CHEMISTRY OF MATERIALS

pubs.acs.org/cm

Article

Defect Diffusion Graph Neural Networks for Materials Discovery in High-Temperature Energy Applications

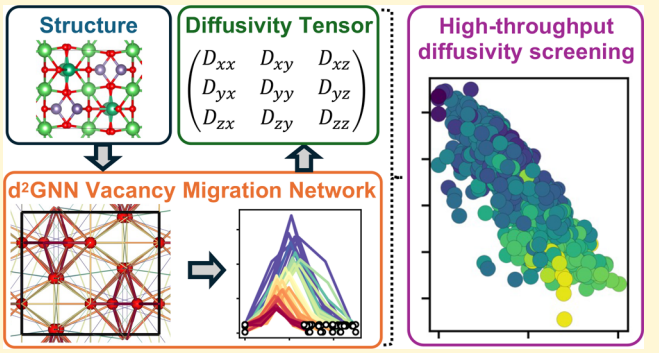
Lauren Way,[▽] Catalin D. Spataru,[▽] Reese E. Jones, Dallas R. Trinkle, Andrew J. E. Rowberg, Joel B. Varley, Robert B. Wexler, Christopher M. Smyth, Tyra C. Douglas, Sean R. Bishop, Elliot J. Fuller, Anthony H. McDaniel, Stephan Lany, and Matthew D. Witman*[,▽]

Cite This: https://doi.org/10.1021/acs.chemmater.5c00021

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The migration of crystallographic defects dictates material properties and performance for a plethora of technological applications. Density functional theory (DFT)-based nudged elastic band (NEB) calculations are a powerful computational technique for predicting defect migration activation energy barriers, yet they become prohibitively expensive for high-throughput screening of defect diffusivities. Without introducing hand-crafted (i.e., chemistry- or structure-specific) descriptors, we propose a generalized deep learning approach to train surrogate models for NEB energies of vacancy migration by hybridizing graph neural networks with transformer encoders and simply using pristine host structures as input. With sufficient training data, computationally efficient and simultaneous inference of vacancy defect thermodynamics and migration activation energies can be obtained to compute temperature-dependent vacancy diffusivities and to down-select candidates for more thorough DFT analysis or experiments. Thus, as we specifically demonstrate for potential water-splitting materials, candidates with desired defect thermodynamics, kinetics, and host stability properties can be more rapidly targeted from open-source databases of experimentally validated or hypothetical materials.



■ INTRODUCTION

Vacancies are the primary mediator for the diffusion of atoms in a large number of solid materials,[1,2] including crystalline oxides. The study of vacancy mobility and diffusion in bulk materials has myriad technological applications, including in devices that depend on low or high mobility. Both metal[3,4] and nonmetal[5,6] oxides are of interest. Passivation layers, thermal and electrical insulators, as well high conductivity thermoelectric, electronic, and electrochemical devices depend on stable atomic structures, which can undergo degradation due to vacancy migration when exposed to high temperatures and temperature gradients. A particularly consequential application is the geological confinement of hazardous waste[7,8] where diffusion can impact the long-term stability of the confinement environment. Vacancy migration can also play a role in other degradation processes, such as corrosion.[9,10] Diffusion is strongly affected by temperature, vacancy concentration, and other factors,[1,2] and significant diffusional anisotropy[8] can serve either as a challenge or a benefit in technological design.

To guide materials design and optimization for targeted vacancy defect diffusion properties, first-principles calculations of defect migration activation energies[11−14] based on nudged elastic band (NEB) calculations[15,16] are often utilized to provide critical atomistic insights. However, they become extremely expensive due to their reliance on density functional theory (DFT), particularly due to the need for large host supercell structural representations to avoid spurious interactions between the periodic defect images. Even more difficult is the prediction of, for example, temperature-dependent vacancy diffusivity tensors,[17] because one must compute the vacancy migration activation energy of all plausible symmetrically inequivalent paths, of which there can be many for compositionally or structurally complex materials. To address the large computational cost of these calculations, surrogate modeling approaches like cluster expansion (CE) have been employed to predict defect migration activation energies in a variety of contexts[18−20] for specific types of structures and chemistries. Meanwhile, other approaches to mitigate this difficulty have focused on a priori estimation of low energy paths to lower the number of first-principles calculations needed subsequently.[21]

Received: February 25, 2025
Revised: August 12, 2025
Accepted: August 14, 2025

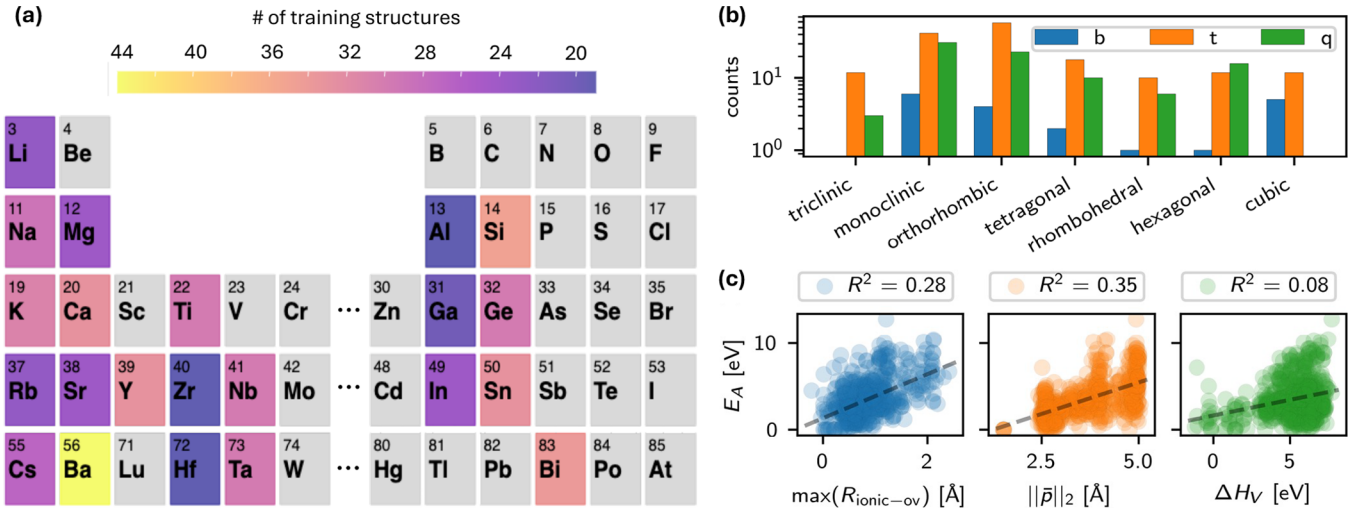


Figure 1. (a) Number of times an element appears in the 272 training data compounds with at least one converged NEB calculation. (b) Distribution of training data compounds among crystal systems and binary, ternary, or quaternary oxide chemical systems. (c) R^2 performance of some simple single variable linear regression models for E_A. Dashed black lines represent the line of best fit.

Substantial work has recently been devoted to the use of various machine learning surrogate models for vacancy or interstitial defect thermodynamic properties,[22−27] and defect migration activation energies have been well correlated with physical descriptors.[28−31] Here, we extend these efforts and derive *generalized* structure−property surrogate models of NEB-calculated defect migration (i.e., yielding both formation energies and activation energies) that can be applied across any structure type or chemistry space with an accuracy that is principally limited by quantity and diversity of NEB training data. Critically, the only input to the model is the relaxed *host structure* and a tuple of indices that represent the start and end atomic sites for the vacancy migration event; i.e., relaxed structures for each NEB image are not used. The basis for this model is a graph neural network architecture, whose convolution functional form could in principle be taken from one of the many variants recently proposed,[32−36] which is then hybridized with a transformer encoder[37] to enforce various physical, NEB-required symmetry constraints in the model's output.

To begin, we create a training database of neutral oxygen vacancy migration activation energies across a diverse structural and chemical space of metal oxides. Our trained model can then screen Materials Project (MP) structures[38] to predict the activation energies of all possible vacancy diffusion paths below a distance cutoff on the path length, and subsequently compute temperature-dependent diffusivity tensors from this data.[17] In addition to generally identifying materials with interesting diffusivity characteristics, utility in a specific materials discovery application is demonstrated via a multidimensional down-select of materials for thermochemical water-splitting (TCH)[39−42] based on predicted oxygen vacancy defect thermodynamics, diffusivity, and host stability. While this data set considers only neutral oxygen vacancies and does not delve into more complex defect-mediated diffusion pathways[31,43] (e.g., interstitial-facilitated or knockout diffusion), we nonetheless expect to identify promising materials that can be investigated on a case-by-case basis with more accurate methods. We therefore also conclude with some perspectives on how the proposed method can be improved with better accuracy and greater applicability for future materials discovery problems.

METHODS

High-Throughput NEB for Oxygen Vacancy Diffusion. We first queried MP20 (using mp_api v0.33.3) for all oxides that only contain cations in the space of {Al, Ba, Bi, Ca, Cs, Ga, Ge, Hf, In, K, Li, Mg, Na, Nb, Rb, Si, Sn, Sr, Ta, Ti, Y, Zr}, yielding a total of ∼3700 unique structures. Selection of this chemical space will be motivated later, where we describe our screening for TCH candidate materials. To compute oxygen vacancy diffusivity via a simple lattice site diffusion mechanism in a single oxide, one must first define an O−O neighbor cutoff radius (5.0 Å in this work) to identify all possible oxygen vacancy migration paths, then compute the migration activation energy (via NEB) for all symmetrically inequivalent paths. Depending on its structural/chemical complexity, a single oxide may easily contain $O(10^2)$ or more inequivalent O−O paths; thus, considering the expense of NEB calculations, exhaustively obtaining the migration activation energies needed for diffusivity predictions in a single material can be arduous. Such an approach would be infeasible for the $O(10^4)$ or more materials typically investigated in high-throughput screening studies.

An alternative approach is to screen materials using a surrogate model, which nonetheless requires a training data set of oxygen vacancy migration activation energies computed with NEB. For an N atom host supercell structure, C_h (where h denotes "host"), with total energy E_h, we define an oxygen vacancy migration path by a tuple of site indices corresponding to the start and end sites, $p = (s, e)$. A vacancy is created at the start or end site and the atomic positions are relaxed at constant volume, yielding structures C_s and C_e with total energies E_s and E_e. These two structures and energies, which are themselves independent of the path by which they are connected, provide the fixed end points for an n-image NEB calculation, which yields a sequence of energies,

$$\{E_{p_i}\} = \{E_{p_0=s}, E_{p_1}, \cdots, E_{p_n}, E_{p_{n+1}=e}\} \quad (1)$$

Here $p_{i \in \{1, \cdots, n\}}$ denotes the index of the image along an n-image NEB calculation, and $p_0 = s$ and $p_{n+1} = e$. Relative formation energies are computed via

$$\Delta H_{p_i} = E_{p_i} - E_h + \mu_O^{ref} \quad (2)$$

where μ_O^{ref} is the reference oxygen chemical potential, such that the first and last values in the sequence correspond to relaxed vacancy formation energies. Therefore, $\Delta H_{p_0} = \Delta H_s$ and $\Delta H_{p_{n+1}} = \Delta H_e$ and are path independent. The forward and reverse migration activation energies can then be computed,

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

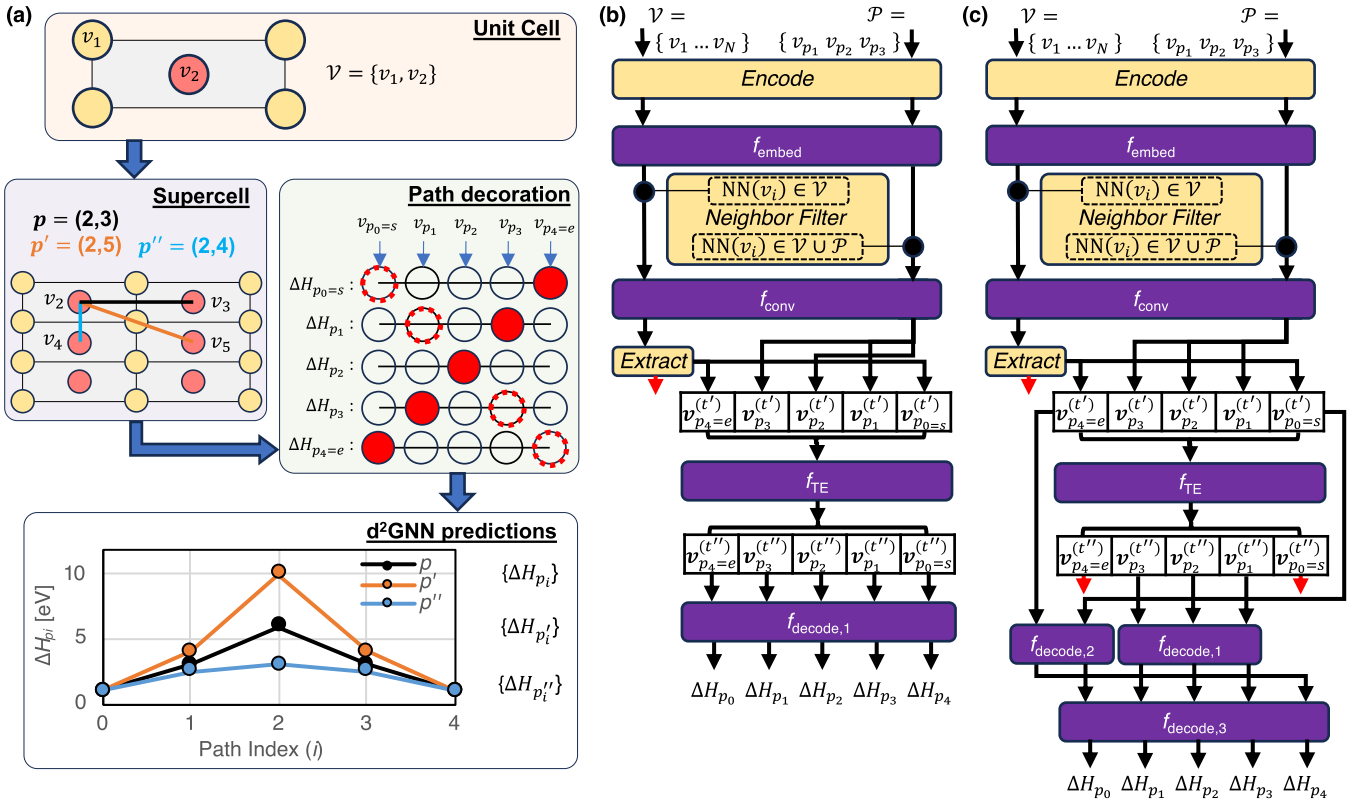


Figure 2. (a) d²GNN utilizes only the relaxed host structure as input. The schematically depicted unit cell contains one inequivalent cation site (ν_1) and one inequivalent anion site (ν_2) on which the vacancy migrates, the supercell contains multiple inequivalent paths ($p \neq p' \neq p''$) between this symmetrically equivalent end point. A given path is decorated with nodes $\nu_{p_{1..n=3}}$ that break the supercell symmetry. The ordering of the energy sequence ($0 \cdots n+1$) is with respect to the vacancy motion (dashed red circle), which is a reverse of the index that represents the occupation of the diffusing species (filled red circle) at a given image coordinate ($n+1 \cdots 0$). (b,c) d²GNN architecture consists of six general steps: graph encoding of the crystal and image nodes, an embedding function for the node features, a nearest neighbor filter that enforces directed message passing in the convolution function, a sequence extraction, a sequence-to-sequence translation via Transformer Encoder (TE), and a decoder step. Purple indicates steps involving learnable parameters, while yellow denotes parameterless operations. Black arrow represent the flow of information, while red arrows represent a discard of features. The improvement of (c) over (b) involves physically constraining the model to preserve eq 5 via independent decoders for the end states and the images.

$$E_A^f = \max(\{\Delta H_{p_1}, \cdots, \Delta H_{p_n}\}) - \Delta H_{p_0}$$
$$E_A^r = \max(\{\Delta H_{p_1}, \cdots, \Delta H_{p_n}\}) - \Delta H_{p_{n+1}} \quad (3)$$

Note that we compute only neutral oxygen vacancy formation and migration activation energies; the merits and limitations of this choice for our desired materials screening applications will be discussed later.

From the ~3700 queried oxides, we a priori prioritized structures and paths for NEB calculations to ensure diversity of chemistry, structure, and path lengths, as shown in Figure 1. We first sort the queried structures based on their chemical system string (e.g., "O−Zr"). Then, among all compounds of a given unique chemical system string, we filter to keep only those with the minimum, median, and maximum canonical energy above the hull (E_H) and assign them a "stability priority" of 0, 1, or 2, respectively. Next, we assign a "complexity priority" equal to the number of elements in the chemical system structure, and we sort the filtered table by "stability priority" and "complexity priority." This final ordered list of prioritized structures provides a simple basis for increasing diversity of calculated training structures (see Figure 1) within a fixed computational budget, for which a priori estimation of how many structures and NEB paths could be calculated was difficult.

We obtained 622 converged NEB calculations spanning 272 structures, each of which represents a unique chemical system. Figure 1a summarizes the chemical diversity of the training data, with color-coding used to represent the number of compounds in which each element appears. Figure 1b counts the number of compounds for each crystal system, grouping them into bins corresponding to binary (b), ternary (t), and quaternary (q) chemical systems.

Before delving into development of complex deep learning-based regression models, it is worthwhile to investigate whether simple linear regression models, crafted with physically motivated features, may provide sufficient accuracy and to provide a benchmark against more complex models. Figure 1c indicates that single variable linear regression models based on simple descriptors can be derived with R^2 as high as ~0.3. Here, $\max(R_{\text{ionic-ov}})$ denotes the maximum in ionic radii overlap (using default pymatgen[44] values) between any atom in the crystal structure and the diffusing O atom along the vector of the linearly interpolated path, $\overline{p} = \overline{r}_s + \eta(\overline{r}_e - \overline{r}_s)$, $\eta \in [0, 1]$, between end and start site coordinates, $\overline{r}_e$ and $\overline{r}_s$, respectively, with $\|\overline{p}\|_2$ denoting the path length. Some interesting qualitative insights are observed, such as an intuitive lower bound to E_A for increasing $\max(R_{\text{ionic-ov}})$ and $\|\overline{p}\|_2$. However, these simple models are generally not accurate enough (all have $R^2 \leq 0.35$) for downstream diffusivity calculations in materials screening applications.

d²GNN Surrogate Model. We now seek a better surrogate model for eq 2, i.e., a single-shot predictor of all NEB energies for a vacancy migration event from one crystallographic site to another. The model should facilitate rapid inference of NEB energies for all symmetrically inequivalent paths in not just one structure but the potentially tens of thousands oftentimes examined in high-throughput screening studies. We therefore seek to construct a model, parametrized by weights θ, that relies only on the relaxed host crystal structure, C_h, and the path tuple p as input,

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

$$\{\Delta H_{p_i}\} = f_{\mathrm{d}^2\mathrm{GNN}}(C_h, p; \theta) \tag{4}$$

The model must make predictions of the final energies without knowledge of the final structure;[24,45] i.e., any structural information derived from the relaxed vacancy crystal structures (C_s or C_e) or image crystal structures (C_{p_i}) cannot be utilized as input to the model. Figure 2a conceptualizes this procedure, whereby symmetrically distinct paths in a unit cell are projected within a supercell, decorated with path image nodes, and subsequently used to predict all sequence energies directly in one shot.

This one-to-many prediction must also address several subtle symmetry constraints, also illustrated schematically in Figure 2a. For symmetrically inequivalent paths, $p \neq p' \neq p''$, within the same crystal structure that share symmetrically equivalent end points, the predicted vacancy formation energies in different sequence predictions must be equivalent,

$$\Delta H_{p_0} = \Delta H_{p_0'} = \Delta H_{p_0''} \tag{5}$$

and therefore independent of the NEB images themselves. The converse is not true, however. Image energies may differ widely between symmetrically inequivalent paths that share equivalent end points,

$$\{\Delta H_{p_1,\cdots,p_n}\} \neq \{\Delta H_{p_1',\cdots,p_n'}\} \neq \{\Delta H_{p_1'',\cdots,p_n''}\} \tag{6}$$

Finally, mirror symmetry should be preserved such that reversing p yields a reversed sequence of energies,

$$\mathrm{rev}(\{\Delta H_{p_i}\}) = f_{\mathrm{d}^2\mathrm{GNN}}(C_h, \mathrm{rev}(p); \theta) \tag{7}$$

These constraints will also be addressed in the following model construction, which summarizes the architectures visualized in Figure 2b,c.

Step 1: Graph Encoding. A graph is constructed for a given symmetrically distinct migration path that consists of $N + n$ nodes for the relaxed host crystal structure's N crystallographic sites, $\mathcal{V} = \{v_1, \cdots, v_N\}$, and the migration path's n images, $\mathcal{P} = \{v_{p_1}, \cdots, v_{p_n}\}$. In other words, $\mathcal{V}$ must be identical across migration path graphs sharing the same base structure, but $\mathcal{P}$ will differ between those paths that are symmetrically distinct. Image nodes' coordinates are linearly interpolated along the displacement vector between v_s and v_e.

Step 2: Embedding. A crystal node's initial feature vector, $\boldsymbol{v}_i^{(0)} = f_{\mathrm{embed}}(\boldsymbol{x}_i)$ is generated by an embedding function on a feature vector that depends only on the elemental identity of site i. Following the embedding approach of SpookyNet,[36] $\boldsymbol{x}_i$ is the ground state electron configuration of the element occupying site i, e.g., for Li with $1s^2 2s^1 2p^0 \cdots$, $\boldsymbol{x}_i = \{2,1,0, \cdots\}$. f_{embed} is a feed forward neural network layer. An image node's initial feature vector is equivalent to that of the path's end points such that $\boldsymbol{v}_{p_i}^{(0)} = \boldsymbol{v}_s^{(0)} = \boldsymbol{v}_e^{(0)}$.

Step 3: Convolutions with Directed Message Passing. GNNs generally use a series of convolution, or message-passing, functions to iteratively update nodes' feature vectors, from which a final property prediction can be made. The challenge here lies in deriving a convolution function that ensures that eqs 5 and 6 are respected. In other words, to preserve eq 5, feature vectors corresponding to crystal sites (which will ultimately be used to predict the vacancy formation energies, ΔH_{p_0} and $\Delta H_{p_{n+1}}$) cannot be influenced by feature vectors corresponding to image nodes (which will be used to predict the image energies, $\Delta H_{p_{\{0\cdots n\}}}$). To ensure eq 6, the converse must be true and the image energies clearly must depend on the path to which they belong, and therefore the end points and their vacancy formation energies. These considerations are also shown schematically in Figure 2.

We therefore propose that all node feature vectors are updated according to an identical convolution (or message passing) function, but that neighbor lists for v_i and v_{p_i} differ in construction and are *directed*. Specifically, a crystal node's neighbor list only consists of other crystal nodes, $\{v_j \in \mathcal{V}\}$, while a path image node's neighbor list may consist of any nodes (either image or crystal), $\{v_j \in \mathcal{V} \cup \mathcal{P}\}$. A series of t convolution functions, $f_{\mathrm{conv}}^{(t)}$, are identically applied to update both crystal and path image nodes,

$$\begin{aligned} \boldsymbol{v}_i^{(t+1)} &= f_{\mathrm{conv}}^{(t)}(v_i^{(t)}, \{v_j^{(t)} \in \mathcal{V}\}) \\ \boldsymbol{v}_{p_i}^{(t+1)} &= f_{\mathrm{conv}}^{(t)}(v_{p_i}^{(t)}, \{v_j^{(t)} \in \mathcal{V} \cup \mathcal{P}\}). \end{aligned} \tag{8}$$

A multitude of convolution architectures with varying complexity have been proposed in recent years.[32−36] While f_{conv} in eq 8 could in principle be based on a variety of these methodologies, given our small training data set, we utilize a low parameter complexity CGCNN[32] convolution function for practical model training purposes. More details of the CGCNN convolutional function and the small modifications used herein are provided in Supplementary Section 4.

Step 4: Sequence Extraction. Graph neural networks for global property predictions often apply a pooling function to aggregate crystal node feature vectors and produce a crystal size-invariant feature vector for downstream manipulation. Or, in the case of local property predictions like defect formation energies,[24] extraction of a single node isolates the feature vector of interest for downstream manipulation and property prediction. Here, after t' total convolutions, we extract an ordered sequence of nodes, $\boldsymbol{X}^{(t')} = \{\boldsymbol{v}_0^{(t')}, \boldsymbol{v}_{p_1}^{(t')}, \cdots, \boldsymbol{v}_{p_n}^{(t')}, \boldsymbol{v}_{n+1}^{(t')}\}$ that featurizes the NEB trajectory for downstream manipulation. Any crystallographic sites other than $p_0 = s$ and $p_{n+1} = e$ are discarded.

Step 5: Sequence-to-Sequence Translation. Next we apply a sequence-to-sequence transformation block to preserve eq 7. This could in principle adopt a variety of forms, although here we choose to exploit the self-attention mechanism of a (low parameter complexity) Transformer Encoder[37] to update the sequence values, $\boldsymbol{X}^{(t'')} = f_{\mathrm{TE}}(\boldsymbol{X}^{(t')})$.

Step 6: NEB Energy Decoder. The final NEB energy sequence, $\{\Delta H_{p_i}\}$, is predicted by one or more decoders, here chosen as a single-input/single-output multilayer perceptron (MLP) as shown in Figure 2. The challenge lies in assigning a frame of reference to the image nodes and their relation to the energy sequence, i.e., visualized in "Path decoration" in Figure 2a. Figure 2b shows the first architecture variation we tested, where the NEB energies are predicted by the same decoder block. However, this formalism does not guarantee the satisfaction of Equation 5, since the self-attention mechanism of the transformer encoder enforces that $\boldsymbol{v}_{p_0}^{(t'')}$ and $\boldsymbol{v}_{p_{n+1}}^{(t'')}$ depend on $\boldsymbol{v}_{p_{i=\{1,\cdots,n\}}}^{(t')}$.

A fundamentally better solution is proposed in Figure 2c, where more than one MLP decoder block is used. Conceptually, a first MLP is used to decode only the image feature vectors, while a second MLP is used to decode the end point feature vectors, but using only the pretransformed sequence of feature vectors. Meanwhile, $\boldsymbol{v}_{p_0}^{(t'')}$ and $\boldsymbol{v}_{p_{n+1}}^{(t'')}$ from the second MLP are discarded. A third and final MLP decoder predicts the final energies. Most importantly, in addition to differentiating the neighbor list construction for path and image nodes (eq 8), this second architecture drops $\{\boldsymbol{v}_{p_0}^{(t'')}, \boldsymbol{v}_{p_{n+1}}^{(t'')}\}$ while updating $\{\boldsymbol{v}_{p_0}^{(t')}, \boldsymbol{v}_{p_{n+1}}^{(t')}\}$ and $\{\boldsymbol{v}_{p_1}^{(t'')}, \cdots, \boldsymbol{v}_{p_n}^{(t'')}\}$ through different decoders to preserve eqs 5 and 6. In other words, the vacancy formation energy (end point) predictions remain independent of the path in which they are being predicted. Meanwhile, using the Transformer encoder as the sequence-to-sequence update function preserves eqs 6 and 7, i.e., the image energies do depend on each other *and* the end points, while also having input reversal symmetry. Thus, the architecture Figure 2c yields a large improvement in the model performance Figure 2b (see Supplementary Section 4).

RESULTS AND DISCUSSION

Model Performance and Validation. The d^{2}GNN model performance is evaluated by nested ($K = 10$, $L = 10$)-fold cross validation. In this scheme, unique structures are randomly split into $K = 10$ outer splits (90% train, 10% test), and all paths in a given structure are assigned accordingly.[46] Thus, no paths from

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

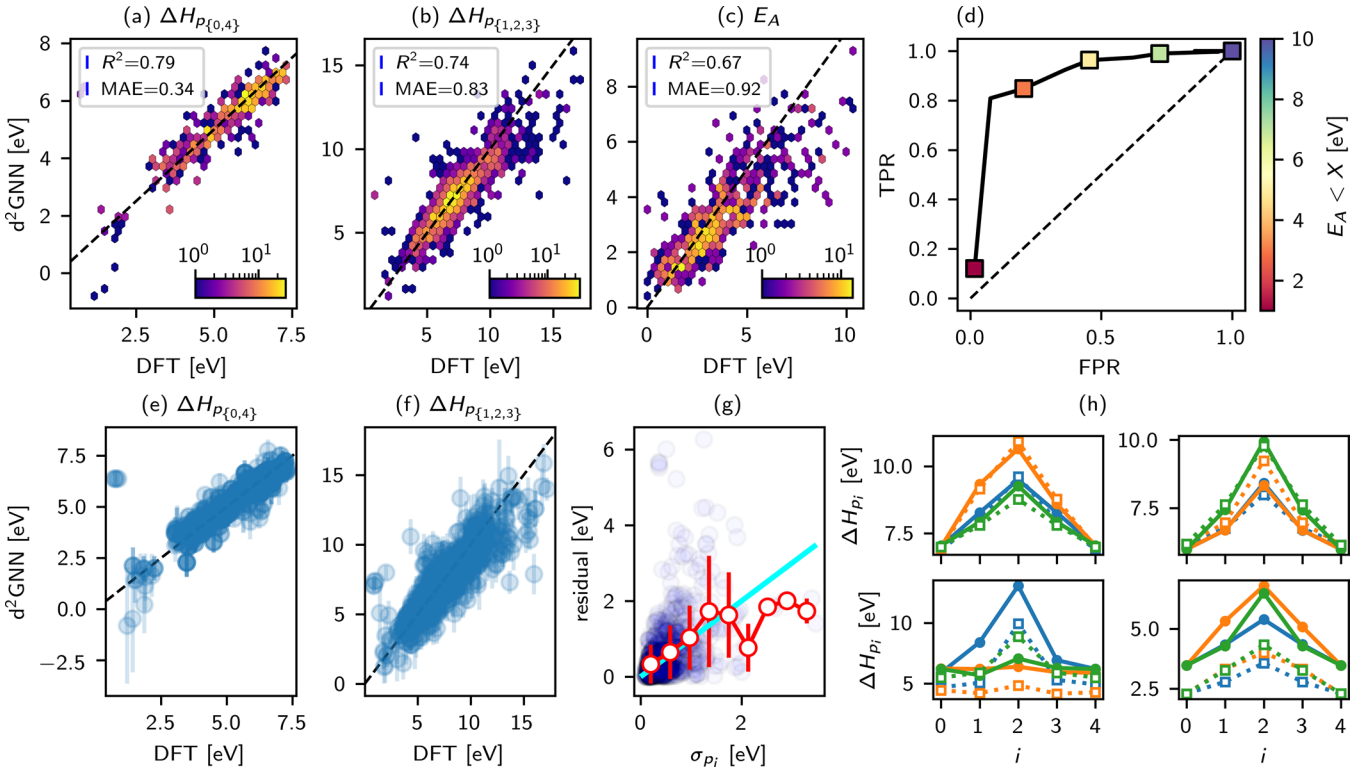


Figure 3. (a) Test set predictions for the vacancy formation energies. (b) Test set predictions for the image energies. (c) Test set predictions for the activation energies. (d) True positive rate vs false positive rate when using the regression model as a classifier for a given threshold of $E_A < X$. (e,f) Test set predictions for the vacancy formation energies and image energies, respectively, with error bars corresponding to σ_{p_i}. (g) Residual error for all predictions vs σ_{p_i}. After binning σ_{p_i}, red circles indicate the average of residuals within a given σ_{p_i} bin, and red error bars indicate the standard deviation of residuals within a given σ_{p_i} bin. (h) d^{2}GNN-predicted NEB energies (dashed) vs DFT-predicted NEB energies (solid) lines for three paths within two well-predicted test set exemplars (top, mp-3952 = BaY_2O_4 and mp-1541753 = Rb_2MgO_2) and two less well-predicted test set exemplars (bottom, mp-30988 = $KBiO_2$ and mp-3870 = $Sr_2Nb_2O_7$) whose absolute ΔH_{p_i} errors are large across all images; however, due to fortuitous cancellation of errors with the i = 0 reference state, these still capture the E_A of the minimum migration activation energy path reasonably well.

the same structure may be split between the train and test sets. For each kth outer train split, the data is randomly divided into L = 10 inner splits (90% train, 10% test). A model is trained for each inner train split (using 10% of the data as validation sets for early stopping), then all inner models are used to predict the outer test set example, $\{\Delta H_{p_i}\}_L$. We take the bootstrapped ensemble average as the final predicted value, $\Delta H_{p_i} = \langle\{\Delta H_{p_i}\}_L\rangle$, and the standard deviation, $\sigma_{p_i} = \sigma(\{\Delta H_{p_i}\}_L)$, as a heuristic metric for the prediction uncertainty.[46,47]

Figure 3a,b show d^{2}GNN vs DFT test set parity for vacancy formation energies and path image energies, respectively. Figure 3c shows d^{2}GNN vs DFT predicted E_A, where some non-systematic error cancellation between the vacancy formation energies and the path image energies reduces the R^2 relative to the individual ΔH_{p_i} predictions. Here we show cross-validated test predictions only for stable materials with hull energies E_H = 0 eV/atom, since model performance begins to degrade for test set predictions of unstable materials with $E_H \gtrsim 0.025$ eV/atom, as discussed in the Supplementary Section 4.

Figure 3d shows the true positive rate (TPR) vs false positive rate (FPR) when using the regression model as a binary classifier to identify whether a given path's E_A is less than a chosen threshold. For low thresholds of $E_A \sim 1$ eV, the model has an 8 times higher TPR than FPR. We therefore expect the model to produce few false positives when performing a preliminary screening for fast-diffusing materials, with the caveat that quantitative comparisons among the most exceptional candidates remains unlikely (i.e., those with connected diffusion pathways with $E_A \ll 1$ eV). The TPR increases rapidly as the threshold is increased to $E_A \sim 2$ eV with only a slight FPR increase.

Figure 3e,f show the test set predictions with error bars corresponding to σ_{p_i}, which is a useful uncertainty metric *on average* that is correlated with the model's residual for a given prediction (Figure 3g).[47] This uncertainty metric is therefore useful for immediately flagging highly out-of-distribution predictions that are likely to be inaccurate, as exemplified by the three lowest test d^{2}GNN-predicted vacancy formation energies that are negative or close to 0 and have very large errors and uncertainties (Figure 3e).

Finally, Figure 3h shows two test set materials with low mean absolute error (MAE) predictions (top row) and two materials with high MAE predictions (bottom row). Detailed investigations on potential sources of error and uncertainty in the d^{2}GNN predictions and the underlying NEB calculations themselves are discussed in Supplementary Section 3 and Supplementary Section 4 (e.g., paths with high atomic overlap or end points corresponding to split vacancies).

Limitations, Merits, and Post Hoc Improvements to the Original Model. Still, the original model exhibits a suboptimal TPR of ~0.13 for $E_A < 1$ eV, and will therefore likely

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

miss many fast diffusing materials during a screening exercise. This is because of the very low likelihood of randomly sampling low $E_A < 1$ eV paths (Figure 4a), for which no simple descriptors

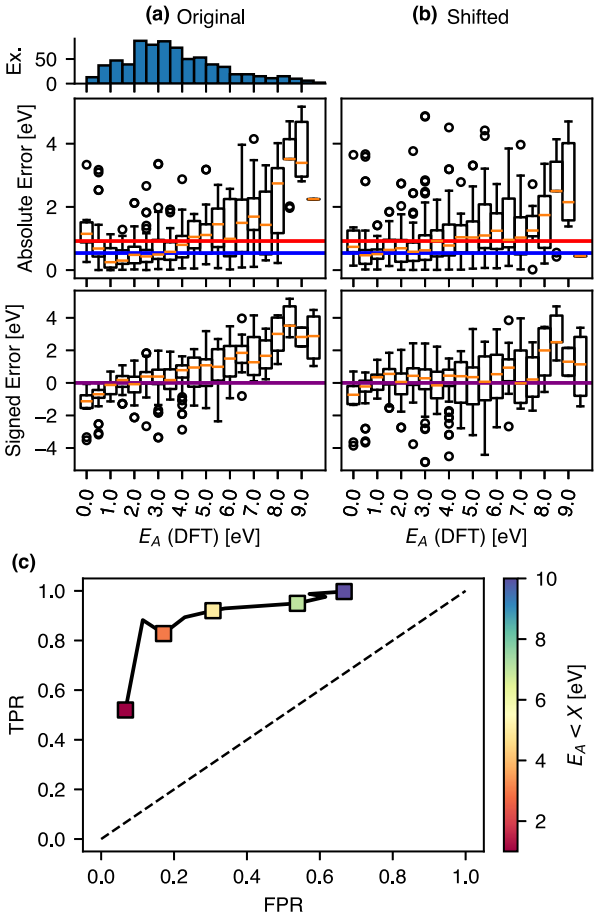


Figure 4. Boxplot distributions of absolute and signed error of CV test set predictions for E_A for the (a) original and (b) empirically shifted model, binned as a function of the truth (DFT) value. The red line corresponds to the overall MAE = 0.92 eV, the blue line corresponds to the MAE = 0.54 eV for DFT-computed paths with $1 < E_A < 4$ eV, and the purple line corresponds to $y = 0$. (a) also shows the histogram distribution of E_A for the training example paths. (c) Empirical shift of eq 9 helps correct for systematic error of the original model in predicting E_A extrema, which boosts TPR for $E_A < 1$ eV by a factor of 4 relative to the original model (Figure 3d).

were a priori discovered that could reliably increase their sampling in our original data set construction (Figure 6c). As a consequence of this under-sampling, the model reverts predictions toward the mean of the training data set. It overpredicts and underpredicts the lowest and highest E_A paths, respectively, as seen in the signed error distributions Figure 4a.

However, for the most commonly sampled paths within $1 \lesssim E_A \lesssim 4$ eV, the MAE of ~0.54 eV is almost half the overall MAE (Figure 3c), which is substantially skewed by the large errors on essentially nondiffusive paths of $E_A \gtrsim 4$ eV. This raises two important observations. First, the model is substantially more accurate than the overall MAE for the majority of paths, which can either compromise a sluggish diffusion network in concert or represent a single rate-limiting path in an otherwise fast network. Second, while the model underestimates, and MAE increases for, entirely nondiffusive paths of $E_A \gtrsim 4$ eV, the absolute accuracy is not relevant here as the model still predicts very high E_A that will correctly prevent these migration hops from contributing to the modeled diffusion network.

Despite the tendency to overestimate the lowest $E_A < 1$ eV paths, the TPR for identifying these paths is a factor of 8 higher than the FPR. A high TPR to FPR ratio is more important than just the absolute TPR value for materials discovery models meant for use in high screening exercises, as it allows one to hone in on interesting materials while avoiding the highly expensive downstream validation (i.e., subsequent DFT calculations or experiments) of unsuccessful/uninteresting materials. In fact, since many positives are still predicted in our high-throughput screening, as discussed in subsequent sections, and these are much more likely to be TPs than FPs, the current version of the model will be useful for a preliminary screen of identifying. Boosting the absolute value of the TPR, i.e., not missing additional promising candidates, will be the priority of future work. This could focus on using the current screening predictions of low E_A paths (next section) as the search space to target low E_A paths with much greater likelihood for additional DFT data collection and subsequent model improvement.

A more immediate but ad hoc solution to boost TPR leverages the systematic error in the model, i.e., the signed error trend in Figure 4a. One could optionally fit a linear regression model that shifts the originally predicted activation energies, E_A^{orig}, to minimize the error with respect to the DFT-calculated activation energies, E_A^{DFT}, via $E_A^{\text{orig}} = m \cdot E_A^{\text{DFT}} + b$. By shifting the model-predicted activation energies according to the optimal $m = 0.62$ and $b = 0.98$,

$$E_{\text{A}} = E_A^{\text{orig}} - b - (m - 1)\cdot E_A^{\text{orig}} \tag{9}$$

one can systematically improve very low and very high E_A predictions (Figure 4b) and 4-fold increase the TPR to 0.52 for the $E_A < 1$ eV threshold. As expected, this empirical shift slightly increases the FPR, but crucially, the TPR to FPR ratio remains unchanged at ~8. Thus, we expect to proportionally waste no more time in the downstream validation of uninteresting materials while significantly boosting the number of potential candidates captured in the screening.

High-Throughput Screening. As shown in Figure 5a for one example material, we first enumerate all possible inequivalent O−O paths in the crystal structure (up to 5 Å), then use d^{2}GNN to predict $\{\Delta H_{p_i}\}$ for all p, and finally assign the predicted activation energies to each possible symmetrically equivalent path in the migration path network. A critical aspect of our screening is the bootstrapped uncertainty metric, σ_{p_i}, which we use to eliminate materials with uncertainty above a chosen threshold. From hereon we only show materials predictions where the average uncertainty across all paths in a given material, $\{p\}$, is below a threshold, $\langle\sigma_{p_i}\rangle_{\{p\}} < 0.75$ eV. Figure 5b shows the distribution of all computed activation energies and vacancy formation enthalpies ($\Delta H_V = \Delta H_{p_{\{0,4\}}}$) across these screened compounds, both of which are highly skewed. The raw screening data, including the high uncertainty predictions, also reveals anomaly/outlier predictions that can be readily identified as out-of-distribution materials with respect to our training data; these are therefore discarded on the basis of uncertainty (Supplementary Section 4).

Our goal is then to predict the temperature-dependent oxygen vacancy diffusivity tensor, D, in a high-throughput manner to

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

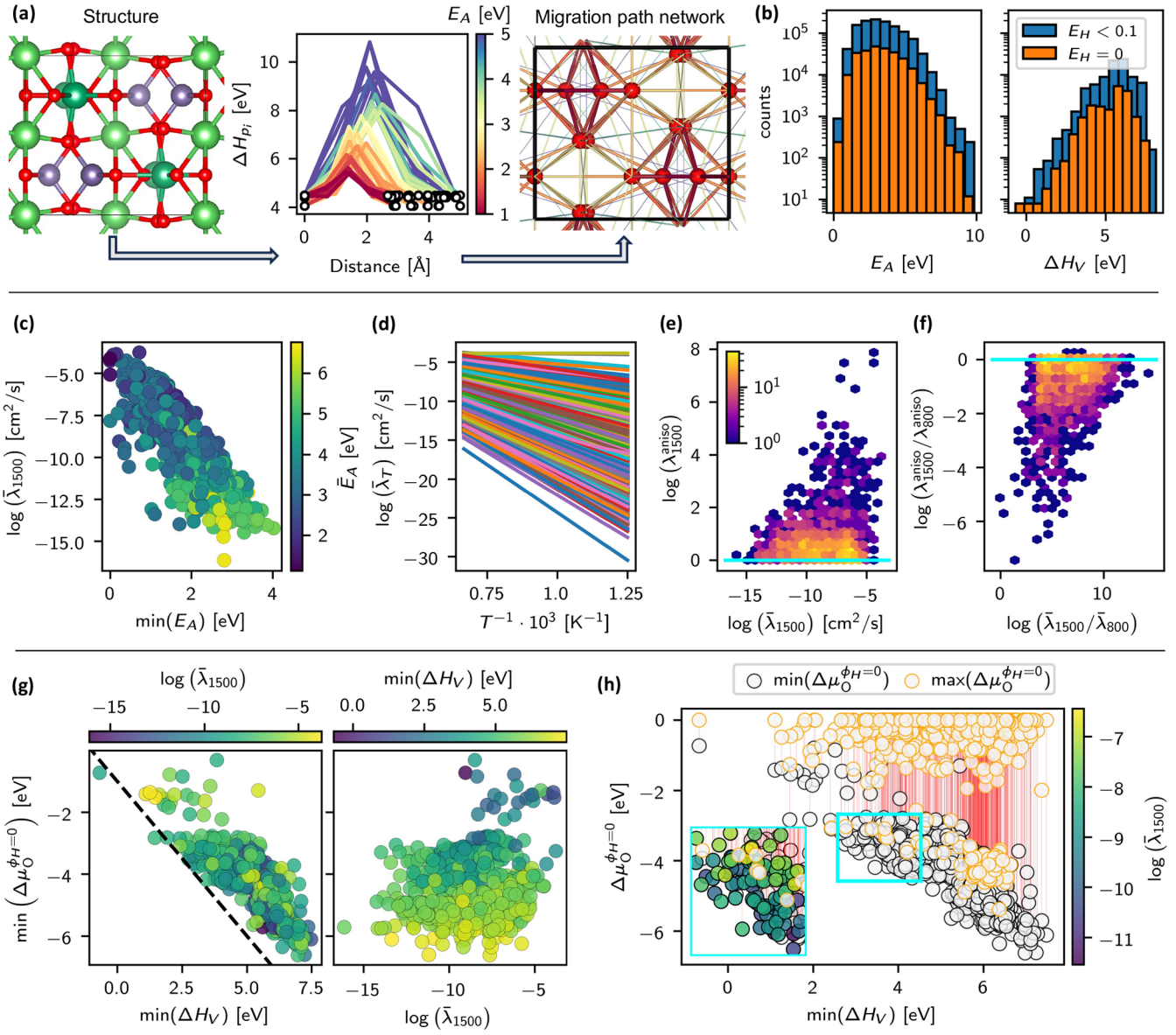


Figure 5. (a) Visualization of the mp-6456 structure (LiNbGeO5), the predicted $\{\Delta H_{p_i}\}$ for all paths less than 5 Å (color-coded by E_A) with white circles for $\Delta H_{p_{[0,4]}}$, and migration path network (red spheres are oxygen atoms, edges are paths color-coded by E_A). (b) Distribution of E_A and ΔH_V across all screened structures, color-coded by E_H threshold. For only near hull materials $E_H < 0.025$ eV/atom: (c) Mean diffusivity vs min(E_A) for all structures, color-coded by $\overline{E}_A$. (d) Arrhenius plot of the mean diffusivity vs inverse temperature. (e) High-temperature diffusion anisotropy vs mean diffusivity and (f) the anisotropy ratio between high and low temperatures vs the change in mean diffusivity, where the colorbar represents the material count. For materials with non-null $\Delta\mu_O^{\phi_H=0}$: (g) minimum host stability vs minimum vacancy formation energy, color-coded by mean diffusivity. The dashed black line depicts $y = -x - 1$, an approximate lower bound to the host vs defect stability correlation. (h) min host stability (gray circles) and max host stability (orange circles) are connected by a red line. Within the cyan inset, materials satisfying the strict set of TCH down-selection criteria are color-coded by mean diffusivity.

identify materials with useful (i.e., fast, slow, aniostropic) diffusivities and to elucidate relevant correlations between vacancy defect thermodynamics, kinetics, host stability, and chemistry. Given the vacancy sites, their energies, the connectivity of sites, and transition state energies, the master equation for the diffusivity is solved directly at each temperature T. The direct solution,[17] implemented in the open source Onsager software (https://github.com/DallasTrinkle/Onsager), is more computationally efficient than a Monte Carlo approach, and accuracy is limited only by roundoff error. Using the d^{2}GNN predictions, we can therefore obtain the temperature dependent oxygen vacancy diffusivity tensors, D_T, for each screened material. Units for all reported diffusivity values are in cm^2/s.

Because our model is less accurate for metastable and unstable materials that are less likely to be able to be synthesized, we show D_T ($T \in [800,1500]$ K) predictions in Figure 5c–f only for materials with $E_H < 0.025$ eV/atom. We denote the eigenvalues of D_T as λ_T, their average as the mean diffusivity, $\overline{\lambda}_T$, and the ratio between the maximum and minimum eigenvalues as the diffusion anisotropy, $\lambda_T^{aniso} = \lambda_T^{max}/\lambda_T^{min}$. For a given material, we plot $\overline{\lambda}_{1500}$ vs the minimum activation energy min(E_A) and color-coded by the average of all paths, $\overline{E}_A$. As expected, a small

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

$\min(E_A)$ is a necessary but insufficient criteria for a large diffusivity, as up to 8 orders of magnitude change in diffusivity can be observed for materials with $\min(E_A) < 0.5$ eV. This indicates the importance of a screening model that can estimate activation energies of all possible (including high E_A) paths in a structure as well. Figure 5d shows an Arrhenius plot of diffusivity and the orders of magnitude change among different materials and across temperatures, some of which deviate from Arrhenius behavior in the modeled temperature range (see Supplementary Section 7 for details). Figure 5e reveals that most materials have near-isotropic diffusivities, but some can have λ_T^{max} as much as 8 orders of magnitude greater than λ_T^{min}; furthermore, among those with the highest mean diffusivities (e.g., $>10^{-6}$ cm^2/s), a higher percentage of materials tend also to have large λ_T^{aniso}. This anisotropy is also highly temperature dependent, i.e., some materials that are isotropic at high temperature become highly anisotropic at lower temperatures (Figure 5f).

Among our screened materials, we can isolate those with interesting predicted diffusivity properties (see Supplementary Section 6), a sampling of which is shown in Figure 6 and summarized in Table 1. These materials were selected for (a) large mean diffusivity, (b) small mean diffusivity, and (c) large mean and highly anisotropic diffusivities, respectively. Materials (a) and (c) respectively contain a 3d-connected and 2d-connected network of low E_A paths, while material (b) only possesses a periodically connected migration network with relatively high E_A. Note that on an absolute scale, 65 of the ~3700 (or ~2%) screened oxides are predicted with $\min(E_A) < 1$ eV. Even though the absolute number of materials with $\min(E_A) < 1$ could likely be higher (possibly by an heuristic factor of 2−6, depending on whether one uses the TPR for the original or empirically shifted model), these identified candidates are highly likely to be true positive rather than false positive identifications because of the high TPR to FPR ratio.

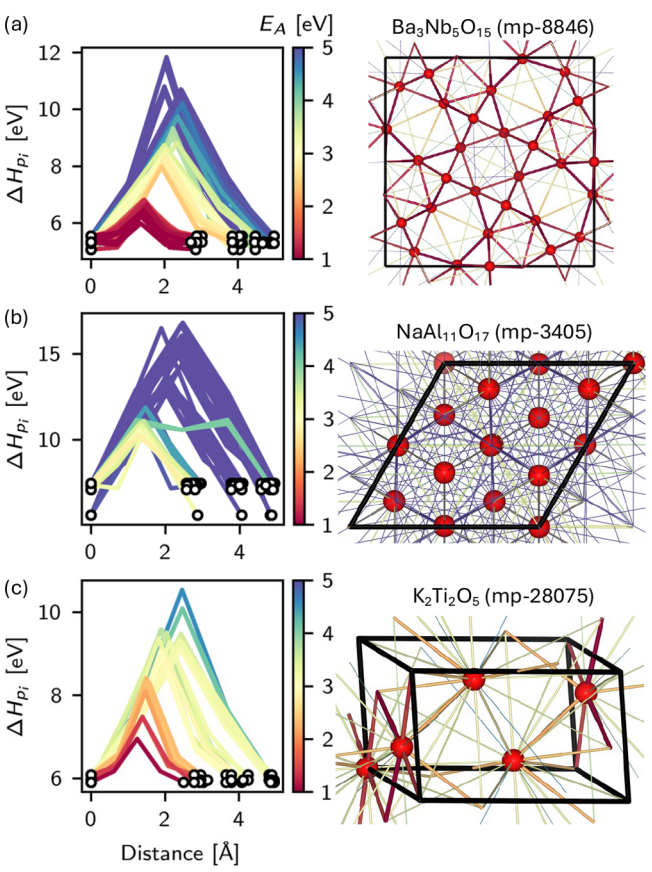


Figure 6. For three example structures with (a) large mean diffusivity, (b) small mean diffusivity, and (c) large mean and highly anisotropic diffusivities, we plot ΔH_{p_i} for all paths less than 5 Å (color-coded by E_A) and visualize the migration path network, with red spheres representing oxygen sites and edges representing possible migration paths (also colored by their E_A).

Table 1. Summary of the Diffusivity Related Predictions (Formula, MP ID, $\min(E_A)$, $\overline{E}_A$, $\max(E_A)$, $\overline{\lambda}_{1500}$, and λ_{1500}^{aniso} for the Three Example Structures Shown in Figure 6

ID	formula	$\min(E_A)$	$\overline{E}_A$	$\log(\overline{\lambda}_{1500})$	$\log(\lambda_{1500}^{aniso})$
mp-8846	$Ba_3Nb_5O_{15}$	0.6	3.3	−5.1	0.2
mp-3405	$NaAl_{11}O_{17}$	2.8	6.6	−16.1	0.2
mp-28075	$K_2Ti_2O_5$	1.0	3.0	−5.7	3.5

Defect Thermodynamics, Kinetics, and Host Stability Correlations. Some materials discovery exemplars necessitate identifying candidates with simultaneous target metrics for vacancy defect thermodynamics, diffusivity, and host stability, all of which may be interdependent. Defining the oxygen chemical potential difference with respect to the reference state, $\Delta\mu_O = \mu_O - \mu_O^{ref}$, and using MP-computed formation energies for all oxides in the screening space, we compute a stability range, denoted $\Delta\mu_O^{\phi_H=0} = [\Delta\mu_O \mid \phi_H(\Delta\mu_O) = 0]$. This yields the $\Delta\mu_O$ range over which a given material is stable, i.e., its grand energy above the hull, ϕ_H, equals zero. Figure 5g shows a strong correlation between the minimum value of the stability range, $\min(\Delta\mu_O^{\phi_H=0})$, vs the minimum vacancy formation energy in a given material, with an approximate lower bound of $\min(\Delta\mu_O^{\phi_H=0}) \approx -1$ eV − $\min(\Delta H_V)$. This trade-off constrains the space in which materials with desired host and defect stability can be designed. Meanwhile, neither criteria strongly correlates with the diffusivity $\overline{\lambda}_{1500}$, which can be tuned more independently.

Thermochemical Hydrogen Production Materials Screening. One example where all three down-selection criteria are important is the identification of metal oxides for thermochemical hydrogen (TCH) production, i.e., water-splitting via a direct two-step thermal redox cycle.[39,41] Criterion #1 is that $\Delta H_V \in [2.5,4.0]$ eV,[48] which has been addressed by a variety of surrogate model-based screenings. Criterion #2 necessitates that $\Delta\mu_O^{\phi_H=0}$ span, or at least intersect, the typical TCH redox oxygen chemical potential range between thermal reduction and water-splitting, $\Delta\mu_O^{TCH} = [-3.0,-2.5]$ eV.[49,50] Finally, Criterion #3 for screening TCH materials, hitherto unexplored via high-throughput screening, is that the oxygen vacancy diffusivity, shown in Figure 5h, be as fast as possible. The $\overline{\lambda}_{1500}$ color-coded materials in the cyan inset box are the small fraction of screened compounds that satisfy both Criteria #1 and #2. As shown by the red line connecting $\min(\Delta\mu_O^{\phi_H=0})$ and $\max(\Delta\mu_O^{\phi_H=0})$, many materials possess an excellent stability range for TCH, but this is inextricably correlated with vacancy formation energies that are too high, which eliminates a large swath of candidates. Only $O(10^2)$ candidates survive these downselection criteria, as examined in more detail in Supplementary Section 8.

The selection of the cationic chemical space in the present work excludes the 3*d* transition metals, with the exception of Ti, which occurs mostly as Ti^{4+} in a d^0 configuration. On one hand,

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

this selection allowed us to limit the scope of this initial work for d^{2}GNN to nonmagnetic DFT calculations. On the other hand, this selection also addresses the hypothesis that high-performance materials for TCH might be found outside the realm of 3*d* transition metal oxides, where the vast majority of recent research efforts has been focused.[51] While numerous discoveries of 3*d* oxides have been made over the past decade, especially including manganates,[40−42] the general trend remains that these oxides split water only under dilute H_2:H_2O conditions.[41,52] This drawback can be attributed to an insufficient reduction entropy, resulting from the presence of repulsive defect interactions that reduce the (atomic) configurational entropy contribution[53] and from the lack of electronic entropy contributions.[49,54]

The strong electron correlations in the 3*d* shell cause localized electronic states, which nevertheless experience significant interactions with the ligands,[55] as well as the existence of multiple oxidation states. As a consequence, the excess electrons introduced by the O vacancy defects tend to be strongly bound to the defect site, resulting in a large ionization energy, which precludes the electronic entropy contribution.[56] While there could be exceptions to this trend for suitable combinations of the transition metal electron configuration and the ligand coordination environment,[57] a potentially rewarding alternative materials search strategy could focus on cations with more delocalized nonmagnetic electronic states, as included in the present work, where a smaller ionization energy could enable water splitting under much more concentrated H_2:H_2O mixtures,[56] which is currently achieved only in CeO_2. While the present work does not directly address the ionization energy and resulting electronic entropy, it does provide the basis for selecting candidate materials for additional electronic structure studies with beyond-DFT methods that are required to address this question. Because previous related work[24] (i.e., a vacancy formation energy-only model) "re-discovered" many known TCH materials containing 3*d* transition metals, we anticipate this work could identify potential candidates but in this unconventional TCH space, within which we are unaware of any existing TCH demonstrations.

CONCLUSIONS

We have provided the generalized d^{2}GNN framework that can perform surrogate model predictions for NEB energies in arbitrary crystal structures and vacancy migration paths. Relying only on a host structure input and the specification of the end points corresponding to the vacancy migration hop, the model facilitates rapid inference on vacancy thermodynamics (formation energies) and kinetics (migration activation energies), while respecting relevant symmetry constraints thanks to an intermediary Transformer encoder and MLP decoder architecture. Then, for a given material and its vacancy sites, their predicted formation energies, the connectivity of sites, and predicted migration activation energies, the master equation for the diffusivity is solved directly at each temperature *T*. Many materials can then be screened in a high-throughput manner to identify candidates with anomalous/interesting vacancy diffusivity properties. As a proof-of-concept, we conducted high-throughput screening of a large chemical space of metal oxides, combining neutral vacancy defect thermodynamics, kinetics, and host stability, to identify promising candidates for thermochemical hydrogen production via two-step water-splitting redox cycles.

Several outstanding challenges remain and point to exciting areas for improvement in future work. The chemical space across which the model has been trained is limited to 14 cations, so expanding the high-throughput NEB calculations to more diverse (but more challenging for DFT) materials like oxides with 3*d* transition metals and rare earths will improve its applicability for materials discovery across different domains. Similarly, our DFT training data only consists of NEB trajectories for oxygen vacancy migration, and this could be extended to different element types to again improve applicability in other materials discovery/modeling domains. The current form of the model also only accounts for vacancy migration events occurring between lattice oxygen sites, while hops between interstitial sites are often favorable.[31] These additional pathways potentially could be accounted for if a suitable modification of the current base model can be derived, but such a modification would also require substantial additional data collection.

As discussed in more detail in Supplementary Section 2, this initial demonstration only focused on neutral oxygen vacancies, which often diffuse more slowly than charged vacancies (e.g., V_O^{2+}) in materials with band gaps. This tendency is apparent in Figure 7, in which we show the correlation between E_A for

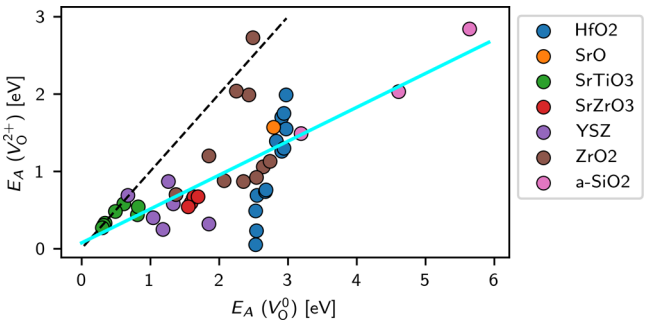


Figure 7. Comparison of calculated migration activation energies for oxygen vacancies in the 0 and +2 charge state extracted from various literature reports.[12,14,58,59] The black dashed line represents $y = x$, and the cyan line represents the line of best fit ($y = 0.43x + 0.07$).

vacancies in the neutral and +2 charge states for several materials studied in previous computational work. E_A is almost universally smaller when vacancies are charged, often significantly so. Therefore, in cases where charged vacancies are thermodynamically favored compared with neutral vacancies at the equilibrium Fermi level, we anticipate that our predictions represent at worst a lower bound on the vacancy diffusivity. Any fast diffusers currently predicted by the neutral vacancy model would remain correctly predicted, although it is clear that some fast diffusers that rely on charged vacancy migration will be missed by the current model. Thus, the TPR rate of our model may decrease somewhat, while, crucially, the FPR will remain unaffected. Therefore, it would be beneficial in future work to repeat certain calculations with charged vacancies[25,60−62] and to estimate the effect of charge on diffusivity predictions in a large-scale screening.

Finally, with the advent of machine learning methods that can model potential energy surfaces of inorganic materials across large portions of the periodic table,[33,63−65] it will be interesting to compare the relative accuracy of this method to that of NEB calculations performed using machine learned force fields, which may not have been explicitly trained on high energy migration barriers.

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

■ ASSOCIATED CONTENT

Data Availability Statement

Code and training data to retrain the model are provided at https://github.com/sandialabs/d2gnn.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.chemmater.5c00021.

Additional details regarding DFT settings, DFT benchmarking, NEB training data statistics, d^2GNN model validation, and screening predictions (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Matthew D. Witman − *Sandia National Laboratories, Livermore, California 94551, United States;* orcid.org/0000-0001-6263-5114; Email: mwitman@sandia.gov

Authors

Lauren Way − *Sandia National Laboratories, Livermore, California 94551, United States*

Catalin D. Spataru − *Sandia National Laboratories, Livermore, California 94551, United States*

Reese E. Jones − *Sandia National Laboratories, Livermore, California 94551, United States;* orcid.org/0000-0002-2332-6279

Dallas R. Trinkle − *Department of Materials Science and Engineering, University of Illinois, Urbana-Champaign, Illinois 61801, United States*

Andrew J. E. Rowberg − *Laboratory for Energy Applications for the Future (LEAF) and Quantum Simulations Group, Lawrence Livermore National Laboratory, Livermore, California 94550, United States;* orcid.org/0000-0002-2928-6917

Joel B. Varley − *Laboratory for Energy Applications for the Future (LEAF) and Quantum Simulations Group, Lawrence Livermore National Laboratory, Livermore, California 94550, United States;* orcid.org/0000-0002-5384-5248

Robert B. Wexler − *Department of Chemistry and Institute of Materials Science and Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, United States;* orcid.org/0000-0002-6861-6421

Christopher M. Smyth − *Sandia National Laboratories, Albuquerque, New Mexico 87123, United States*

Tyra C. Douglas − *Sandia National Laboratories, Albuquerque, New Mexico 87123, United States;* orcid.org/0000-0002-1142-6846

Sean R. Bishop − *Sandia National Laboratories, Albuquerque, New Mexico 87123, United States*

Elliot J. Fuller − *Sandia National Laboratories, Livermore, California 94551, United States;* orcid.org/0000-0001-7912-8688

Anthony H. McDaniel − *Sandia National Laboratories, Livermore, California 94551, United States*

Stephan Lany − *National Renewable Energy Laboratory, Golden, Colorado 80401, United States;* orcid.org/0000-0002-8127-8885

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemmater.5c00021

Author Contributions

$^{\nabla}$L.W., C.D.S., and M.D.W. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge research support from the Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan. The National Renewable Energy Laboratory (NREL) is operated for the DOE under Contract No. DE-AC36-08GO28308. The work at the Lawrence Livermore National Laboratory was performed under the auspices of the U.S. Department of Energy (DOE) under Contract No. DE-AC52-07NA27344. This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Fuel Cell Technologies Office (FCTO) under Award Number DE-EE0010733. The authors gratefully acknowledge research support from the HydroGEN Advanced Water Splitting Materials Consortium, established as part of the Energy Materials Network under the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Fuel Cell Technologies Office, under Award Number DE-EE0010733.

■ REFERENCES

(1) Mehrer, H. *Diffusion in Solids: Fundamentals, Methods, Materials, Diffusion-Controlled Processes*; Springer Science & Business Media: 2007; Vol. *155*.

(2) Shewmon, P. *Diffusion in Solids*; Springer: 2016.

(3) Freer, R. Self-diffusion and impurity diffusion in oxides. *J. Mater. Sci.* **1980**, *15*, 803−824.

(4) Carrasco, J.; Lopez, N.; Illas, F. First principles analysis of the stability and diffusion of oxygen vacancies in metal oxides. *Physical review letters* **2004**, *93*, No. 225502.

(5) Martin, M. In *Diffusion in Condensed Matter: Methods, Materials, Models*; Springer: 2005; pp 209−247.

(6) Pelleg, J. *Diffusion in Ceramics*; Springer: 2016.

(7) Ishigaki, T.; Yamauchi, S.; Kishio, K.; Mizusaki, J.; Fueki, K. Diffusion of oxide ion vacancies in perovskite-type oxides. *J. Solid State Chem.* **1988**, *73*, 179−187.

(8) Van Orman, J. A.; Crispin, K. L. Diffusion in oxides. *Reviews in Mineralogy and Geochemistry* **2010**, *72*, 757−825.

(9) Kofstad, P. *High Temperature Corrosion*. Elsevier Applied Science Publishers: Crown House, Linton Road, Barking, Essex IG 11 8 JU, UK, 1988.

(10) Kofstad, P. Defects and transport properties of metal oxides. *Oxidation of metals* **1995**, *44*, 3−27.

(11) Yang, M. Y.; Kamiya, K.; Magyari-Köpe, B.; Niwa, M.; Nishi, Y.; Shiraishi, K. Charge-dependent oxygen vacancy diffusion in Al2O3-

based resistiverandom-access-memories. *Appl. Phys. Lett.* **2013**, *103*, No. 093504.
(12) Hur, J. H. First principles study of oxygen vacancy activation energy barrier in zirconia-based resistive memory. *Sci. Rep.* **2020**, *10*, 5405.
(13) Roy, A.; Singh, P.; Balasubramanian, G.; Johnson, D. D. Vacancy formation energies and migration barriers in multi-principal element alloys. *Acta Mater.* **2022**, *226*, No. 117611.
(14) Rowberg, A. J. E.; Slomski, H. S.; Kim, N.; Strange, N. A.; Gorman, B. P.; Shulda, S.; Ginley, D. S.; Kweon, K. E.; Wood, B. C. Impact of Sr-Containing Secondary Phases on Oxide Conductivity in Solid-Oxide Electrolyzer Cells. *Chem. Mater.* **2024**, *36*, 6464−6474.
(15) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(16) Sheppard, D.; Xiao, P.; Chemelewski, W.; Johnson, D. D.; Henkelman, G. A generalized solid-state nudged elastic band method. *J. Chem. Phys.* **2012**, *136*, No. 074103.
(17) Trinkle, D. R. Diffusivity and derivatives for interstitial solutes: activation energy, volume, and elastodiffusion tensors. *Philos. Mag.* **2016**, *96*, 2714−2735.
(18) Van der Ven, A.; Ceder, G.; Asta, M.; Tepesch, P. D. First-principles theory of ionic diffusion with nondilute carriers. *Phys. Rev. B* **2001**, *64*, No. 184307.
(19) Rehman, T.; Jaipal, M.; Chatterjee, A. A cluster expansion model for predicting activation barrier of atomic processes. *J. Comput. Phys.* **2013**, *243*, 244−259.
(20) Li, C.; Nilson, T.; Cao, L.; Mueller, T. Predicting activation energies for vacancy-mediated diffusion in alloys using a transition-state cluster expansion. *Phys. Rev. Mater.* **2021**, *5*, No. 013803.
(21) Shen, J. X.; Li, H. H.; Rutt, A.; Horton, M. K.; Persson, K. A. Topological graph-based analysis of solid-state ion migration. *npj Comput. Mater.* **2023**, *9*, 99.
(22) Wexler, R. B.; Gautam, G. S.; Stechel, E. B.; Carter, E. A. Factors Governing Oxygen Vacancy Formation in Oxide Perovskites. *J. Am. Chem. Soc.* **2021**, *143*, 13212−13227.
(23) Choudhary, K.; Sumpter, B. G. Can a deeplearning model make fast predictions of vacancy formation in diverse materials? *AIP Adv.* **2023**, *13*, No. 095109.
(24) Witman, M. D.; Goyal, A.; Ogitsu, T.; McDaniel, A. H.; Lany, S. Defect graph neural networks for materials discovery in high-temperature cleanenergy applications. *Nat. Comput. Sci.* **2023**, *3*, 675−686.
(25) Rahman, M. H.; Gollapalli, P.; Manganaris, P.; Yadav, S. K.; Pilania, G.; DeCost, B.; Choudhary, K.; Mannodi-Kanakkithodi, A. Accelerating defect predictions in semiconductors using graph neural networks. *APL Mach. Learn.* **2024**, *2*, No. 016122.
(26) Mosquera-Lois, I.; Kavanagh, S. R.; Ganose, A. M.; Walsh, A. Machine-learning structural reconstructions for accelerated point defect calculations. *npj Comput. Mater.* **2024**, *10*, 121.
(27) Kavanagh, S. R.; Squires, A. G.; Nicolson, A.; Mosquera-Lois, I.; Ganose, A. M.; Zhu, B.; Brlec, K.; Walsh, A.; Scanlon, D. O. doped: Python toolkit for robust and repeatable charged defect supercell calculations. *J. Open Source Softw.* **2024**, *9*, 6433.
(28) Lee, Y.-L.; Kleis, J.; Rossmeisl, J.; Shao-Horn, Y.; Morgan, D. Prediction of solid oxide fuel cell cathode activity with first-principles descriptors. *Energy Environ. Sci.* **2011**, *4*, 3966.
(29) Lu, G. M.; Witman, M.; Agarwal, S.; Stavila, V.; Trinkle, D. R. Explainable machine learning for hydrogen diffusion in metals and random binary alloys. *Phys. Rev. Mater.* **2023**, *7*, No. 105402.
(30) López, C.; Emperador, A.; Saucedo, E.; Rurali, R.; Cazorla, C. Universal ion-transport descriptors and classes of inorganic solid-state electrolytes. *Mater. Horizons* **2023**, *10*, 1757−1768.
(31) Meng, J.; Sheikh, M. S.; Jacobs, R.; Liu, J.; Nachlas, W. O.; Li, X.; Morgan, D. Computational discovery of fast interstitial oxygen conductors. *Nat. Mater.* **2024**, *23*, 1252−1258.
(32) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, No. 145301.
(33) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31*, 3564−3572.
(34) Park, C. W.; Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **2020**, *4*, No. 063801.
(35) Choudhary, K.; DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput. Mater.* **2021**, *7*, 185.
(36) Unke, O. T.; Chmiela, S.; Gastegger, M.; Schütt, K. T.; Sauceda, H. E.; Müller, K. R. SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nat. Commun.* **2021**, *12*, 7273.
(37) Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I. In *Advances in Neural Information Processing Systems*, ed. by Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; Garnett, R., Advances in Neural Information Processing Systems, 2017; Vol. *30*.
(38) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*, No. 011002.
(39) Chueh, W. C.; Falter, C.; Abbott, M.; Scipio, D.; Furler, P.; Haile, S. M.; Steinfeld, A. High-Flux Solar-Driven Thermochemical Dissociation of CO2 and H 2 O Using Nonstoichiometric Ceria. *Science (80-.).* **2010**, *330*, 1797−1801.
(40) McDaniel, A. H.; Miller, E. C.; Arifin, D.; Ambrosini, A.; Coker, E. N.; O'Hayre, R.; Chueh, W. C.; Tong, J. Sr-and Mn-doped $LaAlO_{3-\delta}$ for solar thermochemical H_2 and CO production. *Energy Environ. Sci.* **2013**, *6*, 2424−2428.
(41) Barcellos, D. R.; Sanders, M. D.; Tong, J.; McDaniel, A. H.; O'Hayre, R. P. BaCe 0.25 Mn 0.75 O 3-δ—a promising perovskite-type oxide for solar thermochemical hydrogen production. *Energy Environ. Sci.* **2018**, *11*, 3256−3265.
(42) Wexler, R. B.; Sai Gautam, G.; Bell, R. T.; Shulda, S.; Strange, N. A.; Trindell, J. A.; Sugar, J. D.; Nygren, E.; Sainio, S.; McDaniel, A. H.; Ginley, D.; Carter, E. A.; Stechel, E. B. Multiple and nonlocal cation redox in Ca−Ce−Ti−Mn oxide perovskites for solar thermochemical applications. *Energy Environ. Sci.* **2023**, *16*, 2550−2560.
(43) Knausgård Hommedal, Y.; Etzelmüller Bathen, M.; Mari Reinertsen, V.; Magnus Johansen, K.; Vines, L.; Kalmann Frodason, Y. Theoretical modeling of defect diffusion in wide bandgap semiconductors. *J. Appl. Phys.* **2024**, *135*, 170902.
(44) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(45) Gibson, J.; Hire, A.; Hennig, R. G. Dataaugmentation for graph neural network learning of the relaxed energies of unrelaxed structures. *npj Comput. Mater.* **2022**, *8*, 211.
(46) Witman, M. D.; Schindler, P. MatFold: systematic insights into materials discovery models' performance through standardized cross-validation protocols. *Digital Discovery* **2025**, *4*, 625−635.
(47) Palmer, G.; Du, S.; Politowicz, A.; Emory, J. P.; Yang, X.; Gautam, A.; Gupta, G.; Li, Z.; Jacobs, R.; Morgan, D. Calibration after bootstrap for accurate uncertainty quantification in regression models. *npj Comput. Mater.* **2022**, *8*, 115.
(48) Emery, A. A.; Saal, J. E.; Kirklin, S.; Hegde, V. I.; Wolverton, C. High-throughput computational screening of perovskites for thermochemical water splitting applications. *Chem. Mater.* **2016**, *28*, 5621−5634.
(49) Lany, S. Communication: The electronic entropy of charged defect formation and its impact on thermochemical redox cycles. *J. Chem. Phys.* **2018**, *148*, No. 071101.
(50) Millican, S. L.; Clary, J. M.; Musgrave, C. B.; Lany, S. Redox Defect Thermochemistry of FeAl2O4 Hercynite in Water Splitting from First-Principles Methods. *Chem. Mater.* **2022**, *34*, 519−528.
(51) Tran, J. T.; Warren, K. J.; Wilson, S. A.; Muhich, C. L.; Musgrave, C. B.; Weimer, A. W. An updated review and perspective on efficient

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX

hydrogen generation via solar thermal water splitting. *WIREs Energy Environ.* **2024**, *13*, No. e528.
(52) Zhai, S.; Rojas, J.; Ahlborg, N.; Lim, K.; Toney, M. F.; Jin, H.; Chueh, W. C.; Majumdar, A. The use of poly-cation oxides to lower the temperature of two-step thermochemical water splitting. *Energy Environ. Sci.* **2018**, *11*, 2172−2178.
(53) Goyal, A.; Sanders, M. D.; O'Hayre, R. P.; Lany, S. Predicting Thermochemical Equilibria with Interacting Defects: $Sr_{1-x}CexMnO_{3-\delta}$ Alloys for Water Splitting. *PRX Energy* **2024**, *3*, No. 013008.
(54) Naghavi, S. S.; Emery, A. A.; Hansen, H. A.; Zhou, F.; Ozolins, V.; Wolverton, C. Giant onsite electronic entropy enhances the performance of ceria for water splitting. *Nat. Commun.* **2017**, *8*, 285.
(55) Rao, C. N. R. Transition metal oxides. *Annu. Rev. Phys. Chem.* **1989**, *40*, 291−326.
(56) Lany, S. Chemical Potential Analysis as an Alternative to the van't Hoff Method: Hypothetical Limits of Solar Thermochemical Hydrogen. *J. Am. Chem. Soc.* **2024**, *146*, 14114−14127.
(57) Lany, S. Semiconducting transition metal oxides. *J. Phys.: Condens. Matter* **2015**, *27*, 283203.
(58) Zhang, L.; Liu, B.; Zhuang, H.; Kent, P.; Cooper, V. R.; Ganesh, P.; Xu, H. Oxygen vacancy diffusion in bulk SrTiO3 from density functional theory calculations. *Comput. Mater. Sci.* **2016**, *118*, 309−315.
(59) Capron, N.; Broqvist, P.; Pasquarello, A. Migration of oxygen vacancy in HfO2 and across the HfO2/SiO2 interface: A first-principles investigation. *Appl. Phys. Lett.* **2007**, *91*, 192905.
(60) Broberg, D.; Bystrom, K.; Srivastava, S.; Dahliah, D.; Williamson, B. A. D.; Weston, L.; Scanlon, D. O.; Rignanese, G. M.; Dwaraknath, S.; Varley, J.; Persson, K. A.; Asta, M.; Hautier, G. High-throughput calculations of charged point defect properties with semi-local density functional theory—performance benchmarks for materials screening applications. *npj Comput. Mater.* **2023**, *9*, 72.
(61) Shen, J. X.; Voss, L. F.; Varley, J. B. Simulating charged defects at database scale. *J. Appl. Phys.* **2024**, *135*, 145102.
(62) Kumagai, Y.; Tsunoda, N.; Takahashi, A.; Oba, F. Insights into oxygen vacancies from highthroughput first-principles calculations. *Physical Review Materials* **2021**, *5*, No. 123803.
(63) Choudhary, K.; DeCost, B.; Major, L.; Butler, K.; Thiyagalingam, J.; Tavazza, F. Unified graph neural network force-field for the periodic table: solid state applications. *Digit. Discovery* **2023**, *2*, 346−355.
(64) Deng, B.; Zhong, P.; Jun, K.; Riebesell, J.; Han, K.; Bartel, C. J.; Ceder, G. CHGNet as a pretrained universal neural network potential for chargeinformed atomistic modelling. *Nat. Mach. Intell.* **2023**, *5*, 1031−1041.
(65) Batatia, I. A foundation model for atomistic materials chemistry. *arXiv (Chem. Phys.)*, **2024**, 2024-03-01. DOI: . (accessed 2024-08-10).

https://doi.org/10.1021/acs.chemmater.5c00021
Chem. Mater. XXXX, XXX, XXX−XXX