for the two-body (X–Y) and

three-body (X–Y–Z) descriptors to the input layer (input vector): D(Si)=(X–Si, X–Si–Si), D(Li)=(X–Li, X–Li–Li), and D(LiSi) for the rest of the contributions, for each X element in the system, (X=Si, Li). We preprocess the input in three steps: (i) We shift each descriptor to have zero mean; (ii) we then apply principal component analysis (PCA) to decorrelate the different descriptors; (iii) finally, we scale each descriptor to have a standard deviation of one. Explicitly, the input to INN, the set of symmetry functions $G_{k,j}$, is standardized and transformed using PCA to the set $P_{k,j}$:

$$P_{k,j} = \frac{G_{k,j}W^T - t_A}{t_S}. \tag{4}$$

where W is the covariant matrix of the linear transformation in PCA, t_A and t_S are the average and standard deviation of the elements in the transformed vectors of the symmetry functions.

Since our preprocessing stage had the extra step of PCA projection, a step that is not used in previous applications of this NN architecture, we present here the derivation of forces. The force $F_{i,\alpha}$ acting on atom i is given by the partial derivative of the total energy with respect to the Cartesian coordinates $R_{i,\alpha}$ of the atom, where $\alpha = x$, y, or z. In the INN potential, the total energy E is calculated as a sum over all atomic energies E_k. This gives the expression for the force $F_{i,\alpha}$:

$$F_{i,\alpha} = -\frac{\partial E}{\partial R_{i,\alpha}} = -\sum_{k=1}^{N} \frac{\partial E_k}{\partial R_{i,\alpha}} \tag{5}$$

$$= -\sum_{k=1}^{N}\sum_{j=1}^{M_k} \frac{\partial E_k}{\partial G_{k,j}} \frac{\partial G_{k,j}}{\partial R_{i,\alpha}} \tag{6}$$

$$= -\sum_{k=1}^{N}\sum_{j=1}^{M_k} \frac{\partial E_k}{\partial P_{k,j}} \frac{\partial P_{k,j}}{\partial G_{k,j}} \frac{\partial G_{k,j}}{\partial R_{i,\alpha}}. \tag{7}$$

an expression in which all derivatives can be calculated analytically. In particular, the derivatives $\partial E_k/\partial P_{k,j}$ are calculated using the recurrence relation of gradients in the backward propagation algorithm[22] without propagating loss function and only using the gradient update routines as it is implemented in Torch7[23].

For the construction of the INN potential that represents the Si atoms, we used the following method, which is illustrated schematically in Fig. 1.

(a) In the first stage, a subset of nodes N_1 is trained using the subsets of the training database that consists of only c-Si. Since there is no contribution from Li interactions to

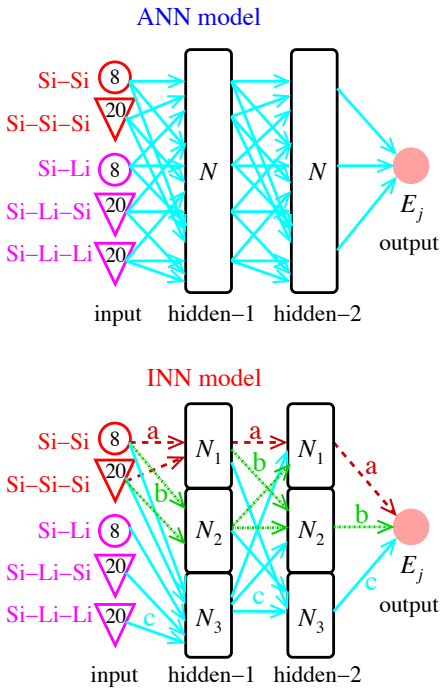

FIG. 1. Schematic representations of the construction of the ANN and INN models for Si: the descriptors consist of radial Si-X terms (circles) and angular Si–X–Y terms (inverted triangles), where X, Y = Si or Li; the ANN model contains N nodes in each of the two hidden layers all trained simultaneously, whereas the INN model contains N_1 nodes for c-Si, N_2 nodes for a-Si, and N_3 nodes for the Si-Li interactions, trained in three separate stages, labeled a (red arrows), b (green arrows) , c (cyan arrows).

descriptors in the c-Si training database, the descriptors D(Li) and D(LiSi) contain zero values, which lead to fixed values in the preconditioned input vector. This allows us to keep the D(Li) and D(LiSi) fixed while the weights of the Si NN are optimized. This stage makes it feasible to build a specialized Si sub-network that can capture the structure of the crystal phase.

(b) In the second stage, the weights of c-Si NN with N_1 nodes are implanted into a NN with additional N_2 nodes and the training is carried out for the combined c-Si and a-Si training database. In this stage, all the weights between input, N_1 nodes and output, as well as the D(Li) and D(LiSi) inputs are fixed while only the weights from/to the N_2 additional nodes are optimized. This ensures that the implanted network for Si-Si interactions generates the correct energy output for the c-Si subset, as well as that for the a-Si subset of structures.

(c) In the third and final stage, the Si-Si and Si-Li sub-networks are combined into a single

NN with N_3 additional nodes to capture the interactions between Li and Si. The same procedure as in the second stage is applied in the last stage, to prevent any change for the weights of the implanted Si-Si nodes ($N_1 + N_2$), and only the weights from/to the additional N_3 nodes are optimized.

In a similar fashion, we constructed the network that represents the Li atoms. Specifically, the pure Li sub-network is trained using the c-Li data set with D(Si) and D(LiSi) fixed during the optimization in stage (a) and is combined with the Li-Si sub-network in stage (c), with only the N_3 additional nodes added to the initial N_1 nodes. Since the pure Li part of the training database has only crystal phases, the Li sub-network is constructed only in two, rather than three stages.

We also considered whether or not there is a need to divide the a-Li-Si dataset into subsets, and then train the corresponding NNs for implanting them into the final Li-Si INN potential. To determine this, we generated three subsets of different Li concentrations, low ($x < 0.75$), medium ($0.75 < x < 1.5$) and high ($x > 1.5$). Upon training of these intermediate NNs using Si and Li implants, we found that the number of additional nodes needed for the final training step of the INN does not depend on the Li concentration because it is always comparable to N_3 of the third stage.

C. Training and validation of ANN and INN potentials

To generate the training database, we selected 85% of the structures and their DFT energies randomly from the thinned database. The rest of the data are reserved as the validation set for choosing hyper-parameters like the number of hidden layers and nodes. We additionally create an independent validation set from MD simulations of a-Si at 1200 K. For a fair comparison, we applied PCA transformations, Eq. (4), and used the same thinned database to train both ANN and INN models. The main differences between the two models is the construction of the INN potential in three stages, as opposed to the construction of the ANN potential which is done in the conventional way, in a one-shot fitting process of the weights, as illustrated in Fig. 1. We found that for the ANN potential, two hidden layers and $N = 30$ nodes in each layer gives the best results, that is, the lowest root-mean-square error (RMSE) for the validation set.

All NNs in this study are trained using the Torch7 NN package[23] and tests of the Li-Si

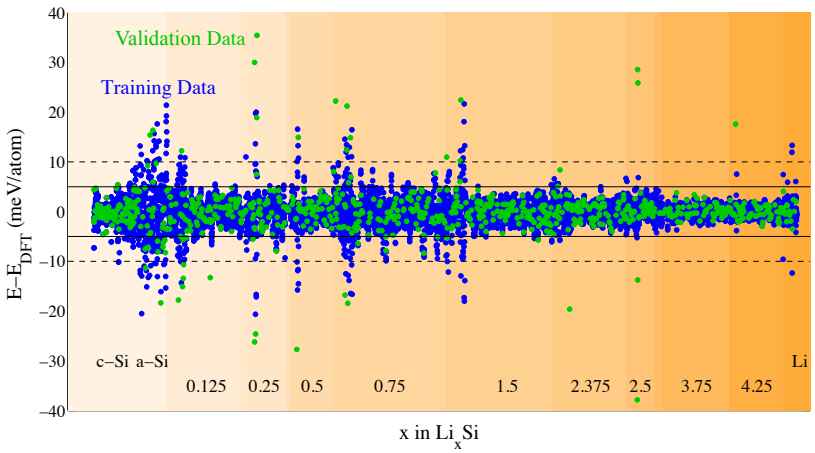

FIG. 2. The errors between the energy predictions of INN and the results of DFT calculations for Li, Si and its amorphous alloys. The colored areas show the Li concentration change where the lightest color does not include any Li and darkest color indicate pure Li structures. While blue dots show the the error at structures in the training database, green dots show validation data errors.

potentials are carried out using LAMMPS program[24] through our own implementation of the potential within the Knowledgeable Interatomic Models (KIM) API[25]. Similar to the ANN model, the choice of two hidden layers gives the best results. We have tested several choices of the number of nodes for each sub-network; we found that the best INN Li-Si potential is obtained with $(N_1, N_2, N_3) = (14, 12, 6)$ for the Si INN and $(N_1, N_3) = (10, 6)$ for the Li INN. Comparing the INN and ANN methods, we find that the the total number of nodes per hidden layer is *lower* for INN (48 nodes, 32 for Si and 16 for Li) than for ANN (60 nodes, 30 for Si and 30 for Li). In this sense, the ANN may be viewed as a special case of INN, in which none of the pre-trained NNs are implanted into the NN architecture. For the present model, this would correspond to $N_3 = 30$ and $N_1 = N_2 = 0$, see Fig 1.

The difference between INN predictions and DFT results for the energy per atom of the various structures and compositions used in the training set are presented in Fig. 2: 94% of the energy values predicted by the INN are within 5 meV/atom of the DFT results, where more than 83% of the structures are amorphous in the thinned database. This level of accuracy matches or exceeds the performance of other binary ANN potentials developed for amorphous structures[26]. As far as the ANN accuracy is concerned, the same level is reached for the training set, that is, 94% of the energy values are within 5 meV/atom of DFT results,

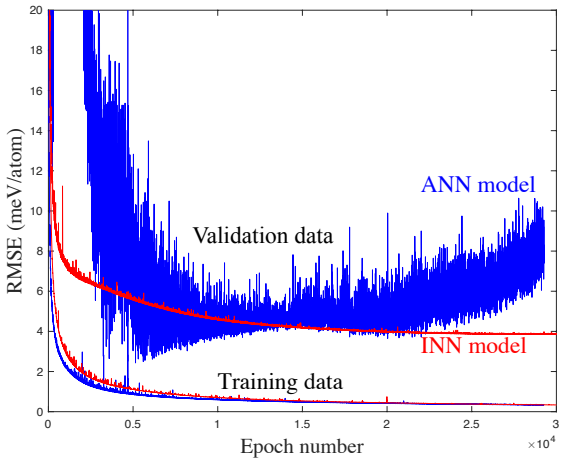

FIG. 3. Root-mean-square errors (in eV/atom) for the total energies of Li-Si structures, from the ANN model (blue curves) and the INN model (red curves), for the training and validation data.

whereas for the validation dataset the ANN predictions have significantly higher error (see Supplemental Material[21] for ANN predictions.)

A crucial consideration in producing NN-based potentials is to avoid overfitting. To explore this issue, we show in Fig. 3 the root-mean-squared error (RMSE) of both models for the training and validation data sets, as a function of the epoch number. While the RMSE for the two models is quite similar for the training data, the behavior for the validation data is strikingly different. A common method for identifying overfitting is to monitor the error of the validation set; non-monotonic behavior of this error signals overfitting. This is indeed the case for the ANN potential, with the error in the validation data decreasing up to a certain epoch number but then increasing. By contrast, the error in the validation data is a monotonically decreasing function of the epoch number for the INN potential. The oscillations in the RMSE of the validation dataset for the ANN model are considerably larger than for the INN model, and can be reduced with a smaller learning rate (see Supplemental Material[21]). However, we found that every attempt at training the ANN potential suffers from overfitting when our thinned database is used. One can use the so called "early stopping" method[15,16], in which the training stops when the error in the validation data starts to increasing, but there is no guarantee that this trick will eliminate overfitting. Indeed, the much larger fluctuations in the error associated with the validation data is a sign that the ANN potential has inherent limitations and cannot avoid overfitting. While ANN models have been reported to work well for many other materials using standard optimization

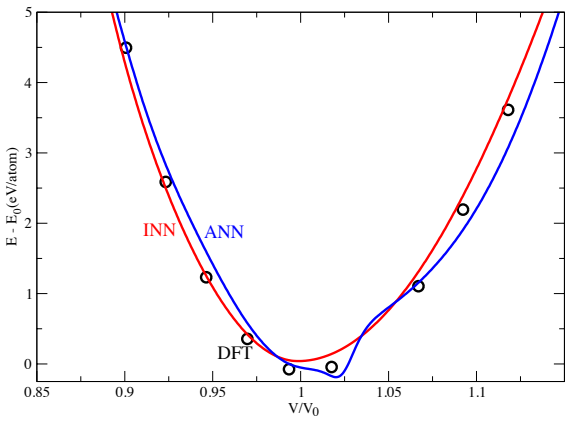

FIG. 4. Comparison of the results from ANN (blue) and INN (red) calculations for the total energy of crystalline bulk Si, in eV/atom relative to the equilibrium energy E_0, as a function of volume, normalized by the equilibrium volume V_0; the DFT energies are shown as black circles.

approaches, the results for the Li-Si system show that this method leads to overfitting since the diversity of the structures, in geometric configurations and concentration of different elements, is significantly increased. We cannot rule out the existence of hyperparameters that would improve the performance of the ANN method; instead, we found that the INN approached developed here is straightforward and leads to significantly less overfitting for the set of hyperparameters we have tried. In fact, recent work in deep learning methods indicates that a model's superiority may fundamentally arise from how easy it is to train, that is, how easy it is to find a set of hyperparameters that work well[27].

The oscillations of the error in the validation data is much smaller for the INN, and comparable to those in error of the training data. In addition, the average error for the validation data is smaller for the INN potential than that for the ANN potential for the entire range of epochs. Finally, we note that the RMSE on the validation set is 4.5 meV/atom and 6.9 meV/atom for the INN and ANN potentials, respectively. The maximum error on the validation set is 18 meV/atom for the INN potential, and 33 meV/atom for the ANN potential.

A signature of overfitting is that potential shows unphysical variations in important quantities. Indeed, as shown in Fig. 4, the total energy of the c-Si phase as a function of volume has unphysical variation near the minimum for the ANN potential; this is clear indication of overfitting for the ANN potential. In contrast to this, the INN potential shows smooth behavior comparable to that of the DFT results. A measure of the similarity of the

TABLE I. Lattice constants a_0, bulk modulus B (calculated using Birch-Murnaghan equation of state[28]) and elastic constants, C_{11}, C_{12}, C_{44} of the bulk phases of Si (cubic diamond) and Li (bcc), as obtained with the ANN and the INN potentials, and comparison to the results of DFT calculations and experiment. The numbers in parentheses for B are obtained from finite differences near the minimum of the total-energy curve.

	Si				Li			
	ANN	INN	DFT	Exp.	ANN	INN	DFT	Exp.
a_0 (Å)	5.55	5.51	5.51	5.43	3.60	3.65	3.65	3.49
B (GPa)	81.4	82.3	82.7	97.6	4.5	11.4	10.8	11.0
	(169.5)	(72.4)	(86.3)		(30.6)	(11.4)	(14.4)	
C_{11} (GPa)	166.6	83.3	149.0	166.0	40.5	12.5	16.7	13.7
C_{12} (GPa)	170.9	67.0	54.9	64.0	25.6	10.8	13.2	11.5
C_{44} (GPa)	37.4	76.2	99.2	79.0	11.9	6.0	7.3	6.9

INN and DFT results is the value of the bulk modulus B, which is the second derivative of the energy with respect to the volume, see Table I). For a fair comparison, we used the ANN potential determined after 7650 steps of training ("early stopping"), at which point the error in the validation data is smallest. For the INN potential, we used the form obtained after 30000 training steps, since the decrease in the error for the validation data is monotonic. We expect that if we had not used the early stopping method for the ANN potential the comparison of the total energy as a function of volume would likely be even more unfavorable than what is shown in Fig. 4.

As another measure of the high accuracy of the INN potential, we report in Table I the predictions for the lattice constants and bulk modulus and elastic constants for the c-Si and c-Li phases, and the corresponding DFT and experimental values. To calculate the bulk modulus, we used the Birch-Murnaghan equation of state (EOS)[28]. For the elastic constants, we apply small displacements to the crystal structures as implemented in LAMMPS and use the Cauchy stress tensor with the applied strain to determine C_{11}, C_{12} and C_{44}. For the DFT calculations, we use the same methods with the SIESTA code. We also present the bulk modulus that is calculated using the finite difference method for comparison in Table I. The difference in lattice constants between INN predictions and DFT calculations is less than 0.1%. Our results for the bulk modulus and elastic constants are overall in good agreement

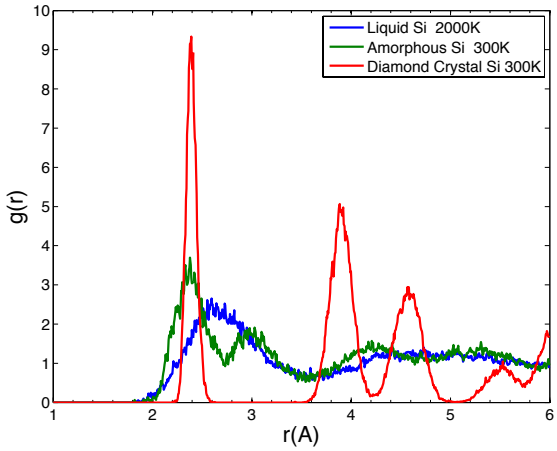

FIG. 5. Features of the crystalline and amorphous bulk phases of Si at 300 K, and of the liquid phase at 2000 K, as obtained from the INN potential.

with the DFT and experimental results for both elements[29,30]. The only exception is the value of C_{11} of Si, for which the INN potential predicts a lower value than the DFT and experimental values. The values obtained for the bulk modulus and elastic constants with the ANN potential are significantly different from both the DFT and experimental values. The main reason for this is that the training in the ANN method can lead to overfitting as discussed above, which produces unrealistic results when calculating derivatives of the energy, as is involved in the calculation of the bulk modulus and elastic constants.

III. APPLICATIONS

We next employ the INN potential to explore the behavior of various phases and processes of the system of interest. Specifically, we consider the structure of the crystalline and amorphous phases of Si at finite temperature (300 K), the structure of the liquid at 2000 K, the formation energy of various Li-Si structures, and the diffusion coefficient for Si and Li atoms in the lithiated amorphous structures. We emphasize that the structure and properties of these phases and processes are not part of the training data set, so in this sense these can be considered predictions of the INN potential that can be compared to either first-principles calculations or to experimental results, as appropriate.

A common check of effective potentials is how well they reproduce the properties of the system at finite temperatures, especially the structures of the liquid and amorphous phases

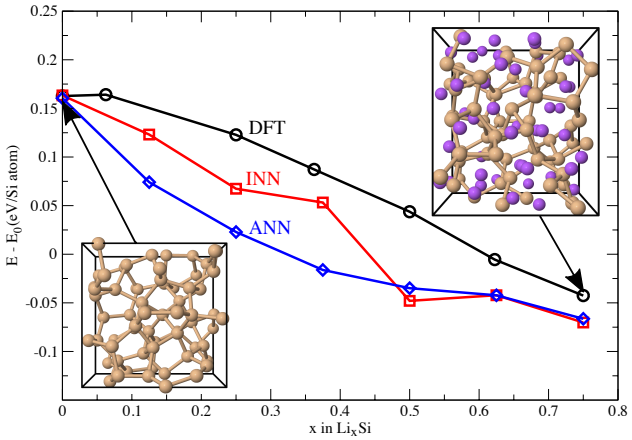

FIG. 6. Formation energy of a-Li$_x$Si structures for $0 \leq x \leq 0.75$, as obtained from DFT calculations and from the ANN and INN models. Each point in the ANN and INN calculations is the average of 5 different configurations at a given value of x. The two insets, at $x = 0$ and $x = 0.75$ show representative atomic structures (yellow spheres: Si atoms, purple spheres: Li atoms).

after quenching. For a comprehensive analysis of the predictions of the Li-Si INN potential, we conduct a series of MD simulations. The liquid phase at 2000 K, the crystalline and amorphous phases at 300 K are obtained with the following procedure: First, a supercell with 1728 Si atoms in the cd phase is generated and relaxed to its zero pressure configuration, which is heated to 2000 K in 30 ps using the NPT ensemble. The amorphous phase of Si is then obtained by quenching the system down to room temperature in 40 ps. For all phases of Si, the radial distribution function (RDF) is calculated by taking an average over an additional 3 ps of simulation using the NVT ensemble. We find that the INN potential, in addition to describing well the solid crystal and amorphous phases of Si, gives a good description of the liquid phase and the procedure of quenching. In Fig. 5 we show the calculated RDF of the liquid phase and compare it to those of the cd and a-Si phases. The liquid and amorphous RDF curves are in good agreement with those reported in experiments[31] and DFT calculations[13]. As mentioned, the training database does not include any information on the liquid phase of Si nor any information for the structure at temperature above 800 K.

A simple application of the Li-Si INN model is the formation energy of lithiated a-Si structures. We calculated the formation energy of structures with Li content up to a-Li$_{0.75}$Si. As was done for the calculation of the formation energy within the DFT approach[13], we add

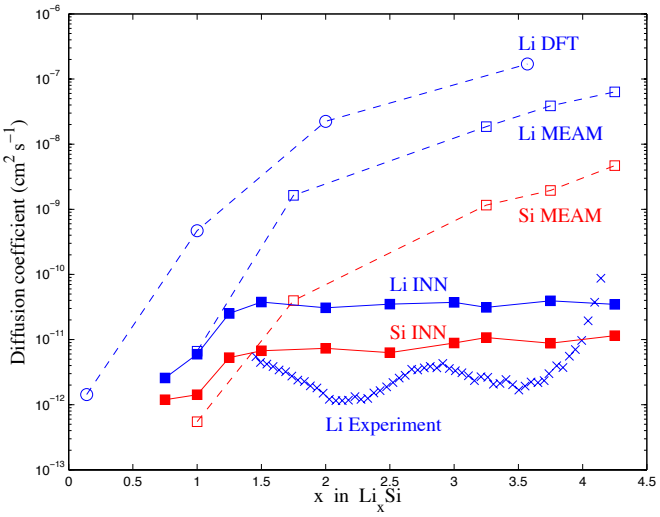

FIG. 7. Diffusion coefficients of Li and Si elements in a-Li-Si alloys. The results of present INN Li-Si potential is compared with the results of MEAM in Ref.[34], DFT calculations in Ref.[35] and the experimental coefficients taken from Ref.[36].

Li atoms into pure a-Si structure that is obtained from DFT calculations. For the sake of direct comparison between DFT results and ANN/INN predictions, we use a periodic unit cell containing 64 Si atoms initially, as was the case for the DFT calculations. Once the total energies of the structures are obtained, the formation energies of a-Li$_x$Si per Si atom is calculated in the usual manner[13]. The procedure is repeated up to 5 times to obtain reasonable average formation energies. We compare our results with the energies determined from *ab-initio* calculations and ANN potential in Fig. 6. While the results from both potentials give lower formation energies than the DFT calculations, the values from the INN potential are closer to the DFT results than those of the ANN potential. The slope of the formation energy (E_f) curve *vs.* the Li content x is an important quantity because it determines the voltage (V) during the lithiation cycle[13], which is given by $V(x) = -\partial E_f/\partial x$[32]. For a comparison between the measured voltages and our results, we take the averages of the slopes between x =0.2 and 0.75, for the DFT, INN and ANN curves; the values we find are 0.31, 0.32, and 0.20 V, respectively. While the DFT and INN potential results are reasonable estimates for the average voltage compared to the experimental value[32,33], 0.3 V, the ANN potential gives a substantially lower voltage which is out of the range of the experimental data in the selected region of x.

A more challenging application of the INN Li-Si potential is to address directly issues related to the performance of actual anodes such as the dynamics of the lithiation process. We carried-out MD simulations at room temperature to determine the diffusion coefficients of Si and Li atoms in the a-Li$_x$Si alloys. We use a timestep of 0.1 fs which ensures a stable total energy with a drift of less than 5% over timescales of 10 ns. Starting from the a-Li$_x$Si structures at various concentrations as obtained from DFT calculations[13], we scaled the periodic cells from the original size to a $2 \times 2 \times 2$ supercell, corresponding to 896 atoms for $x = 0.75$, and 2688 atoms for $x = 4.25$.

In the larger cells, we heat the samples to 300 K over a time of 1 ps, and then perform MD simulations for 2 ps to reduce the order due to repeating the original structures as periodic images. We analyzed the RDF of Si and Li atoms in the MD simulations and found that the INN potential maintains the basic character of the a-Li-Si phases without long-range order, and that Si bonds brake and re-form through the diffusion of Li atoms. We determine the mean-square-displacement (MSD) of atoms in the MD simulations for times up to 40ps. Fitting the late-time part of the MSD curves with a linear term in t, we obtain the diffusion coefficients for 10 different concentrations of a-Li-Si alloys (see Supplemental Material[21] for RDF and MSD curves for $x = 1.25, 2.5, 3.75$ of a-Li$_x$Si). In Fig. 7, we show our results for the diffusion coefficients of Si and Li atoms in the amorphous alloys, and compare with results from DFT calculations[35], simulations with the modified embedded atom method (MEAM) potential[34], and experiment[36]. The predictions from our INN potential close to the experimental values[36–38], within about one order of magnitude. This is a substantial improvement over the other theoretical results, which differ from experiment by many orders of magnitude. While our diffusion coefficients of Li are still higher than the experimental values, the average of the diffusion coefficients is about 3.5×10^{-11} cm^2/s, a value that is in very good agreement with the average Li-ion diffusion coefficient, 2.6×10^{-11} cm^2/s measured using diffusion-induced NMR at 246 K[39].

As a first step towards modeling diffusion dynamics of a-Li-Si, we study the dependence of diffusion on the local atomistic structure for Li atoms using the INN potential. Previous work has shown that the diffusion probabilities (P_D) of individual atoms in disordered structures depend crucially on the density of their neighbors at different distances[40,41]. Larger density of neighbors at the first two peaks of the neighbor shells reduces P_D, whereas larger density of neighbors away from the peaks increases P_D[42,43]. In Fig. 8, we present P_D for a Li atom

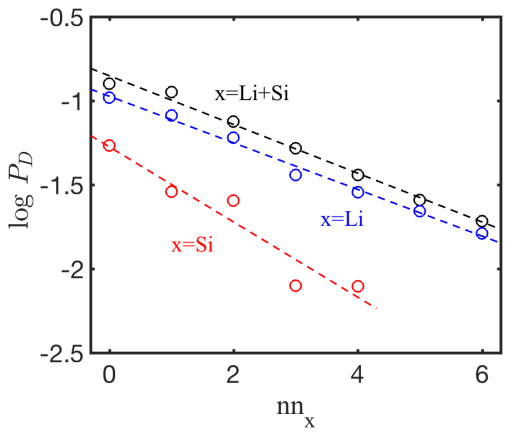

FIG. 8. Logarithm of the probability P_D of a Li atom to diffuse more than 1 Å in the next 200 fs as a function of its number of neighbors, nn_x, for x=Li (other Li neighbors), x=Si (Si neigbhors) and x=Li+Si (either kind of neighbors). The data was taken from an MD simulation of a-$Li_{3.75}Si$ at 300 K and the analysis was carried out over 261 diffusion events. The dashed lines are linear fits to $\log P_D$.

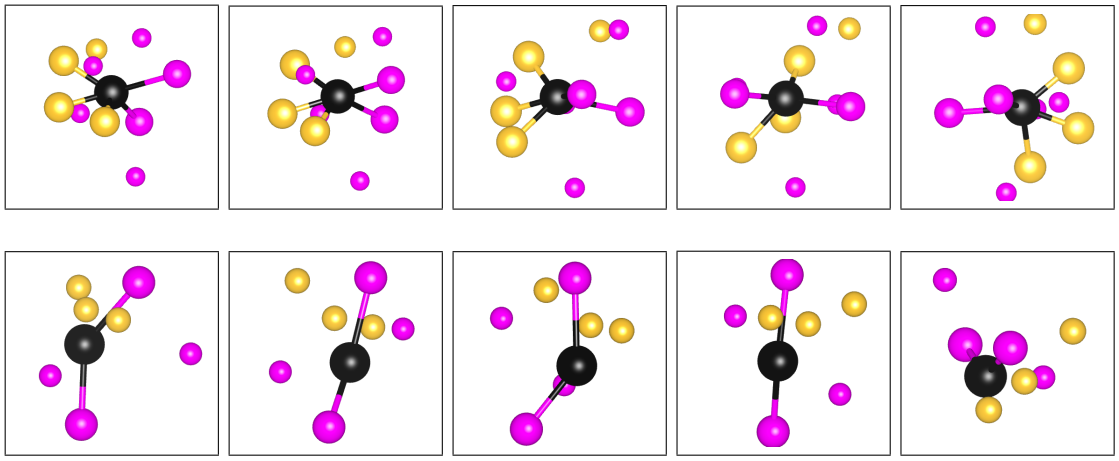

FIG. 9. Atomic structure of a hard (top row) and soft (bottom row) site, each one viewed from five different angles for better perspective. The atom at the center (black) is a Li atom, its nearest neighbor Si (yellow) and Li (purple) atoms are shown linked to it by bonds, and the next neighbors are shown unlinked.

in a-$Li_{3.75}Si$, as a function of nn_x, the number of nearest neighbors lithium (x=Li), silicon (x=Si), and total (x=Li+Si). A neighbor is defined as being within the range of d_{near} and d_{far} from the atom, where $d_{near} = 2.2$ Å for x=Li and 2.1 Å for x=Si, and $d_{far} = 3.1$ Å for x=Li and 2.8 Å for x=Si. As seen in Fig. 8, increasing the number of neighbors reduces P_D

exponentially. In Fig. 9 we show examples of neighborhoods in which the Li atom is not moving (referred to as a "hard spot") and is moving (a "soft spot"), according to the criteria specified in Fig. 8. The increased number of neighbors (both Li and Si) of the Li atom in the hard spot, leads to decreased diffusivity. Further study is needed to understand how the connection between the structure and dynamics lead to the lithiation and crystallization rates observed in experiments.

IV. CONCLUSION

In the present work we developed a new approach to creating effective potentials for systems containing elements that behave very differently in their respective bulk phases; the system we studied here consists of Si, a prototypical covalently bonded solid, and Li, a prototypical metal. The combined system is one of the interesting candidate materials for anodes in high-energy density batteries. The novelty of the approach consists of training parts of a neural network to represent the qualitatively different components of the system (the Si and Li atomic structures) before implanting these parts into the larger network that is trained to capture the energetics of the combined system; thus, we refer to the model as the implanted neural network potential. This method of training the network in stages avoids overfitting, which can hamper the performance of NN potentials. Our results show that the INN potential for the Li-Si system can capture both the pure element phases and amorphous mixed phases. We calculate the diffusion coefficients for Li and Si atoms by performing MD simulations with the INN potential, and find good agreement with available experimental measurements.

Our approach has two distinct advantages over conventional ways of fitting potentials. First, the use of a neural network makes the fitting of complex systems straightforward, without the limitations imposed by having to choose specific functional forms. Second, the idea of pre-training NNs for selective datasets and then implanting these parts of the network to represent the qualitatively different components provides a flexibility that is missing from both full-network and *element-oriented* stratified network trainings. This flexibility is essential in capturing the behavior of systems that are composed of more than one elements and in which each element behaves in a qualitatively different manner. It also helps avoid overfitting, because at each neuron implanting stage there are only few weights to be opti-

mized for the new dataset which greatly reduces the complexity of the optimization problem in NN's with large number of input and hidden nodes. For example, only the weights for N_3 nodes per hidden layer in the Li-Si NN need to be fitted to construct the full Li-Si potential at the final stage, and the complexity of this optimization problem is reduced from $N_3 = 30$ in the ANN model to $N_3 = 6$ in INN model. By reducing the complexity of the fitting problem, we significantly improved the trainability of the neural network potential with the INN approach; this includes the ease of finding good hyperparameters as well as trainability of the neural network itself. It is plausible that a different set of hyperparameters could improve the accuracy of either model.

The approach presented here is readily generalizable to systems with more than two elements, subject to the additional cost incurred by the training of subnetworks. The number of subnetworks that need to be trained is $(2^{n_{\mathrm{el}}} - 1)$, where n_{el} is the number of elements, assuming that each trained subnetwork can then be implanted in the full NN. For the INN Li-Si potential, we would then have to train $2^2 - 1 = 3$ subnetworks, for Li, Si and Si-Li. However, in some cases the subnetworks may need to be subdivided further, as in the case of Si in the present model, which was subdivided into the crystalline and amorphous components to accommodate these two distinct bulk phases with different characteristics; no subdivision was needed for the subnetwork of Li which has only crystalline bulk phases. While the larger number of subnetworks adds to the computational cost of training, the smaller size for each subnetwork in terms of number of nodes reduces this cost significantly, as explained above.

Finally, any NN-based potential can be continuously improved by adding more data to the training set and refitting the weights, without having to change the functional form. For instance, one way to significantly improve the accuracy of the forces produced by the model is to use the DFT calculated forces as part of the training set for a range of configurations; we are currently investigating this type of approach.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the U.S. Army Research Laboratory through the Collaborative Research Alliance (CRA) for Multiscale Multidisciplinary Modeling of Electronic Materials (MSME). B.O. was also supported by the Scientific and

Technological Research Council of Turkey-TUBITAK through project grant 2219. We used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by NSF Grants TG-DMR120073 and TGPHY120021 and the resources of FAS Research Computing at Harvard.

---

[1] M. Armand, J.M. Tarascon, Nature **451**, 652-657 (2008).

[2] M.S. Wittingham, MRS Bull. **33**, 411-419 (2008).

[3] A. Ostadhossein, E.D. Cubuk, G.A. Tritsaris, E. Kaxiras, S. Zhang, and A.C.T. van Duin, Phys. Chem. Chem. Phys. **17**, 3832-3840 (2015).

[4] M.T. McDowell, S.W. Lee, J.T. Harris, B.A. Korgel, C. Wang, W.D. Nix, Y. Cui, Nano Lett. **13**, 758-764 (2013).

[5] A. Ostadhossein, S.Y. Kim, E.D. Cubuk, Y. Qi, and A.C.T. van Duin, J. Phys. Chem. A **120**, 2114-2127 (2016).

[6] J. Xia, and E.A. Carter, Modelling Simul. Mater. Sci. Eng. **24**, 035014 (2016).

[7] J. Behler, and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[8] S. Hajinazar, J. Shao, and A.N. Kolmogorov, Phys. Rev. B **95**, 014114 (2017)

[9] K. Fukushima, Biological Cybernetics **36**, 193 (1980).

[10] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel, Neural Computation **1**, 541 (1989).

[11] N. Artrith, B. Hiller, and J. Behler, Phys. Status Solidi B **250**, 1191-1203 (2013).

[12] E.D. Cubuk, B.D. Malone, B. Onat, A. Waterland, E. Kaxiras, J. Chem. Phys. **147**, 024104, (2017).

[13] E.D. Cubuk, E. Kaxiras, Nano Lett. **14**, 4065-4070 (2014).

[14] J.M. Soler, E. Artacho, J.D. Gale, A. Garcia, J. Junquera, P. Ordejon, D. Sanchez-Portal, J. Phys.: Condens. Matter. **14**, 2745-2779 (2002).

[15] S. K. Natarajan, T. Morawietza, J. Behler, Phys. Chem. Chem. Phys. **17**, 8356-8371 (2015).

[16] J. Behler, International Journal of Quantum Chemistry **115**, 1032-1050 (2015).

[17] J. Behler, J. Chem. Phys., **134**, 074106 (2011).

[18] T. Morawietza, V. Sharma, J. Behler, J. Chem. Phys. **136**, 064103 (2012).

[19] N. Artrith, A.M. Kolpak, Nano Lett. **14**, 2670-2676 (2014).

[20] N. Artrith, A. Urban, Comput. Mater. Sci. **114**, 135-150 (2016).

[21] See Supplemental Material at [URL will be inserted by publisher] for the parameter values of INN potential, the predictions of ANN potential, RMSE of the datasets for ANN and INN, RDF and MSD curves for $x = 1.25, 2.5, 3.75$ of a-Li_xSi.

[22] M. T. Hagan and M. B. Menhaj, IEEE Trans. Neural Netw. **5**, No. 6, 989-993 (1994).

[23] R. Collobert, K. Kavukcuoglu, and C. Farabet, "Implementing Neural Networks Efficiently" in G. Montavon et al. (Eds.): Neural Networks: Tricks of the Trade, 2nd Edition, LNCS 7700, (Springer-Verlag, Berlin Heidelberg, 2012) Chapter 21, 537-557.

[24] S. Plimpton, J. Comp. Phys. **117**, 1 (1995).

[25] E. B. Tadmor, R. S. Elliott, J. P. Sethna, R. E. Miller, and C. A. Becker. Knowledgebase of Interatomic Models (KIM). https://openkim.org, (2011).

[26] G.C. Sosso, M. Salvalaglio, J. Behler, M. Bernasconi, and M. Parrinello, J. Phys. Chem. C **119**, 6428-6434 (2015).

[27] J. Collins, J. Sohl-Dickstein, and D. Sussillo, *Capacity and Trainability in Recurrent Neural Networks*, (ICLR 2017 Conference, 2017), https://openreview.net/pdf?id=BydARw9ex

[28] Murnaghan, F. D. (1944). "The Compressibility of Media under Extreme Pressures". Proceedings of the National Academy of Sciences of the United States of America 30 (9): 244-247. Birch, Francis (1947). "Finite Elastic Strain of Cubic Crystals". Physical Review 71 (11): 809-824.

[29] S.P. Nikanorov, Yu. A. Burenkov, and A. V. Stepanov, Sov. Phys. Solid State **13**, 10 2516-2519 (1971).

[30] G.J. Vazquez, Revista Mexicana de Fisica **23**, 572-578 (1990).

[31] J. Fortner, J. S. Lannin, Phys. Rev. B **39**, 5527-5530 (1989).

[32] V. L. Chevrier and J. R. Dahn, J. Electrochem. Soc. **156** A454-A458 (2009).

[33] V.L. Chevrier, J.R. Dahn, J. Electrochem. Soc. **157**, A392-A398 (2010).

[34] Z. Cui, F. Gao, Z. Cui, J. Qu, J. Power Sources **207**, 150-159 (2012).

[35] C-Y. Chou, G.S. Hwang, J. Power Sources **263**, 252-258 (2014).

[36] N. Ding, J. Xu, Y.X. Yao, G. Wegner, X. Fang, C.H. Chen, I. Lieberwirth, Solid State Ionics **180**, 222-225 (2009).

[37] K. Yoshimura, J. Suzuki, K. Sekine, and T. Takamura, J. Power Sources **146**, 445-447 (2005).

[38] K. Yoshimura, J. Suzuki, K. Sekine, and T. Takamura, J. Power Sources **174**, 653-657 (2007). (1987).

[39] A. Dunst, M. Sternad, V. Epp, and M. Wilkening, J. Phys. Chem. C **119**, 12183-12192 (2015).

[40] E.D. Cubuk, S.S. Schoenholz, J.M. Rieser, B.D. Malone, J. Rottler, D.J. Durian, E. Kaxiras and A.J. Liu, Phys. Rev. Lett. **114**, 108001 (2015).

[41] S.S. Schoenholz, E.D. Cubuk, D.M. Sussman, E. Kaxiras, and A.J. Liu, Nat. Phys. **12**, 469 (2016).

[42] E.D. Cubuk, S.S. Schoenholz, E. Kaxiras, and A.J. Liu, J. Phys. Chem. B **120**, 6139 (2016).

[43] S.S. Schoenholz, E.D. Cubuk, E. Kaxiras, and A.J. Liu, PNAS **114**, 263 (2017).