

Machine Learning-Enabled Design of Point Defects in 2D Materials for Quantum and Neuromorphic Information Processing

Nathan C. Frey, Deji Akinwande, Deep Jariwala, and Vivek B. Shenoy*



Cite This: <https://dx.doi.org/10.1021/acsnano.0c05267>



Read Online

ACCESS |



Metrics & More



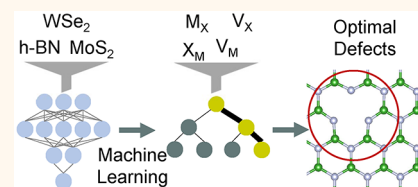
Article Recommendations



Supporting Information

ABSTRACT: Engineered point defects in two-dimensional (2D) materials offer an attractive platform for solid-state devices that exploit tailored optoelectronic, quantum emission, and resistive properties. Naturally occurring defects are also unavoidably important contributors to material properties and performance. The immense variety and complexity of possible defects make it challenging to experimentally control, probe, or understand atomic-scale defect-property relationships. Here, we develop an approach based on deep transfer learning, machine learning, and first-principles calculations to rapidly predict key properties of point defects in 2D materials. We use physics-informed featurization to generate a minimal description of defect structures and present a general picture of defects across materials systems. We identify over one hundred promising, unexplored dopant defect structures in layered metal chalcogenides, hexagonal nitrides, and metal halides. These defects are prime candidates for quantum emission, resistive switching, and neuromorphic computing.

KEYWORDS: machine learning, 2D materials, defects, DFT, quantum emission, resistive switching, neuromorphic computing



At the frontier of atomic-scale engineering, there is a major drive to construct artificial atoms in layered materials to simultaneously exploit the advantages of traditional cold atom systems and the quantum confinement effect in two dimensions. It is now possible to engineer two-dimensional (2D) materials with designer quantum properties, but it remains a tremendous challenge to control defects at the atomic level and manipulate local quantum states. Localized defects in transition metal dichalcogenides (TMDs)¹ and hexagonal boron nitride (h-BN)² have been shown to exhibit single-photon emission (SPE) at cryogenic and room temperatures, as well as resistive switching that enables memristor-based neuromorphic computing architectures.^{3–8} Point defects in 2D materials have advantages over traditional diamond color centers⁹ and memristive systems, namely, high tunability^{10,11} and easy integration into van der Waals heterostructures¹² with other solid-state devices.^{13–15} Defects in 2D materials are also extremely sensitive to local conditions, making them ideal for quantum sensing^{16,17} applications in nanoscale devices and biological systems. However, the space of possible defect structures and host materials is so vast and complex that experimental screening and exploration is a highly inefficient, slow, and nondeterministic approach to address this challenge. This presents an exciting opportunity to use computational tools to systematically explore the defect space, design optimal defects, and develop a more complete understanding of the underlying physics.^{18–20}

Recently, machine learning (ML) has emerged as a useful addition to the toolkit for designing and predicting properties of materials.^{21,22} Graph-based deep learning methods that account for lattice periodicity in crystals have shown promise in mapping bulk crystal structures directly to target properties when large data sets are available.^{23–25} Due to the smaller number of available 2D materials compared to bulk crystals and the resulting lack of data, there have been few studies of machine learning applied to 2D systems.^{26–29} Additional difficulties arise when trying to apply ML to predict quantum properties, which may be due to strong correlations between electrons that are difficult to capture with first-principles methods. Even greater challenges are involved in studying defects due to the exploding combinatorics in the defect search space, but this also provides opportunity to increase the amount of data available by orders of magnitude and use ML^{30–33} to obtain insight into the many outstanding questions about how imperfections affect material behavior.^{30,34–38} An understanding of defects across materials systems is necessary

Received: June 25, 2020

Accepted: September 8, 2020

Published: September 8, 2020



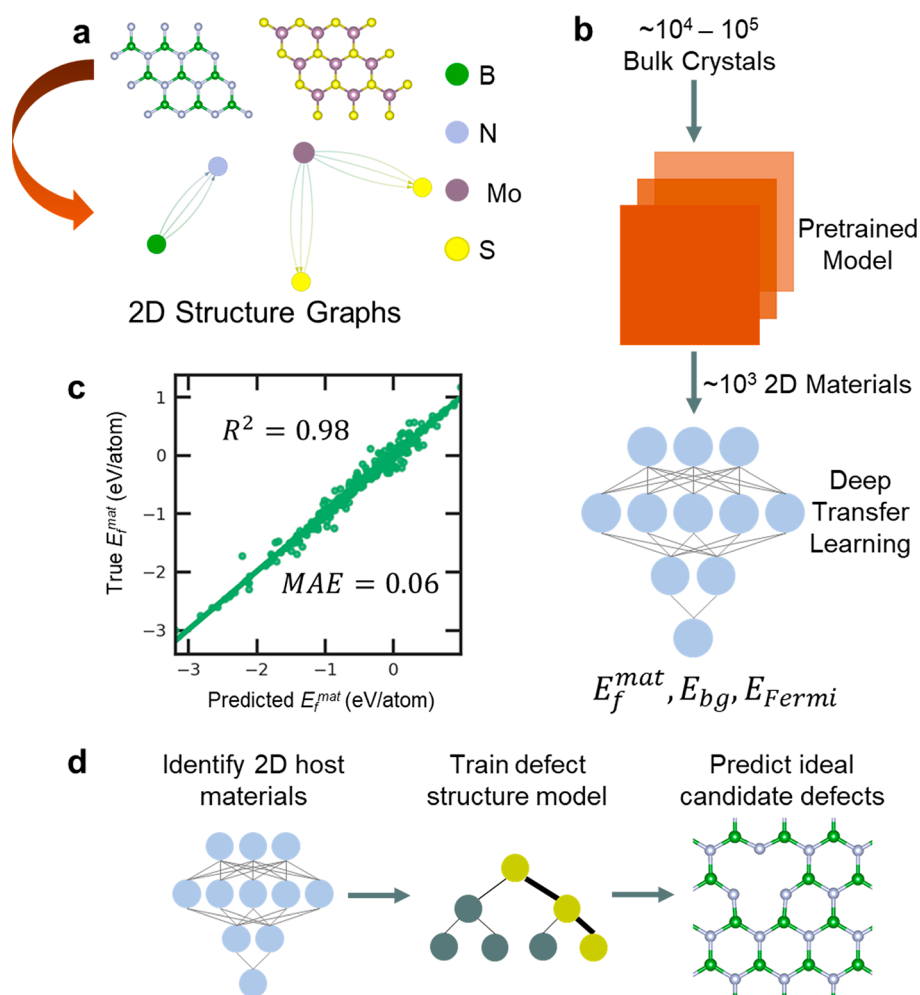


Figure 1. Deep learning property prediction for 2D materials. (a) 2D structures are mapped to structure graphs. (b) Graph network models are pretrained on large data sets of bulk crystals. Transfer learning is used to fine-tune model weights for predicting formation energies, band-gaps, and Fermi energies of 2D materials. (c) Parity plot of DFT calculated 2D material formation energies versus formation energies predicted with deep transfer learning for a test set of 381 materials. R^2 value and mean absolute error (MAE) are given. (d) Schematic of the entire workflow. Deep transfer learning is used to predict 2D host material properties and identify promising hosts, a random forest machine learning model is trained to predict defect structure properties, and finally ideal candidate defects are predicted.

not only for engineering artificial atoms to enable emerging technologies, but also because defects are inextricably linked to material properties and performance.

In this work, we synthesize recent advances in deep learning, machine learning, materials informatics, and *ab initio* materials design to systematically investigate hundreds of 2D materials (both van der Waals and non-van-der-Waals)^{39,40} and quickly identify the most promising defect structures for quantum emission and neuromorphic computing. Starting from nearly 4000 2D material structures, we use transfer learning^{41,42} to leverage models trained on tens of thousands of bulk crystal structures for deep-learning-powered predictions of 2D material properties that are critical to identifying promising defect host systems. After identifying the most promising 2D material hosts, we then generate nearly ten thousand defect structures in TMDs, h-BN, and over 150 2D wide band gap (WBG) materials. We develop a simple, physics-informed representation of defect structures in terms of easily accessible chemical and structural information, requiring no DFT calculations, and train ensemble ML models to create a predictive mapping between these representations and calculated defect properties. Band structures and defect

formation energies (energy difference between the pristine host structure and the dilute defect structure) are computed with density functional theory (DFT) for over 1000 defects across materials in the database to train and test the ML models. The ML approach enables prediction of fundamental defect properties without computationally expensive first-principles calculations. From the subset of considered defect structures, we identify 100 promising deep center defect structures for quantum emission and engineered dopant defects for resistive switching. The results provide ample opportunity for further study of promising candidates as well as a data-driven, physical understanding of defects in 2D materials that is needed to begin to resolve the complex microscopic origins of defect-mediated properties.

RESULTS

2D Material Property Prediction with Deep Transfer Learning. As the first step in designing ideal defect structures, we apply deep learning as a framework for 2D material property prediction and identify optimal host 2D materials. The Computational 2D Materials Database (C2DB)⁴³ contains nearly 4000 2D materials, and it would be highly

desirable to leverage this data to produce powerful, general models that can predict properties of arbitrary 2D materials. However, even 4000 materials are not sufficient to train data hungry deep neural networks (DNNs). To address the lack of data on 2D materials, transfer learning^{41,42} is used for property prediction by starting from networks trained on large data sets of bulk crystals. Here, we make use of graph networks⁴⁴ (a generalization of graph-based neural networks) as implemented in MatErials Graph Network (MEGNet)²⁴ models. Any material, for example, h-BN and MoS₂ (shown in Figure 1a), can be mapped to a graph representation characterized by the atomic numbers of the constituent elements and the spatial distance (bond lengths) between atoms. The graph network maps input graphs to outputs by “learning” the relationship between material structure and some target property. We use three MEGNet models constructed to predict formation energy, band gap, and Fermi energy, and one that classifies metals versus nonmetals. These models were trained on between 10⁴ and 10⁵ crystal structures from the Materials Project database⁴⁵ and as such, the model weights are already tuned to capture material properties. Furthermore, they contain elemental embeddings from the formation energy model using the largest data set (133 000+ materials) to encode chemical trends. Starting from these pretrained models, we fine-tune the model weights by training on the much smaller data set of 2D materials (~10³).

The transfer learning procedure (Figure 1b) enables rapid model training and accurate property prediction for 2D materials, simply by exploiting the learning process for the much larger data set of bulk crystal structures. The purpose of the 2D material property prediction models is to efficiently identify promising host materials, without requiring DFT calculations, whether they are present in a database or not. Not all 2D materials are suitable for quantum emission and resistive switching, so it is important to easily be able to screen candidate host materials for these applications. Predicted quantities such as the host band gap and formation energy are also important for ML predictions of defect properties. Figure 1c shows the parity plot for the test data of formation energy calculated by density functional theory (DFT) versus formation energy predicted by the graph network. The formation energy per atom of a material is given by

$$E_f^{\text{mat}} = \frac{E_{\text{total}} - \sum_i n_i \mu_i}{n} \quad (1)$$

where E_{total} is the total energy of a unit cell, n is the total number of atoms in a unit cell, and n_i and μ_i are the number and chemical potential (referenced to the most stable bulk unary phase) of the i th atomic species, respectively. $E_f^{\text{mat}} > 0$ eV/atom indicates that the material is thermodynamically unstable or metastable. The model achieves an impressive R^2 score of 0.98 and a mean absolute error (MAE) of 0.06 eV/atom on the test data (Table S1). Similar parity plots are given in Figure S1 for band gap (E_{bg}) and Fermi energy (E_{Fermi}) predictions. The metal versus nonmetal classifier (Table S2) has a test set accuracy of 0.84 and an F_1 score of 0.88 (0.73) for metals (nonmetals). The receiver operating characteristic (ROC) curve for nonmetal classification and the confusion matrix for predictions are shown in Figure S1d,e. The model performance statistics are given in Tables S1 and S2. The band gap model performs more poorly ($R^2 = 0.73$) than the others, while still achieving an MAE (0.36 eV) similar to that of the

band gap model trained on the bulk nonmetals (0.33 eV).²⁴ This is expected because the set of 2D nonmetals comprises only 28% (1067 of 3810 2D materials have nonzero calculated band-gaps) of the total data set. We note that it would be highly desirable to construct a single “multi-task” model with multiple outputs. In this work, we were limited by the availability of models pretrained on bulk crystal structures, and by significant differences in the training data sets. For example, the band gap regressor data sets for bulk and 2D crystals are smaller than the formation energy data set, because they include only nonmetals. These models enable rapid and accurate prediction of 2D material properties with deep learning, requiring no feature engineering or ML experience. The workflow presented in this paper is summarized in Figure 1d: deep transfer learning enables efficient prediction of critical host material properties (particularly band gap and formation energy) to identify promising hosts, a random forest machine learning model is trained to predict defect structure properties that are referenced to first-principles calculations, and finally ideal candidate defects are predicted.

Defect Database. Next, we narrow our focus to only those 2D materials that will make optimal hosts for engineered point defects. A good host material should have a wide band gap allowing for isolated deep defect levels and small spin–orbit coupling (SOC).⁴⁶ These conditions are satisfied by screening for nonmagnetic materials with band-gaps greater than 2 eV calculated with the GW approximation. This yields 158 potential WBG materials. We further screen out compounds with heavy elements as needed to reduce the effects of SOC. To identify promising defects in these systems, we establish some screening criteria and specify which are amenable to a high-throughput, machine learning-driven approach, and which require more individual study. By analogy to the NV[−] center in diamond, a deep defect center should have a paramagnetic qubit state with an energy splitting between two spin sublevels that is isolated from the bulk and other bound states.⁴⁶ Further details involving the magnitude of the two-level splitting and transitions between ground and excited states are important for assessing whether the state is optically addressable.^{46,47} For resistive memory applications, the defect state should be reversible, long-lived, and controllable with experimentally feasible switching voltages.^{4,5} By estimating the neutral defect formation energy and the position of defect states relative to the band edges, we evaluate hundreds of defects with machine learning and treat the other criteria as engineerable perturbations. This method has been validated by previous studies^{31,32} on point defects in bulk crystals.

To clarify the approach, we describe the defect state by a simple model Hamiltonian

$$H_{\text{defect}} = H_0 + H_{\text{sb}} + H_{\text{shift}} \quad (2)$$

H_0 includes a minimal description of the electrically neutral defect, namely the formation energy (Figure 2a) and the position of defect levels relative to the conduction and valence bands (Figure 2c). H_{sb} contains symmetry breaking terms (crystal field splitting, spin–orbit coupling, Jahn–Teller distortion, uniaxial strain, etc.) that split degenerate levels. H_{shift} includes terms that preserve the crystal symmetry (biaxial strain, small applied fields, doping) but tune the energy levels and charge states of defect levels in the gap (Figure 2d). All the host structures considered in this work have either D_{3h} (h-BN, AlN), D_{3d} (MgI₂), or C_{3v} (GeS) point group symmetry (neglecting large structural distortions). The point group

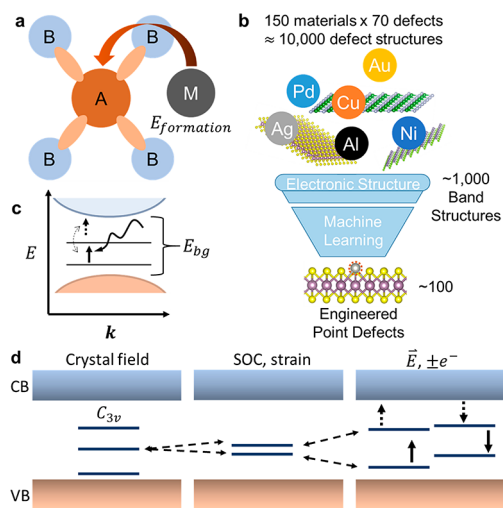


Figure 2. Machine learning a description of point defects in 2D wide band gap semiconductors. (a) Schematic of formation energy for a substitutional defect species, M , replacing an A atom in an A - B binary system. (b) An average of 70 point defects are considered for 150 materials to generate nearly 10 000 defect structures. More than 1000 band structures are computed to train machine learning models, leading to the identification of 100 promising defects. (c) Deep defect center located in the gap acts as a two-level system. (d) Diagram of an engineered two-level system from a deep center defect. Solid (dashed) up or down arrows indicate occupied (unoccupied) spin states.

symmetry immediately gives a first-order description of the defect energy level splitting due to the crystal field. The process of taking a proposed deep center defect and engineering an ideal two-level system is shown in Figure 2d. For example, dopants or vacancies with C_{3v} symmetry have levels split into three irreducible representations denoted a_1 , a_2 , and e . These levels are further split by other symmetry breaking effects, such as SOC and external stimuli such as applied strain, which generate a two-level system with an optically accessible transition. Applied fields and doping will then tune the position of the levels within the gap and the Fermi level. Complexing the dopants with vacancies is one particularly robust strategy to engineer isolated two-level systems.^{35,47} In this work, we focus on characterizing H_0 to obtain statistics, trends, and understanding of point defects across many materials systems with machine learning and without prohibitively expensive hybrid functional calculations. The most promising candidates can then be further studied with higher levels of theory to determine and design the perturbative effects of the symmetry breaking and energy shifting terms.^{10,12,35,47}

With these considerations in mind, we generated over 10 000 defects in the WBG materials by considering all possible vacancies, divacancies, antisites, and common dopants (listed in Table S3). Of these, we computed relaxed defect geometries and band structures for over 1000 quantum point defects (QPDs) and for 140 substitutional metal defects (Table S4) in the atomically thin resistive memory materials MX_2 ($M = \text{Mo}, \text{W}; X = \text{S}, \text{Se}, \text{Te}$) and h-BN. Some representative atomistic structures are shown for defects in h-BN in Figure S2. Figure 2b shows a schematic of the process: the candidate defects are funneled into a subset for electronic structure calculations, which are then used to test ML models

for defect property prediction, and finally to predict ideal defect structures.

Machine Learning Prediction of Defect Properties.

Because detailed defect calculations at a high level of theory are computationally expensive and necessarily low-throughput, here we develop an ML approach to predict defect formation energy and energy-level position *without requiring any DFT calculations*. Figure 3 shows the ML workflow developed in this paper. For any WBG material, first the host material parameters are obtained (Figure 3a). The host material is well-described by automatically generated structural and chemical descriptors^{48–50} and calculated electronic properties such as the band-gaps at the Perdew–Burke–Ernzerhof (PBE),⁵¹ Heyd–Scuseria–Ernzerhof (HSE),⁵² and GW approximation⁵³ levels of theory, which are available in the C2DB or from transfer learning. Because of the scarcity of defect data and the large differences in target properties and defect structure graphs compared to that of pristine host materials, the graph-based approach used above is not tractable here. Instead, defects are described by structural and chemical properties and by percent changes in properties compared to the pristine bulk. For example, one feature may be the mean atomic radius in the structure, $\bar{r}_d$, where the d subscript denotes a defect structure. The defect is described both by $\bar{r}_d$ and by $\Delta\bar{r} = \frac{\bar{r}_d - \bar{r}_b}{\bar{r}_b}$, where $\bar{r}_b$ is the mean atomic radius in the

pristine 2D material. For an Sn_{Se} substitution (Sn occupying an Se site), $\Delta\bar{r} > 0$, while $\Delta\bar{\chi} < 0$ for mean electronegativity. Using percent differences as features, rather than absolute values, facilitates comparing defect structures across material systems that may have significantly different values. This is a form of feature normalization, which is standard in machine learning applications,⁵⁰ although it is not strictly necessary for random forest models.⁵⁴ There are also features that are not averaged over the entire supercell and only relate to local differences at the defect site. This is a simple representation for defect structures that is easily interpretable, well-suited as input to ML models, and requires no first-principles calculations.

The ML approach is broken up into two models: a classifier to predict the existence of deep center defects and a regressor to predict defect formation energies. The existence of deep centers is determined by computing the energy differences between the defect level and the valence band maximum and conduction band minimum, denoted ΔVB and ΔCB (Figure 3b), respectively. For simplicity, a defect is labeled a deep level if $\Delta\text{CB} > k_B T$ and $\Delta\text{VB} > k_B T$ at room temperature ($k_B T \approx 25$ meV). Otherwise, it is a shallow level which is either susceptible to thermal excitations or resonant with the bulk bands. Using this $k_B T$ threshold, 442 (roughly 35%) of the computed QPDs exhibit deep centers. Figure 3c shows the histogram of defect formation energy, E_f (defined in eq 5 in the Methods), values for the QPDs. These computed band structures and E_f values determine the targets for ML prediction, but no first-principles defect calculation data are used as input features.

Before discussing the particulars of the ML model performance, we briefly review the mechanisms at work in the defect that determine the energetics. Understanding the physics of defect formation will inform the feature generation and model training during the ML process. Figure 4a is a schematic that shows the interactions between some dopant atom, M , and the four atoms (blue spheres) to which it is bonded. On the left side, there are the strain-dependent terms

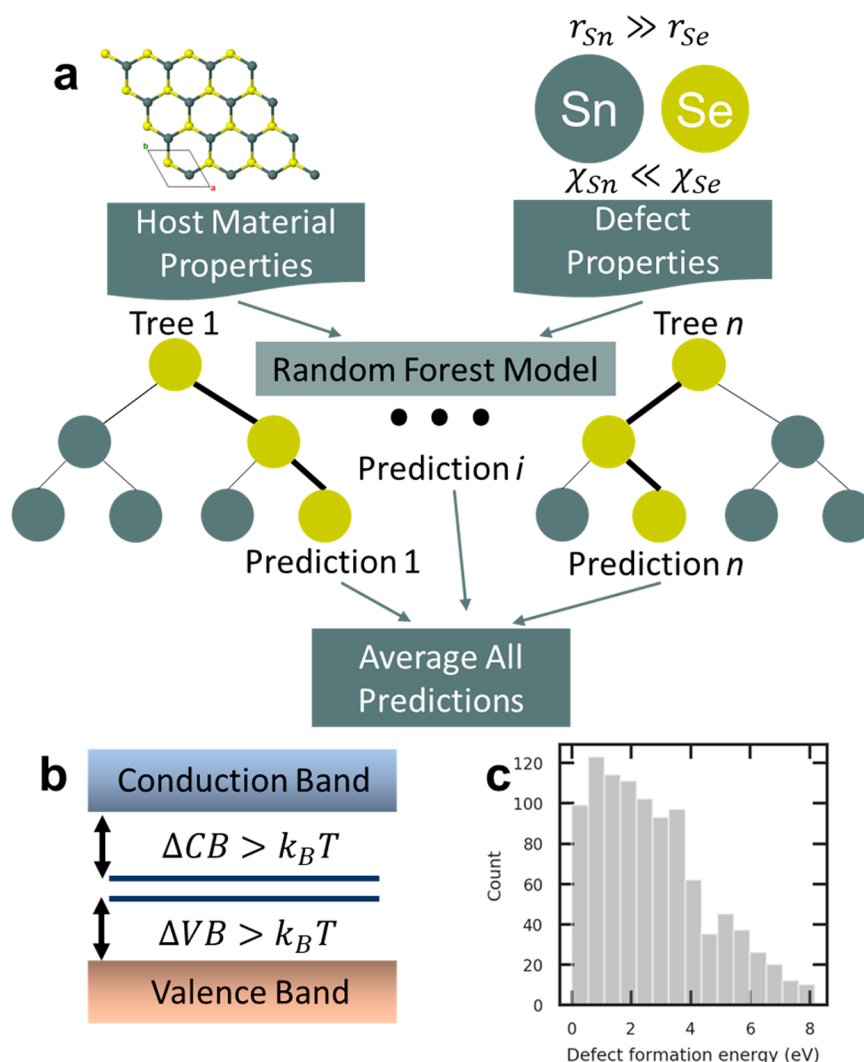


Figure 3. Machine learning model prediction of defect properties. (a) Host material and defect properties that are easily available from online databases or deep transfer learning model predictions are used as inputs to a random forest model. Decision trees in the random forest make predictions that are averaged to generate final predictions. (b) Deep defect center classifier predicts whether defect levels are energetically separated from the valence and conduction bands by at least $k_B T \approx 25$ meV. (c) Distribution of computed defect formation energies.

that include an elastic dipole contribution that is linear in the strain, ϵ , and a strain energy that goes as ϵ^2 . The strain is spontaneously induced to relax the geometry and lower the energy. On the right side, there are terms arising from electron–electron interactions such as Coulomb repulsion and charge transfer that account for the breaking and reforming of bonds, which raises the energy. This simple picture suggests that a minimal description of defects is realized by accounting for local relaxation (strain) and electronegativity (electrostatics).⁵⁵

By constructing feature vectors for defects that encode information about local relaxation and electronic interactions, the ML models are able to classify deep centers and predict E_f . In both cases we use random forest (RF) ensemble models for the benefits in interpretability, performance, and robustness to overfitting.⁵⁴ We split 90% of the data into a training set, making use of bootstrapping⁵⁶ to generate an out-of-bag (OOB) score for validation, and 10% of the data is held as a test set. The final trained RF model for E_f prediction has an OOB score of 0.75 and an R^2 of 0.74 on the test set (Figure 4b). The MAE on the test set is 0.67 eV, which is small

considering that the range of computed E_f values is from 0 to 8 eV. We use the permutation feature importance (Figure 4c) to rigorously inspect how the model is working.⁵⁴ Looking at only the top 10 most important features, there are descriptors clearly related to local relaxation that encode how the defect will induce compression or tension in the lattice, for example, the change in atomic weight of the defect compared to the host atom and the change in van der Waals radius. On the other hand, the change in mean number of p valence electrons and the change in electronegativity relate to electrostatics and bonding. Finally, among the most important features, the chemical potential of the defect species (available from the Materials Project database)⁴⁵ is directly related to E_f .

The model performance is even better for the deep center classifier, with $F_1 = 0.92$ on the test set for classifying deep level defects (Figure S3). The most important features are more directly related to band structure, with the most important being the lowest unoccupied molecular orbital (LUMO) energy. Other features related to changes in electronegativity and column position (electron count) of constituent elements are also weighted heavily.

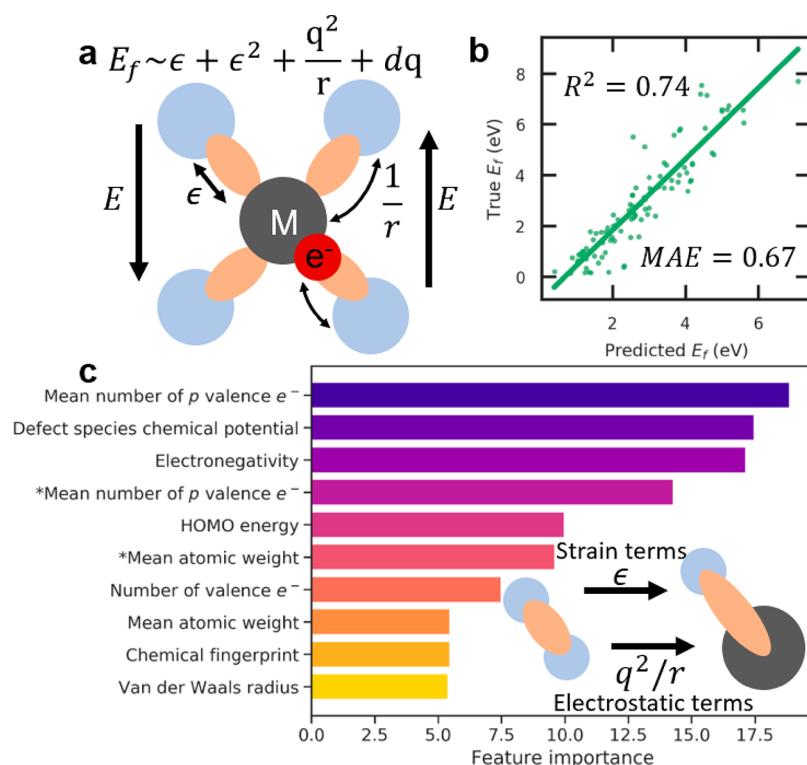


Figure 4. Physical mechanisms underlying defect energetics, model featurization, and model performance. (a) Schematic of a dopant, M , inducing strain and electrostatic interactions that govern formation energy. (b) Parity plot of DFT computed defect formation energy versus predicted formation energy with R^2 value and mean absolute error (MAE). (c) Permutation feature importances in formation energy model. An asterisk (*) denotes the feature is derived from the pristine or defect structure. All other features are computed as differences in values between the defect and pristine structures.

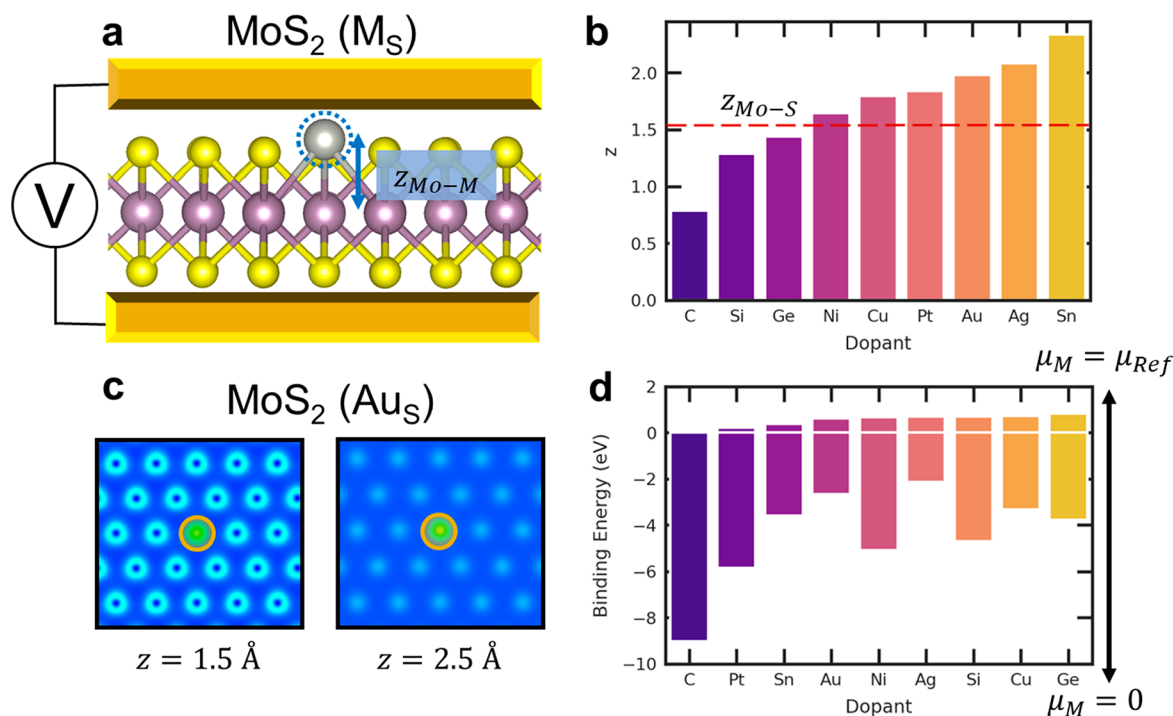


Figure 5. Engineered defects in atomically thin resistive memory devices. (a) Schematic of a device with a semiconducting MoS_2 layer sandwiched between two metallic electrodes. The M_S substitutional defect controls the local resistivity. (b) Calculated distances, z , between different dopants and the Mo plane in MoS_2 (M_S). Colors correspond to magnitude of z value. (c) Cross sections of the charge density in MoS_2 (Au_S). The Au dopant atom is circled in gold. (d) Calculated binding energies for dopants in MoS_2 as a function of the dopant chemical potential. Colors correspond to magnitude of maximum μ_M value.

Importantly, the linear Pearson correlations of all these individual features with E_f and the defect energy level position is quite low (<0.3); in other words, although our physical intuition guides the feature engineering process, it is not possible to construct a simple model to predict defect properties. Linear models such as Lasso⁵⁷ and Ridge regression⁵⁸ fail at predicting E_f , whereas the nonlinear RF performs much better. Despite the complex, nonlinear nature of the models, we can extract mechanistic insight by considering the permutation feature importance and the mean and median values of the most important features when defects are grouped by formation energy or energy level position. If we classify defects as either high or low formation energy based on whether E_f is greater than or less than the median formation energy (2.44 eV), it is possible to quantify the contributions from stress-induced and electronic interactions described above. We find that high energy defects have, on average, 23% larger changes in electronegativity compared to low energy defects. High energy defects also have a 7% larger median chemical similarity value (eq 7 in the Methods) compared to low energy defects, indicating that high energy defect structures are associated with larger distortions. Likewise, deep center defects have a median change in the atomic number that is 39% higher than shallow defect states. The interplay between local relaxation and electronic interactions that determines defect energetics is not trivial, but the RF model is able to capture the defect physics we aimed to describe. The model results support our simple picture that structural and electronic distortions induce deep level defects,^{55,59,60} but large local distortions also raise the formation energy.

Metallic Dopants in Atomically Thin Resistive Memory Materials. Next, we will discuss a subset of engineered defects in TMDs and h-BN that are of importance for their applications in nonvolatile resistance switching (NVRs). These so-called “atomristors” comprise a semi-conducting monolayer sandwiched between two electrodes. A schematic of this device geometry is shown in Figure 5a. A voltage is applied across the electrodes to induce a metal atom, M, from the electrode to hop into a naturally occurring vacancy in the semiconductor, for example, a sulfur vacancy in MoS₂, forming a substitution M_S. Reversing the voltage causes the vacancy to form again. This gives a voltage-controlled local resistivity that forms a kind of memory in an atomically thin device. The defect geometry can be probed *via* scanning tunneling microscopy (STM) to measure the out-of-plane distance $z_{\text{TM-M}}$ between the defect and the plane of the transition metal (TM) atom in the TMD or the plane in h-BN. We define $\Delta z = (z_{\text{TM-M}} - z_{\text{TM-X}})/z_{\text{TM-X}}$ as the change in out-of-plane distance relative to the equilibrium distance between the TM plane and the chalcogen plane; this value represents the local strain and bond breaking/formation, although it is not perfectly correlated with the total binding energy. Because the dopants considered here always preserve the point group symmetry of the material (there is no in-plane displacement of the dopant), and because the out-of-plane displacement can be probed with STM,⁶¹ we have considered Δz as the primary indicator of local strain. With the M atom electronegativity and atomic orbitals, we know from our ML results above that this information can be used to effectively predict the defect properties. In particular, we will focus on neutral and acceptor defects (chalcogen vacancies) with a metallic cation dopant, where electrically neutral substitutions are of interest.

For this subset of defects, we calculate Δz for every combination of host material and metallic dopant (listed in Table S4). Among the dopants we also consider C, Si, and Ge, which exist in few-layer semimetallic phases that could be used as electrodes in a completely van der Waals resistive device. The height profiles for all dopants in MoS₂ are given in Figure 5b, where the dashed red line indicates the equilibrium z between the Mo and S planes. Δz increases linearly with the increasing atomic radius of the dopant (Figure S4). We plot cross sections of the calculated charge density at varying distances from the Mo plane to visualize the defect height profile, as shown for Au_S in MoS₂ in Figure 5c. As z increases, charge density localized on the Au dopant (highlighted with a gold circle in Figure 5c) remains visible, while the charge density localized on neighboring S atoms decreases.

We define the defect binding energy E_{BE} of a metallic dopant, M, as

$$E_{\text{BE}}(\mu_{\text{M}}) = E_{\text{total}} - E_{\text{V}} - \mu_{\text{M}} \quad (3)$$

where E_{total} is the total energy of the supercell with the dopant, E_{V} is the energy of the supercell with the vacancy, and μ_{M} is the chemical potential of the dopant. In experiments, it is difficult to control the dopant chemical potential, so in Figure 5d we plot $E_{\text{BE}}(\mu_{\text{M}})$ for each defect in the MoS₂ for the range from $\mu_{\text{M}} = 0$ to $\mu_{\text{M}} = \mu_{\text{ref}}$ where μ_{ref} is the bulk reference value of the most stable unary phase. $E_{\text{BE}} < 0$ indicates that the dopant defect will occur spontaneously in the presence of the vacancy. $E_{\text{BE}}(\mu_{\text{M}})$ corresponds to the switching voltage (neglecting kinetic barriers), and the plot in Figure 5d shows that by varying the M atom species, a wide range of switching voltages and dopant stabilities can be achieved to serve diverse switching applications from information storage to neuromorphic computing. The binding energy is correlated with the dopant species' atomic radius (or equivalently, Δz), as shown in Figure S5a, where the R^2 value is 0.51. However, the dopant atomic radius does not completely represent bond breaking and formation, so that the binding energy cannot be directly inferred from the dopant atomic radius. We note that the calculated binding energies for MoS₂ with Au and Ag dopants agree well with experimental measurements⁴ of ~ 1 V switching voltages.

Optimal Defect Candidates. To summarize the results of our analysis, we introduce a simple “defect score” metric that succinctly represents a candidate defect's fitness as a potential deep center for quantum emission. The defect score is defined as

$$S = \frac{1}{N} \left(E_{\text{bg}}^{\text{GW}} + \left(\frac{1}{2} s_{\text{d}} + \frac{1}{2} s_{\text{t}} \right) - A - E_{\text{f}} \right) \quad (4)$$

where $E_{\text{bg}}^{\text{GW}}$, s_{d} , s_{t} , A , and E_{f} are the GW band gap, dynamic stability, thermodynamic stability, maximum atomic number in the host (corresponding to SOC), and defect formation energy, respectively. N is an overall normalization factor. Higher scores reflect larger band-gaps, greater stability, smaller defect formation energies, and smaller SOC (smaller maximum atomic number). Although there are not many examples of experimentally verified WBG material quantum emitters, we have compared the average defect scores for three materials (h-BN, h-AlN, and GaN) that have been shown to exhibit quantum emission in either monolayer or crystalline form (Figure S6). We find that the average defect score decreases with reported inverse emission lifetime (τ^{-1}), that is, a higher

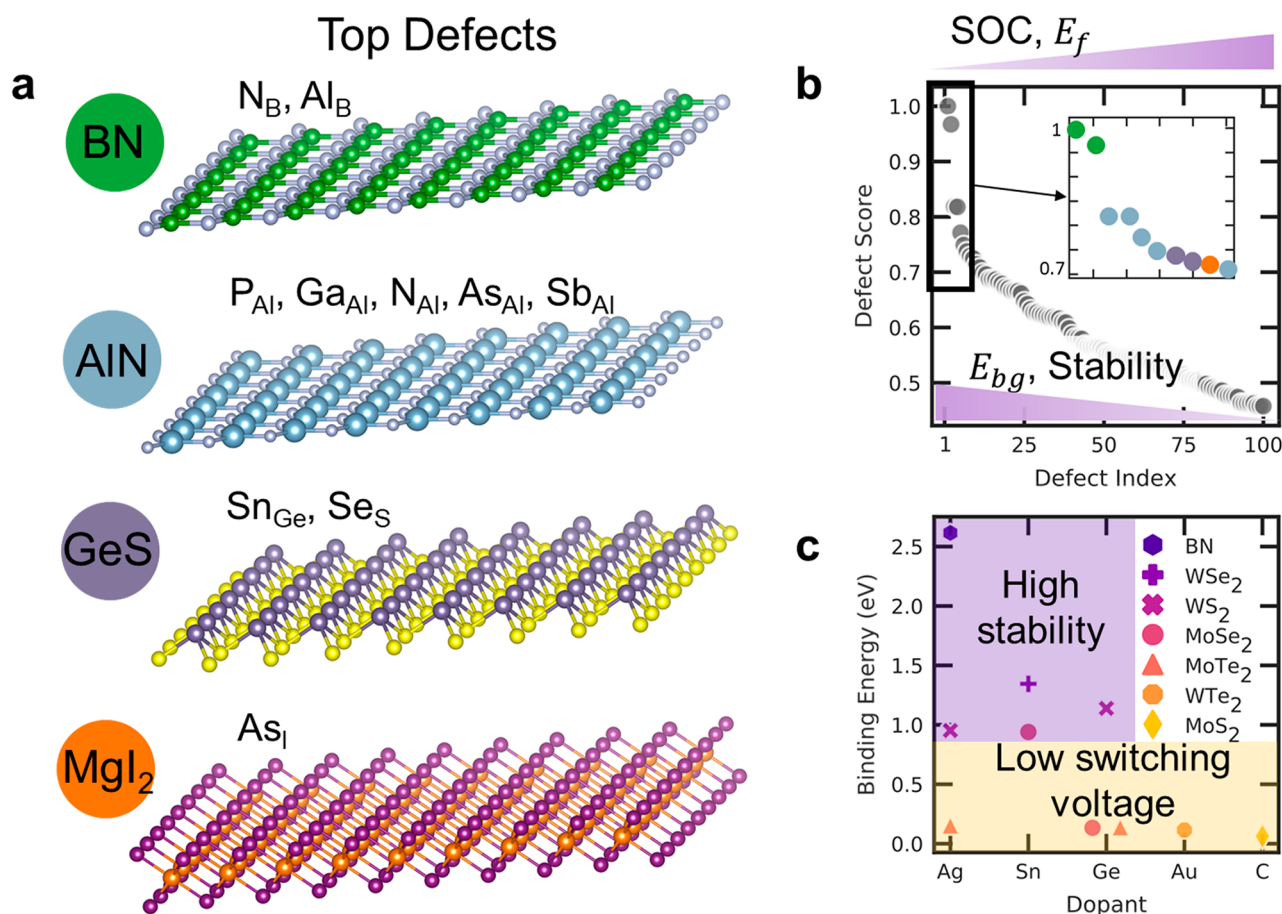


Figure 6. Top identified defect candidates. (a) Ten substitutional defects in h-BN, AlN, GeS, and MgI₂ with the highest defect scores. (b) Plot of defect scores for top 100 defects. The top 10 are highlighted in the inset and color-coded to their corresponding host materials. (c) Highest five and lowest five defects for resistive switching by maximum binding energy. The light purple shaded region indicates high stability defects, while the light yellow shaded region indicates defects with a low switching voltage.

average defect score corresponds to higher emission lifetime. This could be due to materials with larger band-gaps and low E_f deep center defects (higher average defect score) having longer-lived two-level defect states.⁶²

The top 10 dopant defects are shown in Figure 6a. Defects in h-BN score highly, as expected because of its ideal WBG material properties, but we also find optimal defect candidates in AlN, GeS, and MgI₂ (Table S5). Notably, room temperature quantum emission was recently shown in bulk AlN.⁶³ The top 100 defect scores are plotted in Figure 6b, with the top 10 highlighted in the inset. The scores (Table S6) appear Pareto distributed, emphasizing the challenge in identifying promising defect candidates compared to the relatively more abundant, less promising defects. These dopants can be complexed with vacancies to engineer symmetry breaking and construct two-level systems.^{35,47} Other than h-BN, these systems are relatively or completely unexplored for quantum emission applications and are prime candidates for further study.

Finally, we highlight the optimal defect candidates for resistive switching. Figure 6c shows the highest five and lowest five defects by maximum binding energy, $E_{BE}(\mu_M = \mu_{ref})$ (listed in Table S7). The highest binding energy defects are of interest in memory applications for their assumed stability, while the lower binding energy defects require small switching voltages useful for neuromorphic architectures. We have also listed the five defect structures with binding energies in an intermediary regime (between 0.60 and 0.49 eV), such that they strike a

balance between stability and switching voltage. In general, we find that M_X defects in TMDs with larger Δr differences in atomic radii between dopant atoms, M, and chalcogen atoms, X, have larger binding energies (Figure S5a). TMDs with smaller band-gaps and therefore weaker bonds (larger in-plane lattice constants) have lower binding energies for metallic dopants (Figure S5b). This suggests that the unexplored MTe₂ (M = Mo, W) systems are optimal low voltage resistive switching materials, similar to their bulk counterparts.⁶⁴ We also find a number of defects with high values of E_{bg}/E_{BE} (Figure S7), which may form easily and simultaneously offer high ON/OFF current ratios due to the large band gap.⁵ Relatively few experiments^{3–5,61} have been done on 2D monolayer resistive switching, hence our results here provide ample opportunity to explore more optimal host/dopant architectures that span a large range of accessible switching voltages.

CONCLUSIONS

In this work, we have systematically investigated and identified optimal point defects in 2D materials using a combination of deep transfer learning, machine learning, and first-principles calculations. We have leveraged graph networks trained on tens of thousands of bulk crystal structures to enable deep learning for predicting formation energies, Fermi energies, and band-gaps in 2D materials. Nearly 10 000 defect structures were

constructed from over 150 wide band gap semiconductors and layered metal chalcogenides. Band structures and formation energies were calculated for over 1000 of these defects and used to test ensemble machine learning models based on physics-informed featurization. The models used easily accessible descriptors, requiring no electronic structure calculations, to encode information about local relaxation and electronic interactions that captures defect physics. The resulting models were able to predict key defect properties including formation energies and the position of defect levels relative to the valence and conduction bands. We identified the 100 most promising deep center defects for quantum emission applications and 10 optimal defects for nonvolatile resistive switching in atomically thin memristor devices. The systematic, machine learning-enabled exploration of 2D materials and defect structures presented here provides a holistic and simple physical picture of defects that complements the many detailed studies of individual defects in specific materials. By identifying the most promising defect structures and revealing the microscopic strain and electronic mechanisms that govern their behavior, we accelerate the realization of designer artificial atoms on an atomic scaffold for quantum and neuromorphic technologies.

METHODS

Defect Database and First-Principles Calculations. A total of 3810 materials with their corresponding dynamic and thermodynamic stabilities, formation energies, Fermi energies, and band-gaps were obtained from the Computational 2D Materials Database.⁴³ Of those, 158 candidate WBGSC 2D materials were identified by screening for nonmagnetic materials with computed GW band-gaps greater than 2 eV. This yielded 1 unary, 130 binary, and 27 ternary compounds. All DFT calculations were performed with the Vienna *Ab-Initio* Simulation Package,⁶⁵ projector augmented wave (PAW) pseudopotentials,⁶⁶ and the Perdew–Burke–Ernzerhof (PBE)⁵¹ exchange–correlation functional. Host material properties were computed starting from structures in the C2DB and then relaxing them until forces were converged to below 10^{-2} eV/Å with a $10 \times 10 \times 1$ Γ -centered k -point grid and a 520 eV plane wave basis cutoff. For total energy calculations, the k -point grid density was increased to $18 \times 18 \times 1$ and the total energy was converged to 10^{-8} eV. Defect structures were generated with the *pymatgen* code,⁶⁷ considering all possible vacancies, divacancies, antisites, and common semiconductor dopants given in Table S3. All defect structures were modeled as $5 \times 5 \times 1$ supercells to minimize interactions between periodic images of defects.^{47,68} Plane wave basis cutoffs were taken to be the largest cutoff of the pseudopotentials for the constituent elements. With the lattice parameters fixed, ionic positions were relaxed using a $3 \times 3 \times 1$ k -point grid until forces were converged to below 10^{-2} eV/Å. For supercell band structure calculations, 1280 defect structure calculations completed with total energies converged to 10^{-4} eV. Also, 140 additional point defect structures for resistive switching were calculated by considering the metallic dopants listed in Table S4 and the host materials MX_2 ($M = \text{Mo}, \text{W}$; $X = \text{S}, \text{Se}, \text{Te}$) and h-BN. The structure graphs and defect structures are available on figshare at https://figshare.com/collections/Defect_design/4946874.

Neutral defect formation energies^{31,32} were calculated as

$$E_f = E_{\text{total}} - E_{\text{bulk}} - \sum_i n_i \mu_i \quad (5)$$

where E_{total} is the total energy of the defect supercell, E_{bulk} is the total energy of the corresponding pristine monolayer, and the sum is over all atoms that are added ($n_i > 0$) or removed ($n_i < 0$), with n_i and μ_i being the number and chemical potential of species i , respectively.

Deep Learning. Pretrained models were obtained from MEGNet.²⁴ MEGNet models are built using the *keras*⁶⁹ API and *TensorFlow*⁷⁰ backend. All materials in the C2DB were split into an

80:10:10 training, validation, test set. For the band gap regressor, the total data set was reduced to the 1078 nonmetals. Models were trained on 2D materials data until no further improvement was seen in the training and validation losses. Structure graphs were generated using the graph convertors in the pretrained MEGNet models. The models were trained on Nvidia Tesla P100 GPUs. The formation energy neural network model architecture is shown as an example in Figure S1a. The trained graph network models and an example Jupyter notebook are available on GitHub at <https://github.com/ncfrey/defect-design>.

Machine Learning. Machine learning models were implemented with *scikit-learn*,⁷¹ and some descriptors (as defined in Meredig *et al.*⁴⁹) were generated with *matminer*⁴⁸ and *automatminer*.⁵⁰ Defect descriptors were constructed as percent differences between the corresponding descriptor for the bulk structure and the unrelaxed defect structure,

$$f_i = \frac{f_{\text{defect}} - f_{\text{bulk}}}{f_{\text{bulk}}} \quad (6)$$

where f_{defect} and f_{bulk} are values of a feature, f , for defect and bulk structures, respectively. f may be any available feature, such as the average electronegativity of atoms in the structure. No calculations are required to generate the descriptors; only the chemical and structural information on the host and defect structures is needed. We introduce the “chemical similarity fingerprint,” d , defined as

$$d = |v_{\text{bulk}} - v_{\text{defect}}| \quad (7)$$

where $v = (w_1, w_2, \dots, w_n)$ is a vector of all chemical descriptors, w , for a particular monolayer or one of its defect structures. This metric succinctly represents the similarity of a defect structure to its pristine host and the corresponding magnitude of structural distortion that may be induced in the defect structure. For example, MgI with a Br_Γ defect has $d \approx 1$ (indicating a small distortion), whereas GaN with a Bi_N defect has $d \approx 50$ (indicating a large distortion). Machine learning model validation statistics are presented in Table S8.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsnano.0c05267>.

Transfer learning regression model validation in Table S1, transfer learning classification model validation in Table S2, list of considered semiconductor dopants in Table S3, list of dopants considered for resistive switching applications in Table S4, 2D wide band gap semiconductors in Table S5, top 100 deep center point defects and defect scores in Table S6, highest, lowest, and intermediary five defects by binding energy for resistive switching in Table S7, machine learning defect model validation in Table S8, transfer learning graph neural network results in Figure S1, schematics of defect types in Figure S2, permutation feature importance for deep center defect classifier in Figure S3, plot of change in distance versus atomic radius of dopant in MoS_2 in Figure S4, M_X defects in transition metal dichalcogenides in Figure S5, average defect score versus reported inverse lifetimes for single photon emission in Figure S6, plot of host band gap divided by defect binding energy in Figure S7 (PDF)

AUTHOR INFORMATION

Corresponding Author

Vivek B. Shenoy – Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia,

Pennsylvania 19104, United States; Email: vshenoy@seas.upenn.edu

Authors

Nathan C. Frey – Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States; orcid.org/0000-0001-5291-6131

Deji Akinwande – Microelectronics Research Center, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas 78758, United States

Deep Jariwala – Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States; orcid.org/0000-0002-3570-8768

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acsnano.0c05267>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported primarily by contract W911NF-16-1-0447 from the Army Research Office (V.B.S.) and also by Grants EFMA-542879 and CMMI-1727717 from the U.S. National Science Foundation. N.C.F. was supported by the Department of Defense through the National Defense Science & Engineering Graduate Fellowship program. N.C.F. would like to thank C.C. Price for fruitful discussions. D.A. acknowledges the support of the Presidential Early Career Award for Scientists and Engineers (PECASE) through the Army Research Office (W911NF-16-1-0277) and an NSF grant (ECCS-1809017). D.J. acknowledges support for this work by the U.S. Army Research Office under Contract No. W911NF-19-1-0109 and NSF DMR-1905853 and DMR-1720530.

REFERENCES

- (1) Mak, K. F.; Shan, J. Photonics and Optoelectronics of 2D Semiconductor Transition Metal Dichalcogenides. *Nat. Photonics* **2016**, *10*, 216–226.
- (2) Tran, T. T.; Bray, K.; Ford, M. J.; Toth, M.; Aharonovich, I. Quantum Emission from Hexagonal Boron Nitride Monolayers. *Nat. Nanotechnol.* **2016**, *11*, 37–41.
- (3) Kim, M.; Ge, R.; Wu, X.; Lan, X.; Tice, J.; Lee, J. C.; Akinwande, D. Zero-Static Power Radio-Frequency Switches Based on MoS₂ Atomistors. *Nat. Commun.* **2018**, *9*, 2524.
- (4) Ge, R.; Wu, X.; Kim, M.; Shi, J.; Sonde, S.; Tao, L.; Zhang, Y.; Lee, J. C.; Akinwande, D. Atomistor: Nonvolatile Resistance Switching in Atomic Sheets of Transition Metal Dichalcogenides. *Nano Lett.* **2018**, *18*, 434–441.
- (5) Wu, X.; Ge, R.; Chen, P.; Chou, H.; Zhang, Z.; Zhang, Y.; Banerjee, S.; Chiang, M.; Lee, J. C.; Akinwande, D. Thinnest Nonvolatile Memory Based on Monolayer H-BN. *Adv. Mater.* **2019**, *31*, 1806790.
- (6) Sangwan, V. K.; Lee, H. S.; Bergeron, H.; Balla, I.; Beck, M. E.; Chen, K. S.; Hersam, M. C. Multi-Terminal Memtransistors from Polycrystalline Monolayer Molybdenum Disulfide. *Nature* **2018**, *554*, 500–504.
- (7) Zhao, H.; Dong, Z.; Tian, H.; DiMarzi, D.; Han, M.-G.; Zhang, L.; Yan, X.; Liu, F.; Shen, L.; Han, S.-J.; Cronin, S.; Wu, W.; Tice, J.; Guo, J.; Wang, H. Atomically Thin Femtojoule Memristive Device. *Adv. Mater.* **2017**, *29*, 1703232.
- (8) Wang, C.; Wang, C.; Meng, F.; Wang, P.; Wang, S.; Liang, S.; Miao, F. 2D Layered Materials for Memristive and Neuromorphic Applications. *Adv. Electron. Mater.* **2020**, *6*, 1901107.
- (9) Aharonovich, I.; Greentree, A. D.; Prawer, S. Diamond Photonics. *Nat. Photonics* **2011**, *5*, 397–405.
- (10) Noh, J. Y.; Kim, H.; Park, M.; Kim, Y. S. Deep-To-Shallow Level Transition of Re and Nb Dopants in Monolayer MoS₂ with Dielectric Environments. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *92*, 115431.
- (11) Wang, D.; Han, D.; Li, X. B.; Xie, S. Y.; Chen, N. K.; Tian, W. Q.; West, D.; Sun, H. B.; Zhang, S. B. Determination of Formation and Ionization Energies of Charged Defects in Two-Dimensional Materials. *Phys. Rev. Lett.* **2015**, *114*, No. 196801, DOI: [10.1103/PhysRevLett.114.196801](https://doi.org/10.1103/PhysRevLett.114.196801).
- (12) Singh, A.; Manjanath, A.; Singh, A. K. Engineering Defect Transition-Levels through the van der Waals Heterostructure. *J. Phys. Chem. C* **2018**, *122*, 24475–24480.
- (13) Jariwala, D.; Sangwan, V. K.; Lauhon, L. J.; Marks, T. J.; Hersam, M. C. Emerging Device Applications for Semiconducting Two-Dimensional Transition Metal Dichalcogenides. *ACS Nano* **2014**, *8*, 1102–1120.
- (14) Akinwande, D.; Huyghebaert, C.; Wang, C. H.; Serna, M. I.; Goossens, S.; Li, L. J.; Wong, H. S. P.; Koppens, F. H. L. Graphene and Two-Dimensional Materials for Silicon Technology. *Nature* **2019**, *573*, 507–518.
- (15) Atatüre, M.; Englund, D.; Vamivakas, N.; Lee, S. Y.; Wrachtrup, J. Material Platforms for Spin-Based Photonic Quantum Technologies. *Nat. Rev. Mater.* **2018**, *3*, 38–51.
- (16) Degen, C. L.; Reinhard, F.; Cappellaro, P. Quantum Sensing. *Rev. Mod. Phys.* **2017**, *89*, 035002.
- (17) Alkauskas, A.; Basset, L. C.; Exarhos, A. L.; Fu, K. M. C. Defects by Design: Quantum Nanophotonics in Emerging Materials. *Nanophotonics* **2019**, *8*, 1863–1865.
- (18) Aharonovich, I.; Englund, D.; Toth, M. Solid-State Single-Photon Emitters. *Nat. Photonics* **2016**, *10*, 631–641.
- (19) Weston, L.; Wickramaratne, D.; Macko, M.; Alkauskas, A.; Van De Walle, C. G. Native Point Defects and Impurities in Hexagonal Boron Nitride. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2018**, *97*, 214104.
- (20) Narang, P.; Ciccarino, C. J.; Flick, J.; Englund, D. Quantum Materials with Atomic Precision: Artificial Atoms in Solids: *Ab Initio* Design, Control, and Integration of Single Photon Emitters in Artificial Quantum Materials. *Adv. Funct. Mater.* **2019**, *29*, 1904557.
- (21) Ramprasad, R.; Batra, R.; Piliya, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Comput. Mater.* **2017**, *3*, 54.
- (22) Brown, K. A.; Brittman, S.; Maccaferri, N.; Jariwala, D.; Celano, U. Machine Learning in Nanoscience: Big Data at Small Scales. *Nano Lett.* **2020**, *20*, 2–10.
- (23) 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, DOI: [10.1103/PhysRevLett.120.145301](https://doi.org/10.1103/PhysRevLett.120.145301).
- (24) 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.
- (25) Schütt, K. T.; Kessel, P.; Gastegger, M.; Nicoli, K. A.; Tkatchenko, A.; Müller, K.-R. SchNetPack: A Deep Learning Toolbox for Atomistic Systems. *J. Chem. Theory Comput.* **2019**, *15*, 448–455.
- (26) Frey, N. C.; Wang, J.; Vega Bellido, G. I.; Anasori, B.; Gogotsi, Y.; Shenoy, V. B. Prediction of Synthesis of 2D Metal Carbides and Nitrides (MXenes) and Their Precursors with Positive and Unlabeled Machine Learning. *ACS Nano* **2019**, *13*, 3031–3041.
- (27) Cheon, G.; Cubuk, E. D.; Antoniak, E. R.; Blumberg, L.; Goldberger, J. E.; Reed, E. J. Revealing the Spectrum of Unknown Layered Materials with Superhuman Predictive Abilities. *J. Phys. Chem. Lett.* **2018**, *9*, 56.
- (28) Das, S.; Pegu, H.; Sahu, K. K.; Nayak, A. K.; Ramakrishna, S.; Datta, D.; Swayamjyoti, S. Machine Learning in Materials

Modeling—Fundamentals and the Opportunities in 2D Materials. *Synth. Model. Charact. 2D Mater. their Heterostruct.* **2020**, 445–468.

(29) Venturi, V.; Parks, H. L.; Ahmad, Z.; Viswanathan, V. Machine Learning Enabled Discovery of Application Dependent Design Principles for Two-Dimensional Materials. *Mach. Learn. Sci. Technol.* **2020**, 1, 035015.

(30) Medasani, B.; Gamst, A.; Ding, H.; Chen, W.; Persson, K. A.; Asta, M.; Canning, A.; Haranczyk, M. Predicting Defect Behavior in B2 Intermetallics by Merging *Ab Initio* Modeling and Machine Learning. *npj Comput. Mater.* **2016**, 2, 1.

(31) Varley, J. B.; Samanta, A.; Lordi, V. Descriptor-Based Approach for the Prediction of Cation Vacancy Formation Energies and Transition Levels. *J. Phys. Chem. Lett.* **2017**, 8, 5059–5063.

(32) Mannodi-Kanakkithodi, A.; Toriyama, M. Y.; Sen, F. G.; Davis, M. J.; Klie, R. F.; Chan, M. K. Y. Machine-Learned Impurity Level Prediction for Semiconductors: The Example of Cd-Based Chalcogenides. *npj Comput. Mater.* **2020**, 6, 39.

(33) Ferrenti, A. M.; de Leon, N. P.; Thompson, J. D.; Cava, R. J. Identifying Candidate Hosts for Quantum Defects via Data Mining. *npj Comput. Mater.* **2020**, 6, 126.

(34) Walsh, A.; Zunger, A. Instilling Defect Tolerance in New Compounds. *Nat. Mater.* **2017**, 16, 964–967.

(35) Turiansky, M. E.; Alkauskas, A.; Bassett, L. C.; Van de Walle, C. G. Dangling Bonds in Hexagonal Boron Nitride as Single-Photon Emitters. *Phys. Rev. Lett.* **2019**, 123, 127401.

(36) Hayee, F.; Yu, L.; Zhang, J. L.; Ciccarino, C. J.; Nguyen, M.; Marshall, A. F.; Aharonovich, I.; Vučković, J.; Narang, P.; Heinz, T. F.; Dionne, J. A. Revealing Multiple Classes of Stable Quantum Emitters in Hexagonal Boron Nitride with Correlated Optical and Electron Microscopy. *Nat. Mater.* **2020**, 19, 534–539.

(37) Gottscholl, A.; Kianinia, M.; Soltamov, V.; Orlinskii, S.; Mamin, G.; Bradac, C.; Kasper, C.; Krambrock, K.; Sperlich, A.; Toth, M.; Aharonovich, I.; Dyakonov, V. Initialization and Read-Out of Intrinsic Spin Defects in a van der Waals Crystal at Room Temperature. *Nat. Mater.* **2020**, 19, 540–545.

(38) Exarhos, A. L.; Hopper, D. A.; Grote, R. R.; Alkauskas, A.; Bassett, L. C. Optical Signatures of Quantum Emitters in Suspended Hexagonal Boron Nitride. *ACS Nano* **2017**, 11, 3328–3336.

(39) Du, Z.; Yang, S.; Li, S.; Lou, J.; Zhang, S.; Wang, S.; Li, B.; Gong, Y.; Song, L.; Zou, X.; Ajayan, P. M. Conversion of Non-Van der Waals Solids to 2D Transition-Metal Chalcogenides. *Nature* **2020**, 577, 492–496.

(40) Puthirath Balan, A.; Radhakrishnan, S.; Woellner, C. F.; Sinha, S. K.; Deng, L.; Reyes, C. D. L.; Rao, B. M.; Paulose, M.; Neupane, R.; Apte, A.; Kochat, V.; Vajtai, R.; Harutyunyan, A. R.; Chu, C. W.; Costin, G.; Galvao, D. S.; Marti, A. A.; Van Aken, P. A.; Varghese, O. K.; et al. Exfoliation of a Non-Van der Waals Material from Iron Ore Hematite. *Nat. Nanotechnol.* **2018**, 13, 602–609.

(41) Jha, D.; Choudhary, K.; Tavazza, F.; Liao, W.-k.; Choudhary, A.; Campbell, C.; Agrawal, A. Enhancing Materials Property Prediction by Leveraging Computational and Experimental Data Using Deep Transfer Learning. *Nat. Commun.* **2019**, 10, 5316.

(42) Pan, S. J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, 22, 1345–1359.

(43) Hastrup, S.; Strange, M.; Pandey, M.; Deilmann, T.; Schmidt, P. S.; Hinsche, N. F.; Gjerding, M. N.; Torelli, D.; Larsen, P. M.; Riis-Jensen, A. C.; Gath, J.; Jacobsen, K. W.; Mortensen, J. J.; Olsen, T.; Thygesen, K. S. The Computational 2D Materials Database: High-Throughput Modeling and Discovery of Atomically Thin Crystals. *2D Mater.* **2018**, 5, 042002.

(44) Battaglia, P. W.; Hamrick, J. B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; Gulcehre, C.; Song, F.; Ballard, A.; Gilmer, J.; Dahl, G.; Vaswani, A.; Allen, K.; Nash, C.; Langston, V.; et al. Relational Inductive Biases, Deep Learning, and Graph Networks. *ArXiv*, v3, October 17, **2018**, <http://arxiv.org/abs/1806.01261> (accessed 2020-03-26).

(45) 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, 011002.

(46) Weber, J. R.; Koehl, W. F.; Varley, J. B.; Janotti, A.; Buckley, B. B.; Van de Walle, C. G.; Awschalom, D. D. Quantum Computing with Defects. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 8513–8518.

(47) Gupta, S.; Yang, J. H.; Jakobson, B. I. Two-Level Quantum Systems in Two-Dimensional Materials for Single Photon Emission. *Nano Lett.* **2019**, 19, 408–414.

(48) Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N. E. R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; Chard, K.; Asta, M.; Persson, K. A.; Snyder, G. J.; Foster, I.; Jain, A. Matminer: An Open Source Toolkit for Materials Data Mining. *Comput. Mater. Sci.* **2018**, 152, 60–69.

(49) Meredig, B.; Agrawal, A.; Kirklin, S.; Saal, J. E.; Doak, J. W.; Thompson, A.; Zhang, K.; Choudhary, A.; Wolverton, C. Combinatorial Screening for New Materials in Unconstrained Composition Space with Machine Learning. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, 89, 94104.

(50) Dunn, A.; Wang, Q.; Ganose, A.; Dopp, D.; Jain, A. Benchmarking Materials Property Prediction Methods: The Matbench Test Set and Automatminer Reference Algorithm. *ArXiv*, ver. 2, May 8, **2020** <http://arxiv.org/abs/2005.00707> (accessed 2020-05-05).

(51) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, 77, 3865–3868.

(52) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, 125, 224106.

(53) Hybertsen, M. S.; Louie, S. G. First-Principles Theory of Quasiparticles: Calculation of Band Gaps in Semiconductors and Insulators. *Phys. Rev. Lett.* **1985**, 55, 1418–1421.

(54) Breiman, L. Random Forests. *Mach. Learn.* **2001**, 45, 5–32.

(55) Popescu, V.; Zunger, A. Extracting *E* versus *k* Effective Band Structure from Supercell Calculations on Alloys and Impurities. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, 85, 85201.

(56) Banfield, R. E.; Hall, L. O.; Bowyer, K. W.; Kegelmeyer, W. P. A Comparison of Decision Tree Ensemble. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, 29 (1), 173–180.

(57) Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, 58, 267–288.

(58) Hoerl, A. E.; Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, 12, 55–67.

(59) Kent, P. R. C.; Zunger, A. Theory of Electronic Structure Evolution in GaAsN and GaPN Alloys. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, 64, No. 115208, DOI: 10.1103/PhysRevB.64.115208.

(60) Thomas, D. G.; Hopfield, J. J.; Frosch, C. J. Isoelectronic Traps Due to Nitrogen in Gallium Phosphide. *Phys. Rev. Lett.* **1965**, 15, 857–860.

(61) Hus, S. M.; Ge, R.; Chen, P.-A.; Chiang, M.-H.; Donnelly, G. E.; Ko, W.; Huang, F.; Liang, L.; Li, A.-P.; Akinwande, D. Single-Defect Memristor in MoS₂ Atomic-Layer. *ArXiv*, ver. 1, February 4, **2020**, <http://arxiv.org/abs/2002.01574> (accessed 2020-03-20).

(62) Lounis, B.; Orrit, M. Single-Photon Sources. *Rep. Prog. Phys.* **2005**, 68, 1129.

(63) Bishop, S. G.; Hadden, J. P.; Alzahrani, F. D.; Hekmati, R.; Huffaker, D. L.; Langbein, W. W.; Bennett, A. J. Room-Temperature Quantum Emitter in Aluminum Nitride. *ACS Photonics* **2020**, 7, 1636–1641.

(64) Zhang, F.; Zhang, H.; Krylyuk, S.; Milligan, C. A.; Zhu, Y.; Zemlyanov, D. Y.; Bendersky, L. A.; Burton, B. P.; Davydov, A. V.; Appenzeller, J. Electric-Field Induced Structural Transition in Vertical MoTe₂- and Mo_{1-x}W_xTe₂-Based Resistive Memories. *Nat. Mater.* **2019**, 18, 55–61.

(65) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, 54, 11169–11186.

(66) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 1758–1775.

(67) 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.

(68) Freysoldt, C.; Grabowski, B.; Hickel, T.; Neugebauer, J.; Kresse, G.; Janotti, A.; Van De Walle, C. G. First-Principles Calculations for Point Defects in Solids. *Rev. Mod. Phys.* **2014**, *86*, 253–305.

(69) Home-Keras Documentation. <https://keras.io/> (accessed 2020-04-09).

(70) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M.; Jia, Y.; Jozefowicz, R.; Kaiser, L.; Kudlur, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *ArXiv*, ver. 2, March 16, **2016**, <http://arxiv.org/abs/1603.04467> (accessed 2020-04-09).

(71) Pedregosa, F.; Weiss, R.; Brucher, M.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.