

## ARTICLE OPEN



# Sparse representation for machine learning the properties of defects in 2D materials

Nikita Kazeev<sup>1,2,6</sup>, Abdalaziz Rashid Al-Maeni<sup>3,6</sup>, Ignat Romanov<sup>3</sup>, Maxim Faleev<sup>4</sup>, Ruslan Lukin<sup>4</sup>, Alexander Tormasov<sup>4</sup>, A. H. Castro Neto<sup>1,5</sup>, Kostya S. Novoselov<sup>1</sup>✉, Pengru Huang<sup>1</sup>✉ and Andrey Ustyuzhanin<sup>1,2</sup>✉

Two-dimensional materials offer a promising platform for the next generation of (opto-) electronic devices and other high technology applications. One of the most exciting characteristics of 2D crystals is the ability to tune their properties via controllable introduction of defects. However, the search space for such structures is enormous, and ab-initio computations prohibitively expensive. We propose a machine learning approach for rapid estimation of the properties of 2D material given the lattice structure and defect configuration. The method suggests a way to represent configuration of 2D materials with defects that allows a neural network to train quickly and accurately. We compare our methodology with the state-of-the-art approaches and demonstrate at least 3.7 times energy prediction error drop. Also, our approach is an order of magnitude more resource-efficient than its contenders both for the training and inference part.

npj Computational Materials (2023)9:113; <https://doi.org/10.1038/s41524-023-01062-z>

## INTRODUCTION

Atomic-scale tailoring of materials is one of the most promising paths towards achieving new, both quantum and classical properties. Controllable defect engineering, i.e., introduction of vacancies or desired impurities, enables properties modifications and new functionalities in crystalline materials<sup>1</sup>. The opportunities for such controlled material engineering methods got a dramatic boost in the past two decades with the development of the methods for exfoliation of crystal into two-dimensional atomic layer<sup>2</sup>. The reduced dimensionality in layered two-dimensional materials makes it possible to manipulate defects atom by atom and tune their properties down to quantum mechanics limits<sup>3</sup>. Such atomic-scale preparation and fabrication techniques hold promise for the continual development of the semiconductor industry in the post-Moore age and the development of novel technologies such as quantum computing<sup>4</sup>, catalysts<sup>5</sup>, and photovoltaics<sup>6</sup>.

Despite decades of research efforts, knowledge of the structure-property relation for defects in crystals is still limited. Only a small subset of defects in the vast configuration space have been investigated<sup>7</sup>. The properties of complexes of multiple point defects, where quantum phenomena dominate, depend on the composition and configuration of such defects in a strongly non-trivial manner, making their prediction a very hard problem. The diversity and complexity of the problem come from the exchange interaction of defect orbitals separated by discrete lattices<sup>8</sup>. On the other hand, the vast chemical and configuration space prohibits a thorough exploration of such structures by traditional trial and error experiments and even for computationally expensive state-of-the-art quantum mechanic simulations.

The recent development of large materials databases has stimulated the application of deep learning methods for atomistic predictions. Machine learning (ML) methods trained on density functional theory (DFT) calculations have been used to identify

materials for batteries, catalysts, and many other applications. Machine learning methods accelerate the design of the new materials by predicting material properties with accuracy comparable to ab-initio calculations, but with orders of magnitude lower computational costs<sup>9,10</sup>. A series of fast and accurate deep learning architectures have been presented during the last few years. The most successful of them are graph neural networks, such as MEGNet<sup>11</sup>, CGCNN<sup>12</sup>, SchNet<sup>13</sup>, GemNet<sup>14</sup>, etc.

In this work, we propose a method for predicting the energetic and electronic structures of defects in 2D materials with machine learning. Firstly, a machine learning-friendly 2D Material Defect database (2DMD) was established employing high throughput DFT calculations<sup>15</sup>. The database is composed of both structured datasets and dispersive datasets of defects in represented 2D materials such as MoS<sub>2</sub>, WSe<sub>2</sub>, h-BN, GaSe, InSe, and black phosphorous (BP). We use the datasets to evaluate the performance of the previously reported approaches along with ours' which was specially designed to provide accurate description of materials with defects. Our computational experiments show that our approach provides a significant increase in prediction accuracy compared to the state-of-the-art general methods. The high accuracy allows to reproduce the nonlinear non-monotonic property-distance correlation of defects which is a combination of the quantum mechanic effect and the periodic lattice nature in 2D materials. The general methods, on the other hand, mostly fail to predict such property-structure functionals, as we show in subsection "Aggregate performance". Most importantly, our method shows great transferability for a wide range of defect concentrations in the various 2D materials we studied.

Machine learning offers two principal approaches to predicting atomistic properties: graph neural networks (GNN) and physics-based descriptors. Graph neural networks have several valuable properties that make them uniquely suitable for modeling atomic systems: invariance to permutations, rotations, and translation;

<sup>1</sup>Institute for Functional Intelligent Materials, National University of Singapore, 4 Science Drive 2, 117544 Singapore, Singapore. <sup>2</sup>Constructor University Bremen gGmbH, Campus Ring 1, Bremen 28759, Germany. <sup>3</sup>HSE University, Myasnitskaya Ulitsa, 20, 101000 Moscow, Russia. <sup>4</sup>Innopolis University, Universitetskaya St, 1, Innopolis, Republic of Tatarstan 420500, Russia. <sup>5</sup>Centre for Advanced 2D Materials, National University of Singapore, 6 Science Drive 2, 117546 Singapore, Singapore. <sup>6</sup>These authors contributed equally: Nikita Kazeev, Abdalaziz Rashid Al-Maeni. ✉email: kostya@nus.edu.sg; pengru@nus.edu.sg; andrey.u@nus.edu.sg

natural encoding of the locality of interactions. In the recent Open Catalyst benchmark<sup>5</sup>, GNNs solidly outperform the physics-based descriptors. Therefore, in this section, we only focus on GNNs.

Xie et al.<sup>12</sup> is one of the first works to propose applying a convolutional GNN to materials. Wolverton et al.<sup>16</sup> improves on it, by incorporating Voronoi-tessellated atomic structure, 3-body correlations of neighboring atoms, and chemical representation of interatomic bonds. Schütt et al.<sup>13</sup> (Schnet) proposes continuous-filter convolutional layers. Chen et al.<sup>11</sup> (MEGNet) uses a more advanced GNN: message-passing, instead of convolutional. Klicpera et al.<sup>14,17</sup> (GemNet) redresses an important shortcoming of the previous message-passing GNN's: losing of geometric information due to considering only distances as edge features. Zitnick et al.<sup>18</sup> improves handling of angular information. Cloudhary et al. presents an Atomistic Line Graph Neural Network (ALIGNN)<sup>19</sup>, a GNN architecture that performs message passing on both the interatomic bond graph and its line graph corresponding to bond angles. Ying et al.<sup>20,21</sup> introduces Graphormer, a hybrid model between Transformers and GNNs allows for more expressive aggregation operations.

Even though the described models are not evaluated on crystals with defects - they are in-principle capable of handling any atomistic structures, and thus we use the most established ones as baselines. Moreover, we demonstrate our approach using one of the most renowned GNN architectures for materials: MEGNet (see details in subsection "General message-passing graph neural networks for materials").

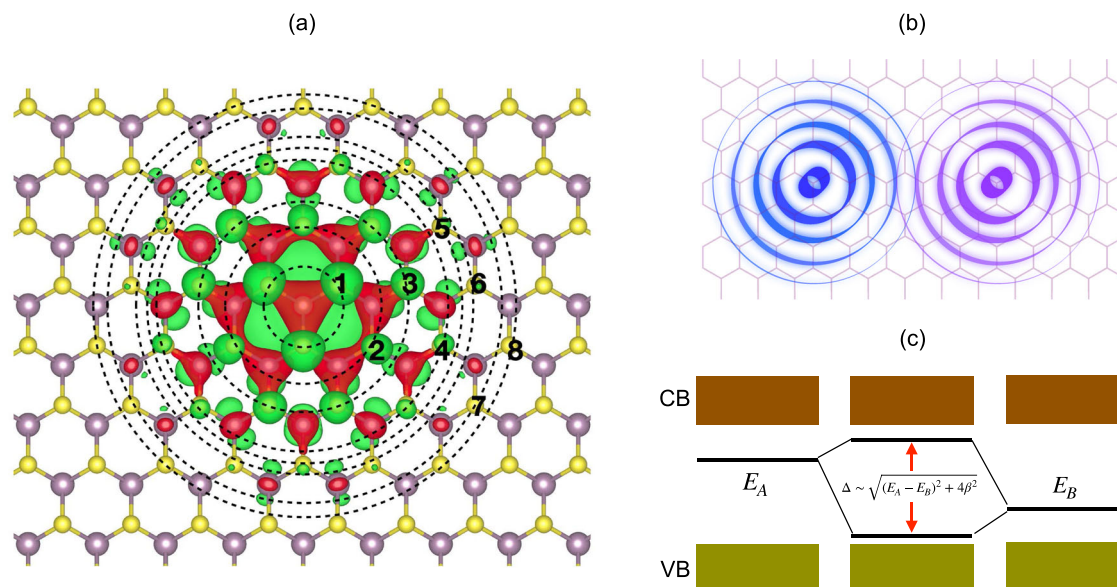
The introduction of a defect site in general creates disturbed electronic states and the wave function associated with such states fluctuates over a distance of a few lattice constants depending on the localization of the electrons in the host lattice, Fig. 1. This results in some localized defect levels in the energy spectrum of the solid. From a quantum embedding theory point of view<sup>22</sup>, the defects could be seen as the active region of interest embedded in the periodic lattice. Accordingly, the defect levels are governed by the interactions of the unsaturated electrons in the background of the valence band electrons. The properties of a defect complex composed of more than one defect sites are governed by the interference of the wave functions of such

electrons<sup>23</sup>. As the result, the formation energy, positions of defect levels, and the HOMO-LUMO gap are non-trivially dependent on the defect configurations. As schematically shown in Fig. 1, two defect states interfere with each other, and the separation of the bonding and antibonding states is governed by the exchange integral of the two states in the screening background of valence electrons. The exchange integral is subtly dependent on the positions of defect components, and such a HOMO-LUMO gap is a complex functional of the defect configuration. It is still a challenge for machine learning to precisely predict such nonlinear quantum mechanic behavior of defects.

Machine learning methods have been proposed for prediction of the formation energies of single point defects<sup>24–26</sup> across different materials, but the authors didn't consider configurations with multiple interacting defects. In a recent preprint<sup>27</sup>, the authors use a model based on CGCNN<sup>12</sup> to conduct a large scale screening of single-vacancy structures for diverse energy applications. In ref. <sup>4</sup>, the authors use MEGNet architecture for prediction of the properties of pristine 2D materials and choosing the ones that make optimal hosts for engineered point defects. Then they use matminer<sup>28</sup> combined with Random Forest<sup>29</sup> for predicting the properties of structures with point defects. A similar descriptor-based approach is used in ref. <sup>30</sup>. We evaluate the descriptor-based approach for our data, as described in subsection "Physics-based descriptors".

ReaxFF<sup>31</sup> potential has been developed for dichalcogenides, and is used for studying defect dynamics<sup>32,33</sup>. The potential is very computationally efficient, and thus allows to probe dynamics on a larger time scale. However, it is not as precise as DFT, and doesn't offer a way to predict the electronic properties.

The paper is structured as follows. Section "Sparse representation of crystals with defects" presents our proposed method for representing structures with defects to machine learning algorithms. Section "Dataset" provides the description of the dataset we use for evaluation. Sections "Aggregate performance", "Quantum oscillations prediction" present the computational experiments, where we compare performances of different methods. Finally, section "Discussion" summarizes our work.



**Fig. 1** The quantum mechanic nature of defects. **a** DFT-simulated defect wave function centered at a single Mo site of MoS<sub>2</sub> crystal lattice. Red and green colors represent the opposite phases of the wave function isosurfaces. The dash circles are an illustration of the electronic orbital shells. **b** A schematic representation of the wave function interference of two defect sites in a crystal lattice. **c** The defect levels are governed by the exchange interaction of defect components, and hence the separation of the defect levels is dominated by the exchange integral according to the valence bond theory.

## RESULTS

### Sparse representation of crystals with defects

For machine learning algorithms, an atomic structure is a so-called point cloud: a set of points in 3D space. Each point is associated with a vector of properties, which at the least contains the atomic number, but may also include more physics-based features, such as radius, the number of valence electrons, etc.

The structures with defects present a challenge to machine learning algorithms. The neighborhoods of the majority of the atoms are not affected by the point defects. In principle, this shouldn't be an obstacle for a perfect algorithm. In practice, however, this comparatively small difference in the full structures is hard to learn. As we demonstrate in section "Aggregate performance" with our computational experiments, state-of-the-art algorithms underperform on crystal structures with defects.

We propose a way to represent structures with defects that makes the problem of predicting properties easy for the ML algorithms leading to better performance. The core idea is presented in Fig. 2: instead of treating a crystal structure as a point cloud of atoms, we treat it as a point cloud of defects. To obtain it, we take the structure with defects, remove all the atoms that are not affected by substitution defects, and add virtual atoms on the vacancy sites.

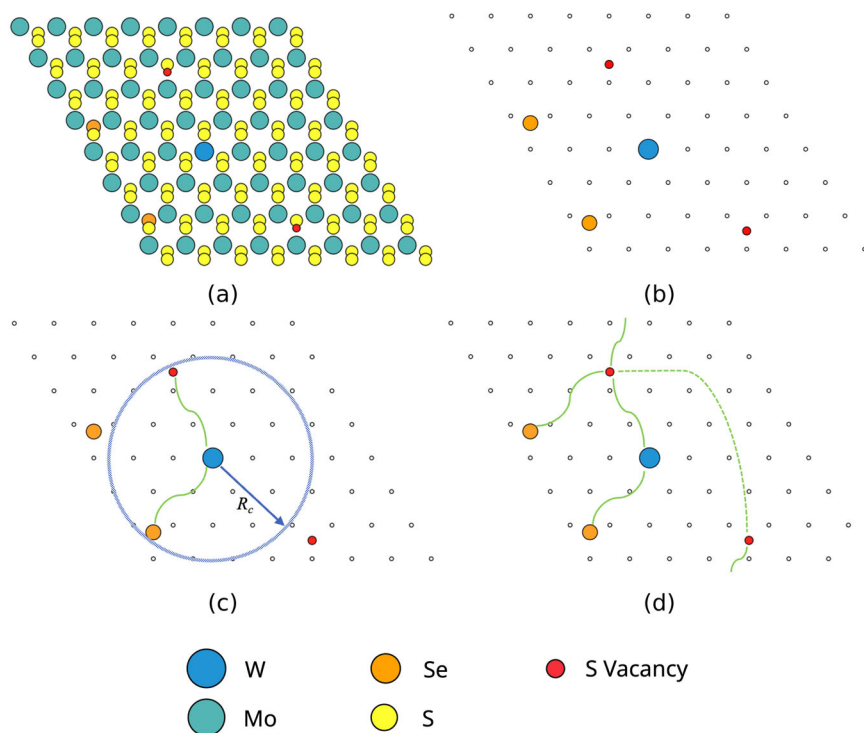
Each point has two parameters in addition to the coordinates: the atomic number of the atom on the site in the pristine structure and the atomic number of the atom in the structure with the defect. Vacancies are considered to have atomic number 0.

The structure of the pristine unit cell is encoded as a global state for each structure using a vector with the set of atomic numbers of the pristine material, e.g. (42, 16) for  $\text{MoS}_2$ . This simple approach is sufficient on our structures. As a future direction, in case generalization between materials is desired, graph embedding would be a logical choice: pristine material unit cell as an

input to a different GNN, which outputs a vector of fixed dimensionality.

Secondly, we propose an augmentation specific to graph neural networks and 2D crystals: adding the difference in  $z$  coordinate (perpendicular to the material surface) as an edge feature. Normally, such a feature would break the rotational symmetry. But in the case of a 2D crystal, the direction perpendicular to the material surface is physically defined and thus can be used.

In a crystal, the replacement of an atom or the introduction of a vacancy causes a major disruption of the electronic states. Given the wave nature of electrons, the introduction of a localized defect creates oscillations in the electronic wave functions at the atomic level similar to a rock thrown into a pond. In the case of crystals, the wave function oscillations may involve one or several electronic orbitals, and the amplitude of those oscillations decays away from the defect at a rate that depends on the nature of those orbitals. This oscillatory nature of the electronic states away from a defect leads to the formation of electronic orbital shells (EOS). We ascribe an EOS index to such shells, that labels the amplitude of the wave function in decreasing order, the  $S$  atom labeled 1 filling the largest amplitude, as we show in Fig. 1. Formally we define EOS orbitals as follows. Firstly, we project all atoms on the  $x$ - $y$  plane, making a truly 2D representation of the material. For binary crystals, for each atom, we draw circles centered on it and passing through the atoms of the other species, numbering them in the order of radius increase. For unary materials (BP in our dataset), the circle radii are multiples of the unit cell size. The circle number is the EOS index of the site with respect to the central atom. The intuition behind those indices is described in the paper<sup>34</sup>, which claims that the atomic electron shells' interaction strength is not monotonic with respect to atom distance, but it oscillates in a way such that minima and maxima coincide with the crystal lattice nodes. To represent those oscillations, we also add parity of the EOS index as a separate feature, which we call *EOS parity*.



**Fig. 2** The transition from full to sparse representation for an example  $\text{MoS}_2$  supercell. **a** A full  $\text{MoS}_2$  structure with one  $\text{Mo} \rightarrow \text{W}$ , two  $\text{S} \rightarrow \text{Se}$  substitutions, and two  $\text{S}$  vacancies; **b** sparse representation, containing only defects types and coordinates; **c** intermediate point of the graph construction: cutoff radius centered on the metal substitution defect; **d** the final graph with sparse representation. A dashed green line indicates a virtual graph edge that effectively connects one node with another via a periodic boundary condition.

## Incorporating sparse representation into a graph neural network

Our proposed representation fits into the graph neural networks (GNN) framework (described in subsection “General message-passing graph neural networks for materials” as follows:

1. Graph nodes correspond to point defects, not to all the atoms in a structure;
2. Threshold for connecting nodes with edges is increased;
3. Node attributes contain the atomic number of the atom on this site in the pristine structure, and the atomic number of the atom in the structure with defect, with 0 for vacancies;
4. Edge attributes contain not only the Euclidean distance between point defects corresponding to the adjacent vertices, but also EOS index, EOS parity index, and  $Z$  plane distance;
5. Input global state contains the chemical composition of the crystal as a vector of atomic numbers.

## Dataset

We established a machine learning friendly 2D material defect database (2DMD)<sup>15</sup> for the training and evaluation of models. The datasets contain structures with point defects for the most widely used 2D materials: MoS<sub>2</sub>, WSe<sub>2</sub>, hexagonal boron nitride (h-BN), GaSe, InSe, and black phosphorous (BP). The types of point defects are listed in the Table 1. Supercell details are available in Supplementary Table 1 and example defect depictions in Supplementary Fig. 1.

The datasets consist of two parts: low defect concentration of structured configurations and high defect concentration of random configurations. The low defect concentration part consists of 5933 MoS<sub>2</sub> structures and 5933 WSe<sub>2</sub> with all possible configurations in the 8x8 supercell for defect types depicted in Fig. 3. We used `pymatgen`<sup>35</sup> to find the configurations, taking into account symmetry. The high-density dataset contains a sample of randomly generated substitution and vacancy defects for all the materials. For each total defect concentration 2.5%, 5%, 7.5%, 10%, and 12.5% 100 structures were generated, totaling 500 configurations for each material and 3000 in total. Overall, the dataset contains 14866 structures with 120–192 atoms each. The datasets as designed could provide training data for AI methods both the fine features of quantum mechanic nature and those features associated with different elements, crystal structures, and defect concentrations. We used Density Functional Theory (DFT) for computing the properties, the details are described in the subsection “DFT computations”

We use two target variables for evaluating machine learning methods: defect formation energy per site and HOMO–LUMO gap.

Formation energy, i.e., the energy required to create a defect is defined as

$$E_f = E_D - E_{\text{pristine}} + \sum_i n_i \mu_i \quad (1)$$

**Table 1.** Point defect types present in the 2DMD datasets.

Material	Substitutions	Vacancies
MoS <sub>2</sub>	S→Se; Mo→W	Mo; S
WSe <sub>2</sub>	Se→S; W→Mo	W; Se
h-BN	B→C; N→C	B; N
GaSe	Ga→In; Se→S	Ga; Se
InSe	In→Ga; Se→S	In; Se
BP	P→N	P

where  $E_D$  is the total energy of the structure with defects,  $E_{\text{pristine}}$  is the total energy of the pristine base material,  $n_i$  is the number of the  $i$ -th atoms removed from ( $n_i > 0$ ) or added to ( $n_i < 0$ ) the supercell to/from a chemical reservoir, and  $\mu_i$  is the chemical potential of the  $i$ -th element, computed with the same DFT settings. Finally, to make the results better comparable across examples with different numbers of defects, we normalize the formation energy by dividing it by the number of defect sites:

$$E'_f = E_f / N_d, \quad (2)$$

where  $N_d$  is the number of defects in the structure.

The electronic properties of defects are characterized by the energy spacing between the highest occupied states and the lowest unoccupied states. For the sake of representation, we adopt the terminologies of HOMO–LUMO gap for the separation of defect levels. Defects in some of the materials (BP, GaSe, InSe, h-BN) have unpaired electrons and hence non-zero magnetic momentum. Therefore, DFT was computed taking into account two channels of spin-up and spin-down bands, resulting in the majority and minority HOMO–LUMO gaps. For evaluating the machine learning algorithms, we took the minimum of those gaps as the target variable.

## Aggregate performance

We split the dataset into 3 parts: train (60%), validation (20%), and test (20%). The split is random and stratified with the respect to each base material. For each model, we use random search for hyperparameter optimization; we generate 50 hyperparameter configurations, train the model with each configuration on the train part, and select the best-performing configuration by evaluating quality on the validation part. The search spaces and optimal configurations are present in Supplementary Discussion 1. To obtain the final result, we train each model with the optimal parameters on the combination of train and validation parts and evaluate the quality on the unseen test part. We do this 12 times to estimate the effects of the random initialization. We use unrelaxed structures as inputs and predict the energy and HOMO–LUMO gap after relaxation. To account for the material class imbalance, we use weighted mean absolute error (MAE) as the quality metric during both training and evaluation:

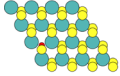
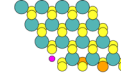
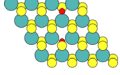
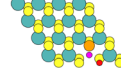
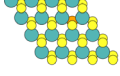
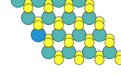
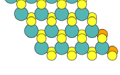
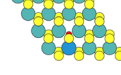
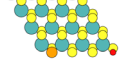
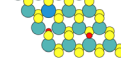
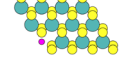
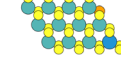
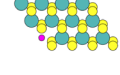
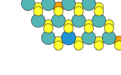
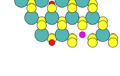
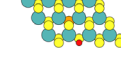
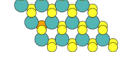
$$\text{MAE} = \frac{\sum_{i=1}^N |\bar{y}_i - y_i| w_i}{\sum_{i=1}^N w_i}, \quad (3)$$

where  $w_i$  is the weight assigned to each example;  $y_i$  is the predicted value;  $\bar{y}_i$  is the true value;  $N$  is the number of the structures in the dataset. The purpose of using weights is to prevent the combined error value from being dominated by the low defect density dataset part, as it's 4 times more numerous compared to the high defect density part. The weights are computed as follows:

$$w_{\text{part}} = \frac{N_{\text{total}}}{C_{\text{parts}} N_{\text{part}}}, \quad (4)$$

where  $w_{\text{dataset}}$  is the weight associated with each example in a dataset part,  $N_{\text{total}} = 14866$  is the total number of examples,  $C_{\text{parts}} = 8$  is the total number of dataset parts (2 low-density and 6 high-density),  $N_{\text{part}} \in \{500, 5933\}$  is the number of examples in the part (500 for low defect density parts, 5933 for the high defect density parts).

We compare the performance of our sparse representation combined with MEGNet<sup>11</sup> to several baseline methods: MEGNet, SchNet<sup>13</sup>, and GemNet<sup>14</sup> on full representation, and CatBoost<sup>36</sup> with `matminer`-generated features 4.3. The results are presented in Table 2. For energy prediction, our model achieves 3.7× less combined MAE compared to the best baseline, with 2.2×–6.0× difference in individual dataset parts. For HOMO–LUMO gap, using sparse representation doesn't lead to an increase in overall

Mo	S	S	Num	Example	Mo	S	S	Num	Example
	vac		1		vac	Se	Se	743	
	vac	vac	19		vac	vac	Se	1415	
	Se		1		W			1	
	Se	Se	19		W	vac		15	
	vac	Se	29		W	vac	vac	743	
vac			1		W	Se		15	
vac	vac		15		W	Se	Se	743	
vac	vac	vac	743		W	vac	Se	1415	
vac	Se		15						

W

Se

Mo Vacancy

Mo

S

S Vacancy

**Fig. 3 Defect types in low density dataset.** The ‘Mo’ and two ‘S’ columns denote the type of site that is being perturbed either by substituting the listed element, or a vacancy (vac). ‘Num’ column contains the number of structures with defects of the type in the dataset. Finally, ‘Example’ column presents a structure with such defect. Note that the actual supercell size is  $8 \times 8$ , here we draw  $4 \times 4$  window centred on the defects to conserve space. Drawings of example structures with  $8 \times 8$  supercells are available in Supplementary Fig. 1.

prediction quality. The prediction quality for  $\text{MoS}_2$  and  $\text{WSe}_2$  is improved by a factor of 1.3–4.8, but this is outweighed by a factor of 1.06–1.15 increase in MAE for the other materials. Coincidentally, the combined MAEs are similar, being averaged over the absolute error values.

In terms of computation time, when trained on a Tesla V100 GPU, MegNet with sparse representation took 45 minutes; MegNet with full representation 105 minutes; GemNet 210 minutes; SchNet 100 minutes; CatBoost 0.5 minutes. Low memory footprint and GPU utilization allow to fit 4 simultaneous runs with sparse representation on the same GPU (16 GiB RAM) without losing speed, this is not possible for the GNNs running on full representation. Computing matminer features can’t be done on a GPU, and costs 7.5 CPU core-minutes per structure and 1860 core-hours for the whole 2DMD dataset. Model configurations are listed in Supplementary Discussion 1.

### Quantum oscillations prediction

In addition to the overall performance, we specifically evaluate the models with the respect to learning quantum oscillations. We use the  $\text{MoS}_2$  with one Mo and one S vacancy as the test dataset, and the rest of the 2DMD dataset as the training dataset. No sample weighting is used. We train every model 12 times with both optimal hyperparameters found via random search and the default parameters.

As seen in the Table 2, sparse representation performs especially well on the low-density data. This behavior extends nicely to the 2-vacancy data, as shown in Fig. 4. The baseline approaches fail to meaningfully learn the dependence, while sparse representation succeeds perfectly, including the non-monotonous reduction at  $5 \text{ \AA}$ .

As shown in the Supplementary Discussion 3, the result is similar for untuned hyperparameters.

### Ablation study

The ablation study investigates how much each proposed improvement contributes to the final result. The performance values are presented in Table 3.

To conduct the ablation study, we took the optimal configurations for MEGNet with sparse and full representations found by random search. We then took the configuration for the sparse representation turned off our enchantments one-by-one, trained and evaluated the resulting models. We use a value averaged over 12 experiments, same as in Table 2 to estimate training stability.

For formation energy, just enabling the z coordinate difference in sparse representation edges allows the Sparse-Z model to outperform the Full model everywhere except h-BN; adding pristine atom species (Sparse-Z-Were) as the node features contributes the most of the remaining gain. The most likely explanation for the importance of the pristine species for h-BN is

**Table 2.** Performance of the different methods in terms of the mean absolute error (MAE).

Formation energy per site MAE, meV; lower is better						
Material	Density	SchNet	GemNet	MEGNet (Full)	CatBoost	Sparse (MEGNet)
Combined	Both	631 ± 31	483 ± 91	158 ± 47	164 ± 5	<b>43 ± 5</b>
BP	High	2088 ± 72	1490 ± 429	198 ± 211	382 ± 30	<b>80 ± 10</b>
GaSe	High	245 ± 12	230 ± 41	107 ± 25	103 ± 4	<b>48 ± 7</b>
InSe	High	268 ± 19	247 ± 26	95 ± 27	137 ± 5	<b>35 ± 2</b>
MoS <sub>2</sub>	High	321 ± 100	535 ± 206	136 ± 22	136 ± 5	<b>23 ± 5</b>
WSe <sub>2</sub>	High	536 ± 123	575 ± 181	112 ± 33	162 ± 6	<b>23 ± 4</b>
h-BN	High	1442 ± 68	697 ± 315	496 ± 229	363 ± 17	<b>127 ± 16</b>
MoS <sub>2</sub>	Low	65 ± 5	44 ± 14	58 ± 11	12.6 ± 0.4	<b>4 ± 1</b>
WSe <sub>2</sub>	Low	85 ± 22	42 ± 9	65 ± 16	16.3 ± 0.8	<b>6 ± 1</b>
HOMO – LUMO gap MAE, meV; lower is better						
Material	Density	SchNet	GemNet	MEGNet (Full)	CatBoost	Sparse (MEGNet)
Combined	Both	224 ± 111	166 ± 42	<b>112 ± 3</b>	<b>117 ± 1</b>	<b>112 ± 3</b>
BP	High	208 ± 20	176 ± 10	<b>170 ± 4</b>	<b>174 ± 2</b>	187 ± 9
GaSe	High	309 ± 83	196 ± 11	<b>178 ± 8</b>	<b>173 ± 4</b>	194 ± 11
InSe	High	214 ± 69	178 ± 22	<b>156 ± 7</b>	<b>155 ± 1</b>	<b>167 ± 15</b>
MoS <sub>2</sub>	High	204 ± 121	174 ± 111	54 ± 4	71 ± 4	<b>39 ± 4</b>
WSe <sub>2</sub>	High	186 ± 177	268 ± 182	47 ± 3	106 ± 6	<b>38 ± 4</b>
h-BN	High	244 ± 24	227 ± 6	233 ± 4	<b>208 ± 3</b>	260 ± 14
MoS <sub>2</sub>	Low	187 ± 180	46 ± 42	30 ± 2	26.7 ± 0.8	<b>5.7 ± 0.2</b>
WSe <sub>2</sub>	Low	236 ± 224	64 ± 46	32 ± 5	18.3 ± 0.6	<b>8.1 ± 0.6</b>

Sparse (MEGNet) is our representation implemented in the MEGNet model with all the improvements enabled. SchNet, GemNet, MEGNet (full) have full structures as their input with no additional features. CatBoost uses matminer featurizers. All the models were trained on the same dataset, containing stratified samples of all the parts of 2DMD dataset. The dataset splitting strategy is described in detail in subsection “Aggregate performance”. Combined refers to the whole test sample, with the error contributions weighted according to equation (4). The individual material/density combinations refer to the subsets of the combined test dataset. Error indicates the standard deviation of the results obtained from 12 experiments with the same datasets and model parameters, but different random initialization. Bold font indicates the best results for each dataset, taking the uncertainty into account.

that both atoms can be substituted to C, without this additional information, the model can't distinguish between B and N substitutions. Adding EOS improves expected prediction quality and stability by a small amount for the low-density datasets.

For HOMO–LUMO gap, Sparse-Z-Were and Sparse-Z-Were-EOS perform similarly to Full in terms of the combined metric, outperform it by a factor of 4 for low-density data. EOS again improves prediction quality and stability by a small amount for the low-density datasets.

## DISCUSSION

2D crystals present an incredible potential for the future of material design. Their two-dimensional nature makes them prone to chemical modification, which further increases their tunability for a variety of applications. However, the search space for possible configurations is vast. Thus, the ability to predict the properties of such crystals efficiently becomes a vital task. In this paper, we focus on predicting the properties of such crystals blended with defects, substitutions, and vacancies. State-of-the-art machine learning algorithms struggle to learn the properties of crystals' defects accurately. We propose using a sparse representation combined with graph neural network architectures like MEGnet and show that it dramatically improves energy prediction quality. Our studies demonstrate that the prediction error drops 3.7 times compared to the nearest contender. Moreover, the representation is compatible with any machine learning algorithm based on point clouds. Computationally the training of a graph

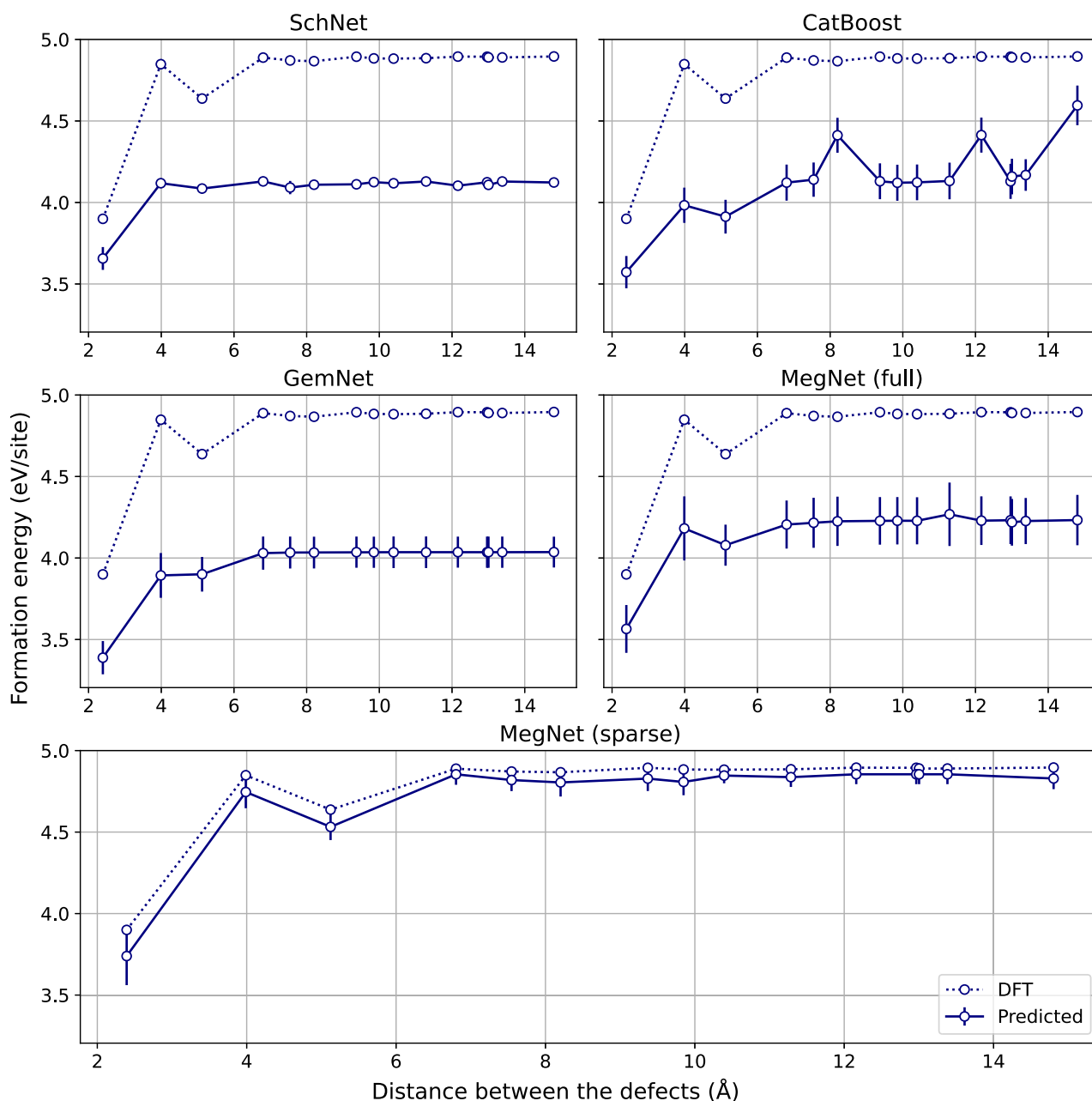
neural network using sparse representation takes at least 4x less memory and 8x less GPU operations compared to the full representation. Thus, we conclude that our approach gives a practical and sound way to explore a vast domain of possible crystal configurations confidently.

We see two principal directions for future work. Firstly, 3D materials. Sparse representation can be used as is for ordinary 3D crystals with point defects. Secondly, generalization to unseen materials. In our paper, we consider setup where each base material is present in the training dataset. Combining sparse defect representation with advanced base material representation<sup>24,25,27</sup> opens up an enticing possibility for predicting properties of defect complexes in new materials, without having to prepare a training dataset with defects in those new materials.

## METHODS

### DFT computations

Our calculations are based on density functional theory (DFT) using the PBE functional as implemented in the Vienna Ab Initio Simulation Package (VASP)<sup>37–39</sup>. The interaction between the valence electrons and ionic cores is described within the projector augmented (PAW) approach<sup>40</sup> with a plane-wave energy cutoff of 500 eV. The initial crystal structures were obtained from the Material Project database, and the supercell sizes and the computational parameters for each material can be found in Supplementary Table 1. Since very large supercells are used for the calculation of defects, the Brillouin zone was sampled using  $\Gamma$ -point



**Fig. 4** Predicted and DFT formation energy as a function of the distance between the defects for  $\text{MoS}_2$  with one Mo and one S vacancy. Error bars represent the standard deviation across 12 models trained from different random initialisation.

only Monkhorst-Pack grid for structural relaxation and denser grids for further electronic structure calculations. A vacuum space of at least 15 Å was used to avoid interaction between neighboring layers. In the structural energy minimization, the atomic coordinates are allowed to relax until the forces on all the atoms are less than 0.01 eV/Å. The energy tolerance is  $10^{-6}$  eV. For defect structures with unpaired electrons, we utilize standard collinear spin-polarized calculations with magnetic ions in a high-spin ferromagnetic initialization (the ion moments can of course relax to a low spin state during the ionic and electronic relaxations). Currently, we are focusing on basic properties of defects at the level of single-particle physics and did not include spin-orbit coupling (SOC) and charged states calculations. Since the materials we considered are normal nonmagnetic semiconductors and none of them are strongly correlated systems, we did not employ the GGA+U method. A comparison of a few selected the computed

values to the ones available in the literature is available in Supplementary Table 2.

#### General message-passing graph neural networks for materials

There are many types of graph neural networks (GNN). In this section, we outline the message-passing neural network proposed by Battaglia et al.<sup>41</sup>. Those became rather popular for analyzing material structure<sup>11</sup>.

To prepare a training sample, a graph is constructed out of a crystal configuration: atoms become graph nodes, and graph edges connect nodes at distances less than a predefined threshold. The connections respect periodic boundary conditions, i.e., for a significantly large threshold, an edge can connect a node to its image in an adjacent supercell. Specific property vectors also characterize nodes and edges. A node contains the atomic

**Table 3.** Performance of various combinations of our proposed features.

Formation energy per site MAE, meV; lower is better						
Material	Density	Full	Sparse	Sparse-Z	Sparse-Z-Were	Sparse-Z-Were-EOS
Combined	Both	158 ± 47	432 ± 29	404 ± 24	<b>42 ± 4</b>	<b>43 ± 5</b>
BP	High	198 ± 211	116 ± 18	104 ± 16	<b>75 ± 9</b>	<b>80 ± 10</b>
GaSe	High	107 ± 25	76 ± 9	66 ± 10	<b>48 ± 7</b>	<b>48 ± 7</b>
InSe	High	95 ± 27	81 ± 8	69 ± 8	<b>34 ± 3</b>	<b>35 ± 2</b>
MoS <sub>2</sub>	High	136 ± 22	223 ± 21	96 ± 15	<b>23 ± 4</b>	<b>23 ± 5</b>
WSe <sub>2</sub>	High	112 ± 33	147 ± 27	63 ± 11	<b>23 ± 2</b>	<b>23 ± 4</b>
h-BN	High	496 ± 229	2741 ± 202	2792 ± 184	<b>123 ± 14</b>	<b>127 ± 16</b>
MoS <sub>2</sub>	Low	58 ± 11	37 ± 15	19 ± 4	<b>5 ± 2</b>	<b>4 ± 1</b>
WSe <sub>2</sub>	Low	65 ± 16	30 ± 9	21 ± 5	<b>7 ± 2</b>	<b>6 ± 1</b>
HOMO – LUMO gap MAE, meV; lower is better						
Material	Density	Full	Sparse	Sparse-Z	Sparse-Z-Were	Sparse-Z-Were-EOS
Combined	Both	<b>112 ± 3</b>	148 ± 5	121 ± 2	<b>114 ± 7</b>	<b>112 ± 3</b>
BP	High	<b>170 ± 4</b>	197 ± 10	192 ± 9	<b>188 ± 18</b>	187 ± 9
GaSe	High	<b>178 ± 8</b>	237 ± 12	205 ± 16	207 ± 16	194 ± 11
InSe	High	<b>156 ± 7</b>	170 ± 8	172 ± 14	<b>160 ± 9</b>	<b>167 ± 15</b>
MoS <sub>2</sub>	High	54 ± 4	110 ± 10	<b>45 ± 2</b>	<b>40 ± 5</b>	<b>39 ± 4</b>
WSe <sub>2</sub>	High	47 ± 3	173 ± 13	66 ± 4	<b>39 ± 6</b>	<b>38 ± 4</b>
h-BN	High	<b>233 ± 4</b>	265 ± 16	271 ± 11	267 ± 21	260 ± 14
MoS <sub>2</sub>	Low	30 ± 2	12 ± 1	7.2 ± 0.4	<b>6.0 ± 0.7</b>	<b>5.7 ± 0.2</b>
WSe <sub>2</sub>	Low	32 ± 5	19 ± 1	11.7 ± 0.6	<b>8.4 ± 0.6</b>	<b>8.1 ± 0.6</b>

Full—full structure is used as MEGNet input. Sparse—sparse structure used as MEGNet input; Sparse-Z adds z coordinate differences in edges; Sparse-Z-Were adds the atomic species of the pristine material on the defects site as a node feature; Sparse-Z-Were-EOS add EOS index and EOS parity as edge features. Bold font indicates the best results for each dataset, taking the uncertainty into account.

number, and an edge contains the Euclidean distance between the atoms it connects.

A layer of a message-passing neural network transforms a graph into another graph with the same connectivity structure, changing only the nodes, edges, and global attributes. The layers are stacked to provide an expressive deep architecture. Let  $G = (V, E, u)$  be a crystal graph from the previous step, the nodes are represented by vectors  $V = \{\mathbf{v}_i\}_{i=0}^{|V|}$ , where  $\mathbf{v}_i \in \mathbb{R}^{d_v}$  and  $|V|$  is the number of atoms in the supercell. The edge states are represented by vectors  $\{\mathbf{e}_k\}_{k=0}^{|E|}$ , where  $\mathbf{e}_k \in \mathbb{R}^{d_e}$ . Each edge has a sender node index  $v^s \in \{0, \dots, |V| - 1\}$ , a receiver node  $v^r \in \{0, \dots, |V| - 1\}$  and a vector of edge attributes. An edge is represented by a tuple  $(\mathbf{v}_k^s, \mathbf{v}_k^r, \mathbf{e}_k)$ , where the superscripts  $s, r$  denote the sender and the receiver nodes respectively. The global state vector  $\mathbf{u} \in \mathbb{R}^{d_u}$  represents the global state of the system. In the input graph, the global state is used to provide the algorithm with information about the system as a whole, in the case of our sparse representation, the composition of the base material. In the output graph, the global state contains the model predictions of the target variables. A message-passing layer is a mapping from  $G = (V, E, \mathbf{u})$  to  $G' = (V', E', \mathbf{u}')$ , this mapping is based on update rules for nodes, edges, and global state. Edge update rule operates on the information from the sender  $\mathbf{v}_k^s$ , receiver nodes  $\mathbf{v}_k^r$ , edge itself  $\mathbf{e}_k$ , and the global state  $\mathbf{u}$ . We can represent this rule by a function  $\phi_e$ :

$$\mathbf{e}'_k = \phi_e(\mathbf{v}_i^s, \mathbf{v}_i^r, \mathbf{e}_k, \mathbf{u}). \quad (5)$$

Node update rule aggregates the information from all the edges  $E_{\mathbf{v}_i} = \{\mathbf{e}'_k | \mathbf{e}_k \in \text{neighbors}(\mathbf{v}_i)\}$  connected to the node  $\mathbf{v}_i$ , the node itself  $\mathbf{v}_i$  and the global state  $\mathbf{u}$ . We can represent this rule by

function  $\phi_v$ :

$$\bar{\mathbf{e}}_i = \frac{1}{|E_{\mathbf{v}_i}|} \sum_{\text{neighbors}(\mathbf{v}_i)} \mathbf{e}'_k \quad (6)$$

$$\mathbf{v}'_i = \phi_v(\bar{\mathbf{e}}_i, \mathbf{v}_i, \mathbf{u}) \quad (7)$$

Finally, the global state  $\mathbf{u}$  is updated based on the aggregation of both nodes and edges alongside the global state itself and process them with  $\phi_u$ :

$$\bar{\mathbf{u}}_v = \frac{1}{|V|} \sum_{i=0}^{|V|-1} \mathbf{v}'_i \quad (8)$$

$$\bar{\mathbf{u}}_e = \frac{1}{|E|} \sum_{k=0}^{|E|-1} \mathbf{e}'_k \quad (9)$$

$$\mathbf{u}' = \phi_u(\bar{\mathbf{u}}_v, \bar{\mathbf{u}}_e, \mathbf{u}) \quad (10)$$

The functions  $\phi_v, \phi_e, \phi_u$  are fully-connected neural networks. The model is trained with ordinary backpropagation, minimizing the mean squared error (MSE) loss between the predicted values in  $\mathbf{u}$  of the output graph, and the target values in the training dataset.

### Physics-based descriptors

To make a complete comparison, we also evaluate a classic setup, where physics-based descriptors are combined with a classic machine learning algorithm for tabular data, CatBoost<sup>36</sup>. The numerical features we extract from the crystal structures using the matminer package<sup>28</sup> are outlined in Table 4.



**Table 4.** Matminer feature groups used for a defect configuration description.

Features	Description
Density	Structure density, volume per atom, maximum packing fraction and maximum packing efficiency
Structural Complexity <sup>43</sup>	Shannon information entropy of a structure
XRD Powder Pattern <sup>44</sup>	Powder diffraction of a structure
Orbital Field Matrix <sup>45</sup>	Representation based on the valence shell electrons of neighboring atoms
JarvisCFID <sup>46</sup>	Pairwise radial, nearest neighbor, bond-angle, dihedral-angle and core-charge distributions

## DATA AVAILABILITY

The datasets analyzed during this study are available at <https://research.constructor.tech/p/2d-defects-prediction>.

## CODE AVAILABILITY

Code used to calculate the results of this study is available under Apache License 2.0 at [https://github.com/HSE-LAMBDA/ai4material\\_designIt](https://github.com/HSE-LAMBDA/ai4material_designIt) can be run online at <https://research.constructor.tech/p/2d-defects-prediction>.

Received: 23 May 2022; Accepted: 9 June 2023;

Published online: 26 June 2023

## REFERENCES

- Lin, Y.-C., Torsi, R., Geohagan, D. B., Robinson, J. A. & Xiao, K. Controllable thin-film approaches for doping and alloying transition metal dichalcogenides mono-layers. *Adv. Sci.* **8**, 2004249 (2021).
- Novoselov, K. S. et al. Two-dimensional gas of massless dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
- Aharonovich, I., Englund, D. & Toth, M. Solid-state single-photon emitters. *Nat. Photonics* **10**, pp.631–641 (2016).
- Frey, N. C., Akinwande, D., Jariwala, D. & Shenoy, V. B. Machine learning-enabled design of point defects in 2d materials for quantum and neuromorphic information processing. *ACS Nano* **14**, 13406–13417 (2020).
- Chanussot, L. et al. Open catalyst 2020 (oc20) dataset and community challenges. *ACS Catal.* **11**, 6059–6072 (2021).
- Wang, Z. et al. Novel 2d material from amqs-based defect engineering for efficient and stable organic solar cells. *2D Mater.* **6**, 045017 (2019).
- Bertoldo, F., Ali, S., Manti, S. & Thygesen, K. S. Quantum point defects in 2d materials—the qpod database. *npj Comput. Mater.* **8**, 56 (2022).
- Freysoldt, C. et al. First-principles calculations for point defects in solids. *Rev. Mod. Phys.* **86**, 253 (2014).
- Smith, J. S., Isayev, O. & Roitberg, A. E. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203 (2017).
- Stocker, S., Gasteiger, J., Becker, F., Günnemann, S. & Margraf, J. T. How robust are modern graph neural network potentials in long and hot molecular dynamics simulations? *Mach. Learn.: Sci. Technol.* **3**, 045010 (2022).
- 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.* **31**, 3564–3572 (2019).
- Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
- Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. SchNet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018).
- Gasteiger, J., Becker, F. & Günnemann, S. Gemnet: Universal directional graph neural networks for molecules. *Adv. Neural Inf. Process. Syst.* **34**, 6790–6802 (2021).
- Huang, P. et al. Unveiling the complex structure-property correlation of defects in 2d materials based on high throughput datasets. *npj 2D Mater. Appl.* **7**, 6 (2023).
- Park, C. W. & Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **4**, 063801 (2020).
- Klicpera, J., Groß, J. & Günnemann, S. Directional message passing for molecular graphs. Published at the International Conference on Learning Representations (ICLR) 2020. Preprint at <https://arxiv.org/abs/2003.03123> (2020).
- Shuaibi, M. et al. Rotation invariant graph neural networks using spin convolutions. Preprint at <https://arxiv.org/abs/2106.09575> (2021).
- Choudhary, K. & DeCost, B. Atomistic line graph neural network for improved materials property predictions. *npj Comput. Mater.* **7**, 1–8 (2021).
- Ying, C. et al. Do transformers really perform badly for graph representation? *Adv. Neural Inform. Process. Syst.* **34**, 28877–28888 (2021).
- Shi, Y. et al. Benchmarking graphormer on large-scale molecular modeling datasets. Preprint at <https://arxiv.org/abs/2203.04810> (2022).
- Sun, Q. & Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **49**, 2705–2712 (2016).
- Huang, P. et al. Carbon and vacancy centers in hexagonal boron nitride. *Phys. Rev. B* **106**, 014107 (2022).
- Deml, A. M., Holder, A. M., O'Hayre, R. P., Musgrave, C. B. & Stevanović, V. Intrinsic material properties dictating oxygen vacancy formation energetics in metal oxides. *J. Phys. Chem. Lett.* **6**, 1948–1953 (2015).
- Choudhary, K. & Sumpter, B. G. A deep-learning model for fast prediction of vacancy formation in diverse materials. Preprint at <https://arxiv.org/abs/2205.08366> (2022).
- Wexler, R. B., Gautam, G. S., Stechel, E. B. & Carter, E. A. Factors governing oxygen vacancy formation in oxide perovskites. *J. Am. Chem. Soc.* **143**, 13212–13227 (2021).
- Witman, M., Goyal, A., Ogitsu, T., McDaniel, A. & Lany, S. Materials discovery for high-temperature, clean-energy applications using graph neural network models of vacancy defects and free-energy calculations. Preprint at <https://chemrxiv.org/engage/chemrxiv/article-details/63b7181c1f24031e9a1789e0>.
- Ward, L. et al. Matminer: An open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).
- Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
- Manzoor, A. et al. Machine learning based methodology to predict point defect energies in multi-principal element alloys. *Front. Mater.* **8** <https://www.frontiersin.org/article/10.3389/fmats.2021.673574> (2021).
- Ostadossein, A. et al. Reaxff reactive force-field study of molybdenum disulfide (MoS<sub>2</sub>). *J. Phys. Chem. Lett.* **8**, 631–640 (2017).
- Patra, T. K. et al. Defect dynamics in 2-d MoS<sub>2</sub> probed by using machine learning, atomistic simulations, and high-resolution microscopy. *ACS Nano* **12**, 8006–8016 (2018).
- Banik, S. et al. Learning with delayed rewards—A case study on inverse defect design in 2d materials. *ACS Appl. Mater. Interfaces* **13**, 36455–36464 (2021).
- Shtyov, A. V., Abanin, D. A. & Levitov, L. S. Long-range interaction between adatoms in graphene. *Phys. Rev. Lett.* **103**, 016806 (2009).
- Ong, S. P. et al. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. CatBoost: unbiased boosting with categorical features. *Adv. Neural Inform. Process. Syst.* **31**, 6639–6649 (2018).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
- Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
- Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
- Battaglia, P. W. et al. Relational inductive biases, deep learning, and graph networks. Preprint at <https://research.google/pubs/pub47094/> (2018).
- Kostenetskiy, P. S., Chulkevich, R. A. & Kozyrev, V. I. HPC resources of the higher school of economics. *J. Phys.: Conf. Ser.* **1740**, 012050 (2021).
- Krivovichev, S. V. Structural complexity of minerals: information storage and processing in the mineral world. *Mineral. Mag.* **77**, 275–326 (2013).
- Ong, S. P. et al. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
- Lam Pham, T. et al. Machine learning reveals orbital interaction in materials. *Sci. Technol. Adv. Mater.* **18**, 756–765 (2017).
- Choudhary, K., DeCost, B. & Tavazza, F. Machine learning with force-field-inspired descriptors for materials: Fast screening and mapping energy landscape. *Phys. Rev. Mater.* **2**, 083801 (2018).

## ACKNOWLEDGEMENTS

This research/project is supported by the Ministry of Education, Singapore, under its Research Centre of Excellence award to the Institute for Functional Intelligent Materials (I-FIM, project No. EDUNC-33-18-279-V12). K.S.N. is grateful to the Royal Society (UK, grant number RSRP\R\190000) for support. This research was supported in part through computational resources of HPC facilities at HSE University<sup>42</sup>. The article was prepared within the framework of the project "Mirror Laboratories" HSE University, RF. This research has been financially supported by The Analytical Center for the Government of the Russian Federation (Agreement No. 70-2021-00143 dd. 01.11.2021, IGK 000000D730321P5Q0002). P.H. acknowledges the supports of the National Key Research and Development Program (2021YFB3802400) and the National Natural Science Foundation (52161037) of China. The computational work for this article was performed on resources at the National Supercomputing Centre of Singapore (NSCC) and NUS HPC. The research used computational resources provided by Constructor AG.

## AUTHOR CONTRIBUTIONS

N.K. conceived the sparse representation and conducted computational experiments with sparse representation, CatBoost, and SchNet; generated structures with high defect density. A.R.A.M. implemented EOS, conducted computational experiments with sparse representation using EOS, MEGNet, and GemNet architectures and is considered a 'co-first author'. I.R. implemented MEGNet in pytorch, and conducted computational experiments with it. MF&RL jointly generated the low-density structures with defects and performed initial computational experiments using CatBoost+matminer and SchNet. A.T. contributed to discussion. A.H.C.N. proposed EOS. P.H. did the DFT computations. K.S.N. and A.U. jointly supervised the work. All authors contributed to the debate and analyzes of the data, the writing of the paper, and approved the final version.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01062-z>.

**Correspondence** and requests for materials should be addressed to Kostya S. Novoselov, Pengru Huang or Andrey Ustyuzhanin.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023