Predicting Magnetic Properties of van der Waals Magnets using Graph Neural Networks

Peter Minch,[1] Romakanta Bhattarai,[1] Kamal Choudhary,[2] and Trevor David Rhone[1]

[1] *Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY, 12180*

[2] *Materials Science and Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899, USA*

(Dated: November 4, 2024)

We study two-dimensional (2D) magnetic materials using state-of-the-art machine learning models that use a graph-theory framework. We find that representing materials as graphs allows us to better learn structure-property relationships by leveraging both the chemical properties of the constituent atoms and the connectivity between those atoms. Graph neural network models are capable of predicting global properties of crystal structure (i.e. graph-wise properties) and local properties of the constituent atoms (i.e. node-wise properties). We embed physical constraints into our model by simultaneously making predictions of local and global properties. In particular, we use the Atomistic Line Graph Neural Network (ALIGNN) architecture. We train the ALIGNN model on data comprising local and global magnetic moments of 314 2D structures of the form $CrA^{ii}B^{i}B^{ii}X_6$, based on monolayer $Cr_2Ge_2Te_6$, calculated from first-principles. By learning the relationships between both local and global magnetic properties, we demonstrate an improvement over models that only consider global magnetic properties.

I. INTRODUCTION

Emergent phenomena play a fundamental role in the study of condensed matter physics [1, 2]. This is especially true in systems with reduced dimensionality, such as two-dimensional (2D) materials. Over the last two decades, the number of studies of 2D materials has exploded. In recent years, several 2D materials have been experimentally determined to exhibit long-range magnetic order [3–6]. These 2D magnets have a number of properties that make them ideal for applications in spintronics [7], data storage [8], and sensing [9].

For any given application, one must identify a material with properties that satisfy specific design criteria. However, because chemical space (i.e. all possible material candidates), can be enormous, conducting a search for a material with a set of desirable properties can be slow and expensive using serial experiments or first-principles calculations[10–12]. Thus, faster screening methods are necessary to accelerate materials design and discovery. We leverage data from first-principles calculations to train surrogate machine learning (ML) models [13, 14]. ML models can estimate materials properties several orders of magnitudes faster than first-principles calculations, and thus may be used to accelerate materials screening. Recently, graph neural networks are being increasingly used to predict materials properties [15, 16]. These models describe materials using graphs, where nodes represent the atoms in a crystal structure and the edges represent the chemical bonds. Graph neural networks have been used successfully in the prediction of formation energy [17, 18], electronic density of states [19], and phonon density of states [20]. The architecture of these models is well suited for describing material properties due to the fact that they encode properties of lattice sites, as well as the connectivity of the lattice sites. The model architecture facilitates the learning of complex material properties [21].

For 2D magnetic materials, it is important to identify the ground state magnetic order. Materials with different types of long-range magnetic orderings have vastly different applications [22–25]. Previous efforts to build models to identify the ground state magnetic order have treated the task as a classification problem [26–30]. Another approach is to predict the energy difference between competing magnetic configurations [31–34]. Other studies have focused on the prediction of total magnetic moment and applied a threshold to determine ferromagnetic order. [32, 35–38]. These approaches, however, are limited in their predictive power. The total magnetic moment approach is able to identify both ferromagnetic materials and materials with a large magnetic moment, however, these models are unable to distinguish between nonmagnetic and antiferromagnetic materials, which is crucial for the development of spintronic devices. The classification and energy difference approaches are able to distinguish between nonmagnetic, antiferromagnetic, and ferromagnetic materials, but they make no prediction regarding magnetic moment. Furthermore, traditional approaches do not account for the detailed atomic-level interactions that determine the magnetic properties, leading to oversimplified predictions. These limitations stem from the fact that these approaches treat magnetic ordering as a macroscopic or global property of the material. However, the difference between nonmagnetic and antiferromagnetic materials exists at the atomic level. Thus, it is necessary to make predictions with atomic resolution in order to simultaneously identify the magnetic ordering and magnetic moment.

Identifying magnetic ordering with microscopic resolution allows for finer detail in characterization of the magnetic properties of materials. For instance, to identify potential altermagnetic materials, it is necessary to

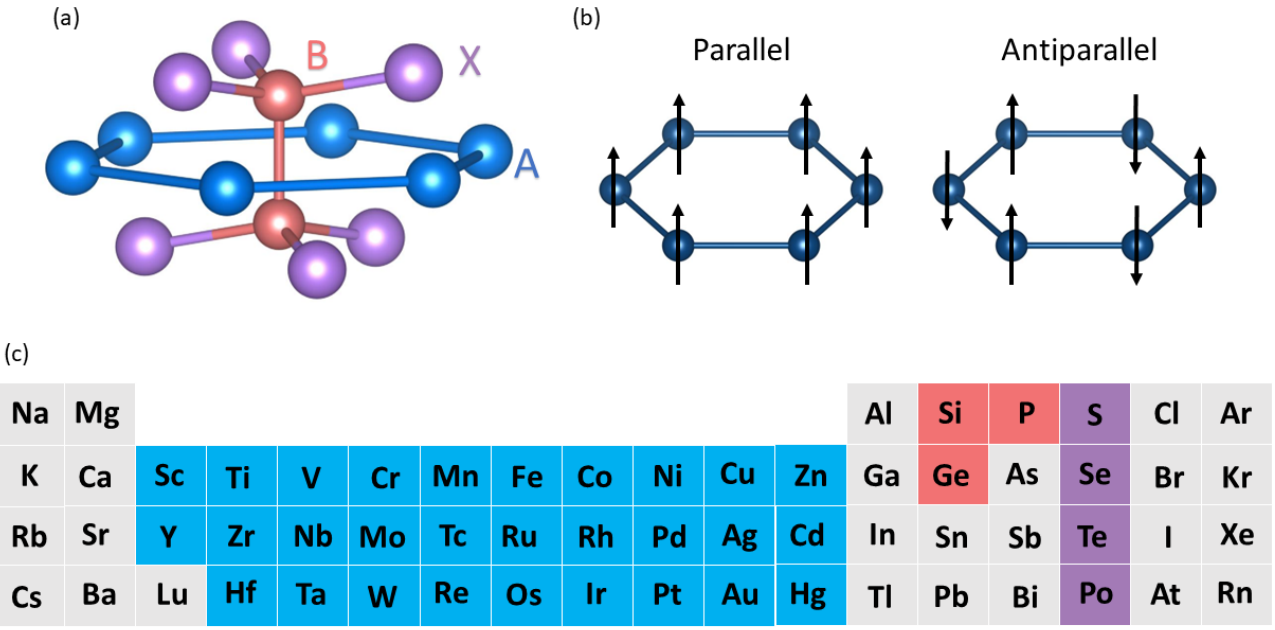

FIG. 1. (a) Side view of the structure of a $A^{i}A^{ii}B^{i}B^{ii}X_6$ lattice. (b) Spin configurations labelled parallel (P) and antiparallel (AP). Only the A sites are shown. (c) Elements used for chemical substitution at the A^{ii} site (blue), B sites (red), and X sites (magenta). The A^{i} site is decorated with Cr.

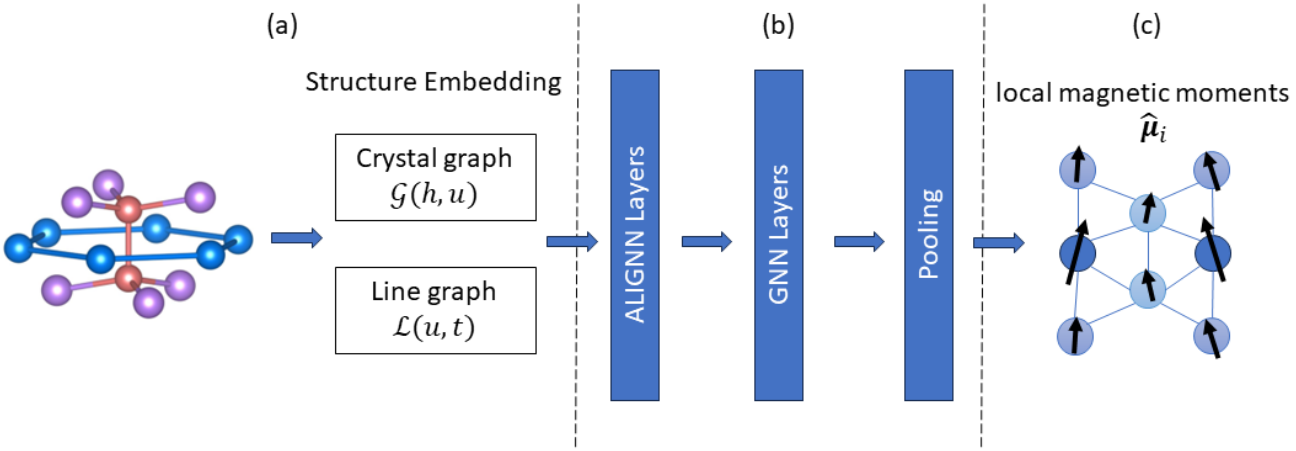

FIG. 2. Schematic of the ALIGNN model used to predict the magnetic moment. (a) The crystal structure is embedded into two graphs, the crystal graph $\mathcal{G}$ and the line graph $\mathcal{L}$. The crystal graph's nodes are labelled with atomic descriptor vectors $\boldsymbol{h}$ and edges are labelled with the bond feature vectors $\boldsymbol{u}$. The line graph's nodes are labelled with bond features $\boldsymbol{u}$ and edges are labelled with the bond feature vectors $\boldsymbol{t}$. (b) These graphs are updated via a series ALIGNN layers, GNN layers, and a fully connected layer that outputs (c) the local magnetic moment $\hat{\boldsymbol{\mu}}_i$ for each atomic site. The global magnetic moment $\hat{M}$ is calculated by taking the sum of these local magnetic moments.

distinguish between different types of antiferromangetic ordering [39]. Because previous machine learning methods have treated magnetic ordering as a global property, they are unable to make distinctions that require nanoscale resolution. Thus, we propose a new method for modelling magnetic ordering with atomic detail.

To address the limitations of previous methods, we use the atomistic line graph neural network (ALIGNN) architecture [17] in order to make microscopic predictions of magnetic properties. The ALIGNN architecture is well-suited to magnetic property prediction due to the fact that it encodes not only bond distances, but also bond angles. Bond angles play a fundamental role in determining the magnetic behavior in materials [40]. In addition, we train our model to simultaneously make predictions of the local magnetic moment of the constituent atoms and the global magnetic moment. By simultaneously making these predictions, we impose the constraint that the local magnetic moments must sum to the global magnetic moment. This constrains the model to more accurately capture the underlying physics. We find that through this framework, we not only improve prediction accuracy, but also provide a more detailed understanding of magnetic behavior at the atomic level. To demonstrate this, we apply this approach to predict of the local and global magnetic moments of 314 2D structures based on $Cr_2Ge_2Te_6$ calculated from first-principles.

II. METHODS

A materials database based on monolayer $Cr_2Ge_2Te_6$ was generated by making atomic substitutions at the lattice sites yielding crystal structures of the form $A^iA^{ii}B^iB^{ii}X_6$ (see Figure 1a,b). Here, A^i = Cr, while A^{ii} sites comprise various transition metals (see Figure 1c); B site substitutions are combinations of Ge, Si, or P; X sites are chalcogens. The total number of possible structures generated from these substitutions is $\sim 10^4$. Using first-principles, we calculate the magnetic properties of a subset of 314 of these structures. An example of a structure in this database is $CrVSiGeSe_6$. For each structure, two spin configurations are considered when determining the ground state: the spins on the A sites are either parallel or antiparallel, as shown in Figure 1b. These calculations result in a dataset containing 3140 atomic magnetic moment datapoints.

The calculations were carried out using the VASP package [41], which implements DFT based on the projector augmented wave (PAW) method. We used the GGA-PBE exchange functionals [42]. The plane-wave energy cutoff was set to 450 eV. We used a Γ-centered $10 \times 10 \times 1$ mesh to sample k-space. The self-consistent DFT calculations were converged to 10^{-6} eV. We utilized spin-polarized DFT in the noncollinear mode with spin-orbit coupling turned on for determining the ground state spin configuration. Additionally, the atomic positions and unit cell shape were relaxed until the forces on each atom were smaller than 10^{-2} eV/Å. It should be noted that commercial software is identified to specify procedures. Such identification does not imply recommendation by National Institute of Standards and Technology (NIST).

We use ALIGNN to accelerate predictions of material properties. In ALIGNN, a crystal structure is represented as a graph by mapping atoms at the lattice sites to nodes and bonds to the edges of the graph. Each node in the crystal graph is assigned 10 input node features based on its atomic species: electronegativity, group number, period number, covalent radius, number of s, p, and d valence electrons, first ionization energy, electron affinity, and the number of unpaired electrons. In addition, we add an indicator variable, labelling A^i sites with +1 and A^{ii} sites with -1. This indicator variable breaks the permutation symmetry of the model in the case where $A^i = A^{ii}$, allowing the model to predict antiparallel spin configurations [43]. The inter-atomic bond distances are used as edge features with radial basis function up to 8 Å cut-off and 12 nearest-neighbors (N). This crystal graph is then used for constructing the corresponding line graph that encodes interatomic bond-distances as nodes and bond-angles as edge features. ALIGNN uses edge-gated graph convolution for updating nodes as well as edge features using a propagation function (f) for layer (l), atom features (h), node (i), and neighboring nodes (N_i), details of which can be found in Refs. 17 and 44:

$$h_i^{(l+1)} = f(h_i^l, \{h_j^l\}_{j \in N_i}) \tag{1}$$

The hyperparameter of the model, including learning rate and batch size, were optimized using a 10-fold cross validation (see the Supplemental Material for details [45], see also references [46–48] therein).

There are two classes of targets that our model is trained on: Global (graph-wise) targets and local (atom-wise) targets. Global targets are properties of the crystal structure, such as the formation energy, total magnetic moment, and are predicted using the pooled graph-wise output of the model. Local targets are properties associated with each atomic site within a given structure, such as forces and atomic magnetic moments. Local targets are predicted using the node-wise outputs of the model.

More specifically, For a given structure S and atomic site i in the unit cell of S, we define the local magnetic moment on site i to be $\boldsymbol{\mu}_S^i$. The total magnetic moment for a unit cell of S is given by $M_S = |\sum_{i \in S} \boldsymbol{\mu}_S^i|$. We train our model using M_S as a global target and $\boldsymbol{\mu}_S^i$ as local targets. We define two loss terms for the global and local targets called GL and LL respectively. These loss functions are given by:

$$\text{GL} = \sum_S |\hat{M}_S - M_S| \tag{2}$$

$$\text{LL} = \sum_S \sum_{i \in S} |\hat{\mu}_S^{i;x} - \mu_S^{i;x}| + |\hat{\mu}_S^{i;y} - \mu_S^{i;y}| + |\hat{\mu}_S^{i;z} - \mu_S^{i;z}| \tag{3}$$

The predictions for the local and global targets are coupled by incorporating both terms in the model's total loss function. The total loss is given by:

$$\text{Total Loss} = \alpha \text{GL} + (1-\alpha)\text{LL} \tag{4}$$

Here $\hat{M}_S = |\sum_{i \in S} \hat{\boldsymbol{\mu}}_S^i|$ is the prediction for total magnetic moment per unit cell for structure S and $\hat{\mu}_S^{i;x}$, $\hat{\mu}_S^{i;y}$, $\hat{\mu}_S^{i;z}$ are the x, y and z components of the predicted local magnetic moments at site i. α is a free parameter that controls the weighting of the graph-level target versus the node-level targets. We choose this parameter through hyperparameter tuning (see Supplemental Material for details [45]).

We assess the model accuracy on a number of tasks: (I) The global magnetic moment (M). (II) The magnitudes of the local A-site magnetic moments (μ_A). (III) Local A-site magnetic moment z-component, $\hat{\mu}_S^{\text{A};z}$ vs. $\mu_S^{\text{A};z}$. The model may also be used to make predictions of the ground state spin configurations. This may be done by comparing the relative signs of the z components of the A sites. By doing this, we can determine the ground state spin configurations (see Supplemental Material for details [45]).

Candidate structures are obtained by decorating the lattice sites of the archetype crystal structure, $Cr_2Ge_2Te_6$. The resulting unrelaxed structure will be different from the structure after relaxing with DFT. We assume that the best model performance is obtained from the relaxed structure. However, relaxing each input structure using DFT would be computationally expensive. We would prefer to design an ML framework that does not require DFT calculations to generate the input to the model. For this reason, we investigate how well the unrelaxed structures perform as the input to the model. Thus, we train separate models on (i) a set of unrelaxed structures and (ii) a set of DFT relaxed structures. The model trained on the unrelaxed structures is labelled ALIGNN-unrelaxed and the model trained on the DFT relaxed structures is labelled ALIGNN-relaxed.

In addition, we attempted to relax decorated structures without DFT with the assumption that we will have better model performance. We explored two methods of approximating the relaxation. (i) We used the ALIGNN-FF model [44]. This model optimizes the atomic structure by predicting the classical force field of the structure. (ii) We trained a model to predict the relaxed unit cell area of the structure. This model consists of an ALIGNN model based on our dataset of CGT structures that is trained to predict the relaxed unit cell area of a unrelaxed structure. The lattice parameters of the unrelaxed structures are then scaled up to match the relaxed unit cell areas. A similar approach has been adopted in previous work as a starting point for DFT relaxation [49]. We train models using these approximately relaxed structures obtained via these methods, labelled (i) ALIGNN-FF-relax and (ii) ALIGNN-AS. We compare the performance of these models against ALIGNN-relaxed and ALIGNN-unrelaxed.

III. RESULTS AND DISCUSSION

We trained the ALIGNN-FF-relax, ALIGNN-AS, ALIGNN-unrelaxed, and ALIGNN-relaxed models outlined above. We tested the performance of these models on tasks (I)-(III) outlined above. The performance for these regression tasks is measured using the mean absolute error (MAE) and the R^2 coefficient.

The performance of the ALIGNN-unrelaxed model is shown in Figure 3. This model achieves a test R^2 score of 0.864 and an MAE of 0.27 μ_B for the prediction of global magnetic moment. Previous work that predicted global magnetic moments using chemical descriptors that used the same dataset achieved an MAE of 1.26 μ_B [32]. In Ref. 32, the model had difficulty generalizing across different X sites, performing well only for X = Te and poorly for X = Se and X = S.

The performance of the ALIGNN-unrelaxed model in predicting the global magnetic moment of X = Te, Se, and S is presented in Figure 4. The model generalizes well across the different X sites. Comparing these results to those from Ref. 32, we find that our model generalizes significantly better across different X sites. Ref. 32 found an increase in model performance of up to an MAE of 0.3 μ_B by focusing the model on X = Te as well as adding additional data points, totaling 262 structures of the form $CrA^{ii}B^{i}B^{ii}T_6$. The ALIGNN-unrelaxed model is able to match this performance for X = Te structures when trained on a mixed dataset of 314 structures of the form $CrA^{ii}B^{i}B^{ii}X_6$, of which only 81 structures have X = Te. We surmise this improvement in generalization is due to the geometric information encoded in graph-based models as well as the additional information learned through training on local magnetic moments.

Figures 3b and 3c show the ALIGNN-unrelaxed model performance in predicting the magnitudes of the A site magnetic moments and their z-components respectively. The model achieves a test R^2 score of 0.97 and an MAE of 0.11 μ_B for the prediction of the magnitudes of the A site magnetic moments. The model also achieves a test R^2 score of 0.98 and MAE of 0.13 μ_B for the prediction of the z-components of the A site magnetic moments. These metrics demonstrate that the model is able to make accurate predictions of the local magnetic moments.

We highlight several outliers in the prediction of μ_A in Figure 3b. These are predictions for the Mn atom in structures of the form $CrMnB_2X_6$ that have an antiparallel spin configuration. The true magnitudes for these atoms range from 3.7 to 4.2 μ_B. The model predictions, however, range from 1.3 to 2.2 μ_B. These outliers are likely due to the small amount of Mn atoms in the dataset. Notably, Mn is the only type of A site atom in the dataset that has a larger magnetic moment than the Cr atoms. It is likely that with a larger dataset,

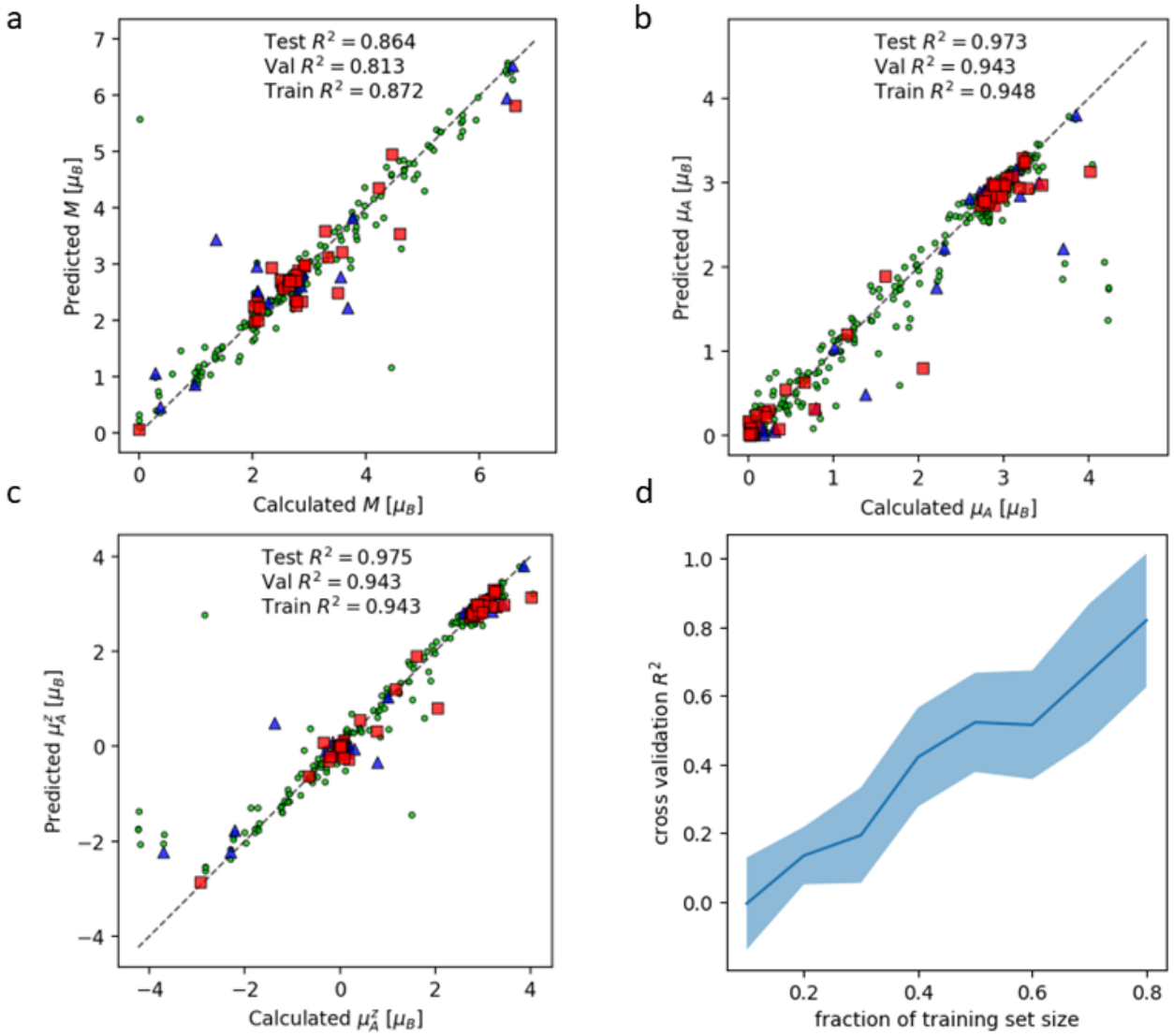

FIG. 3. ALIGNN-unrelaxed model performance on various magnetic property predictions. (a) Total magnetic moment, (b) A site magnitude, and (c) A site z-component of the magnetic moment in μ_B. Green circles indicate the training set, Blue triangles indicate the validation set, and red squares represent the test set. R^2 scores for the training, validation, and test sets are also shown. (d) Change in performance for the global prediction for varying training set size, measured via a 5-fold cross validation.

there would be sufficient structures with Mn atoms for the model to improve the prediction accuracy of these outliers. We note that Mn magnitude (μ_{Mn}) outliers in the A site magnitude predictions are not reflected in the total magnetic moment plot. This is due to the fact that despite getting μ_{Mn} incorrect, the model still predicts the correct total magnetic moment for these structures. The reason for this is that for the structures, the predicted value of $|\mu^z_{\mathrm{Cr}} + \mu^z_{\mathrm{Mn}}| \approx M$ is still close to the calculated values, despite the fact that the predicted value of μ_{Mn} is incorrect. We also note that we observe better predictions in $CrMnB_2X_6$ structures when the spin configurations are parallel, with the μ_{Mn} predictions ranging between 3.2 and 3.8 μ_B. When the spins are parallel, it is impossible for the model to predict the correct total magnetic moment, while predicting the local moments incorrectly. This highlights the usefulness of predicting the local and global moments together. The model has a tendency to underestimate μ_{Mn}, and will do so when its prediction is unconstrained, as it does in the antiparallel case. However, in the parallel case, the total magnetic moment constrains the prediction of the μ_{Mn}, guiding the model to make more accurate predictions.

The model performance for predicting the global mag-

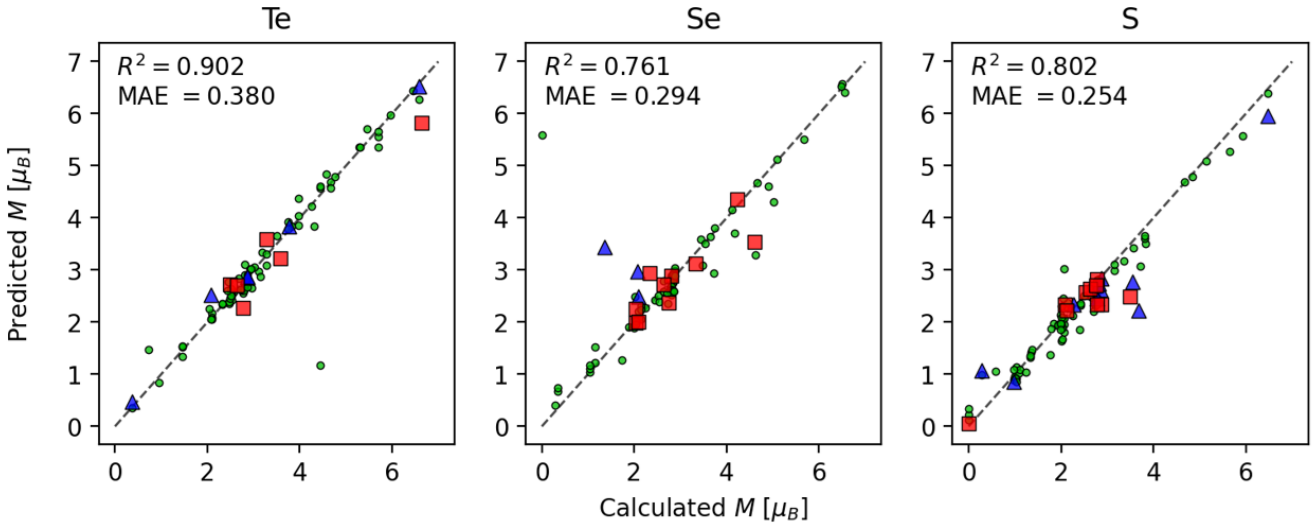

FIG. 4. Model performance for the predictions of the total magnetic moment, μ [μ_B] for X = Te, Se, and S are compared with the corresponding DFT calculations. Green circles indicate the training set, blue triangles indicate the validation set, and red squares represent the test set. R^2 score and MAE for the test set are shown.

netic moment increases with training set size, see Figure 3d. Notably, while the model performs well with a relatively small dataset, that is 314 structures with 3140 atomic magnetic moments, the model performance should improve with the addition of more data.

The relative performance for the ALIGNN-FF-relax, ALIGNN-AS, ALIGNN-unrelaxed, and ALIGNN-relaxed models is shown in Figure 5. One would expect that the performance of the model would improve as the structures used to create the input graphs get closer to the relaxed structures. That is, we would expect based on the assumption that an input relaxed structure would give the best result, the order of the models from most to least accurate to be ALIGNN-relaxed, ALIGNN-FF-relax, ALIGNN-AS, and ALIGNN-unrelaxed. However, we find that there is no appreciable difference in the performance of the four models. All four models perform within a margin of error from each other in all tasks. The structural information encoded in graphs based on the unrelaxed structures is sufficient to achieve the performance shown in Figure 3.

We also assess the model performance on the prediction of the local magnetic moment of the B and X sites (see Supplemental Material [45]). We find that the properties at these sites are much more difficult for the models to learn than those at A sites. This is likely due to the fact that the B and X site magnetic moments cover a range of between 0 and 0.2 μ_B, which is much smaller than the magnitude of the A site moments. Thus, the B and X sites do not contribute as much to the overall loss function due to the smaller role they play in determining the global magnetic moment. We do however obtain a reasonable prediction performance for the X site moments. The idealized ALIGNN-relaxed model achieves an average R^2 of 0.37 for predicting the X site magnitudes. We find that none of our models are able to learn to predict the B site moments, with all the R^2 scores for predicting B site magnitudes being negative. The fact that the model is successful for the X site moments, while failing for the B site moments indicates that the X sites play a more important role in the magnetic behavior of CGT-based structures. The greater importance of the X sites than the B sites in these structures has also been observed in a study of the anisotropy of these structures in Ref. 34. They suggest that this importance is the result of the hybridization between X and A site orbitals. This hybridization facilitates the super-exchange pathways between A sites and plays a fundamental role in determining the magnetic ground state of these structures [50].

TABLE I. Comparison between the ALIGNN-unrelaxed model's performance and various chemical descriptor based models for the total magnetic moment prediction task.

Model	Test R^2 score
ALIGNN	0.86
Linear	0.47
Adaboost	0.08
Kernel Ridge	0.37
Random Forest	0.34

To verify the ALIGNN model's improvement over more traditional machine learning methods, we tested several chemical descriptor based models. To form the chemical descriptors, we use the average, standard deviation, maximum, minimum, and maximum difference of each of

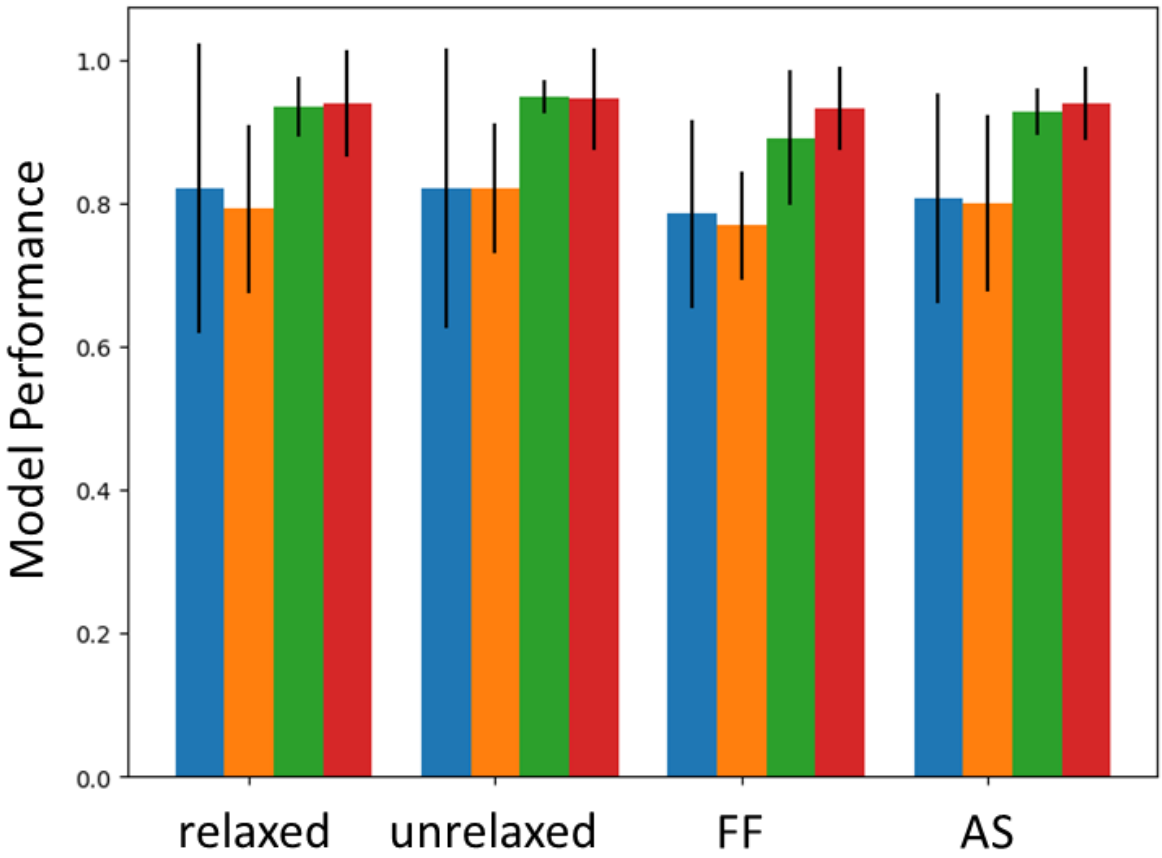

FIG. 5. Comparison of performance of the DFT-relaxed, unrelaxed, force-field (FF) relaxed, and area-scaled (AS) models in predictions for global magnetic moment (blue), spin configuration classification (orange), A site magnetic moment magnitude (green), and the z-component of the A site magnetic moment (red). Performance for regression tasks are assessed via a 10-fold cross-validation R^2 score. Performance for the classification task is measured via 10-fold cross-validation macro-averaged F_1 score.

the input node features described in the main text. We train a random forest model, Adaboost model, a kernel ridge regressor, and linear regressor on the total magnetic moment prediction task (see Supplemental Material for additional details [45]). We find that all these models are unable to achieve the level of accuracy of the ALIGNN model. A summary of the performance of these models is shown in Table I.

It should be noted that this framework may be extended to other datasets that include a variety of crystal structure types as well as bulk structures [17, 19, 20, 28]. This is because graph-based models are structure-agnostic and may be applied to any crystal system. We anticipate that by training on a dataset containing a wide variety of crystal structure types, the transferability of the model will increase.

There are, however, some limitations of the current model framework. Notably, if the ground state magnetic configuration has a nonzero propagation vector or is incommensurate with the unit cell, the model will be unable to capture these magnetic configurations. In these structures, the magnetic unit cell is larger than the crystallographic unit cell [51]. Since the model makes predictions per atom of the crystallographic unit cell, it is incapable of representing magnetic configurations of this kind.

One method to address this in future work may be to predict the magnetic propagation vector in addition to predicting the magnetic moment of each atom in the crystallographic unit cell [52]. This would allow the model output to represent a magnetic unit cell that is larger than the crystallographic unit cell. An alternate approach would be to not directly predict the magnetic configuration, but instead predict the exchange coupling

parameters of the crystal. This approach would allow the model to identify not only commensurate and incommensurate magnetic configurations, but also to identify more complex magnetic states such as spin textures and quantum spin liquids.

IV. CONCLUSIONS

In this work, we presented a study of 2D magnetic materials using the ALIGNN framework. In particular, we trained an ALIGNN model on DFT-calculated data comprising 314 2D structures of the form $CrA^{ii}B^{i}B^{ii}X_6$. We constrained the model by simultaneously making predictions on the global and local magnetic moments. By imposing the physical constraint that global moment prediction must equal the sum of the local moments, the model more accurately captures the underlying physics. We demonstrate that this method generalizes among different structures better than previous methods. We obtained improved prediction accuracy for a set of structures with different X sites.

V. CODE AVAILABILITY

The modified ALIGNN code for magnetic moments predictions as well as the data used in this work to train the model are available on Github at `https://github.com/quantum-intelligence/ALIGNN-magnetic-moments`.

ACKNOWLEDGMENTS

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This material is based upon work supported by the National Science Foundation CAREER award under Grant No 2044842. KC acknowledges funding from the CHIPS Metrology Program, part of CHIPS for America, National Institute of Standards and Technology, U.S. Department of Commerce.

[1] C. Gong and X. Zhang, Two-dimensional magnetic crystals and emergent heterostructure devices, Science **363**, eaav4450 (2019).

[2] Y. Li, B. Yang, S. Xu, B. Huang, and W. Duan, Emergent phenomena in magnetic two-dimensional materials and van der waals heterostructures, ACS Applied Electronic Materials **4**, 3278 (2022).

[3] B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden, *et al.*, Layer-dependent ferromagnetism in a van der waals crystal down to the monolayer limit, Nature **546**, 270 (2017).

[4] M. Bonilla, S. Kolekar, Y. Ma, H. C. Diaz, V. Kalappattil, R. Das, T. Eggers, H. R. Gutierrez, M.-H. Phan, and M. Batzill, Strong room-temperature ferromagnetism in vse2 monolayers on van der waals substrates, Nature nanotechnology **13**, 289 (2018).

[5] C. Gong, L. Li, Z. Li, H. Ji, A. Stern, Y. Xia, T. Cao, W. Bao, C. Wang, Y. Wang, *et al.*, Discovery of intrinsic ferromagnetism in two-dimensional van der waals crystals, Nature **546**, 265 (2017).

[6] Y. Deng, Y. Yu, Y. Song, J. Zhang, N. Z. Wang, Z. Sun, Y. Yi, Y. Z. Wu, S. Wu, J. Zhu, *et al.*, Gate-tunable room-temperature ferromagnetism in two-dimensional fe3gete2, Nature **563**, 94 (2018).

[7] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, Spin and pseudospins in layered transition metal dichalcogenides, Nature Physics **10**, 343 (2014).

[8] D. A. Wahab, M. Augustin, S. M. Valero, W. Kuang, S. Jenkins, E. Coronado, I. V. Grigorieva, I. J. Vera-Marun, E. Navarro-Moratalla, R. F. Evans, *et al.*, Quantum rescaling, domain metastability, and hybrid domain-walls in 2d cri3 magnets, Advanced Materials **33**, 2004138 (2021).

[9] V. O. Jimenez, V. Kalappattil, T. Eggers, M. Bonilla, S. Kolekar, P. T. Huy, M. Batzill, and M.-H. Phan, A magnetic sensor using a 2d van der waals ferromagnetic material, Scientific Reports **10**, 4789 (2020).

[10] Z. Yao, Y. Lum, A. Johnston, L. M. Mejia-Mendoza, X. Zhou, Y. Wen, A. Aspuru-Guzik, E. H. Sargent, and Z. W. Seh, Machine learning for a sustainable energy future, Nature Reviews Materials **8**, 202 (2023), publisher: Nature Publishing Group.

[11] D. Torelli, H. Moustafa, K. W. Jacobsen, and T. Olsen, High-throughput computational screening for two-dimensional magnetic materials based on experimental databases of three-dimensional compounds, npj Computational Materials **6**, 158 (2020).

[12] D. Wines, K. Choudhary, and F. Tavazza, Systematic dft+ u and quantum monte carlo benchmark of magnetic two-dimensional (2d) crx3 (x= i, br, cl, f), The Journal of Physical Chemistry C **127**, 1176 (2023).

[13] C. Xin, B. Song, G. Jin, Y. Song, and F. Pan, Advancements in High-Throughput Screening and Machine Learning Design for 2D Ferromagnetism: A Comprehensive Review, Advanced Theory and Simulations **6**, 2300475 (2023), _eprint:

https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.202300475.
[14] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. Billinge, *et al.*, Recent advances and applications of deep learning methods in materials science, npj Computational Materials **8**, 59 (2022).
[15] T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. **120**, 145301 (2018).
[16] V. Fung, J. Zhang, E. Juarez, and B. G. Sumpter, Benchmarking graph neural networks for materials chemistry, npj Computational Materials **7**, 1 (2021), publisher: Nature Publishing Group.
[17] K. Choudhary and B. DeCost, Atomistic Line Graph Neural Network for improved materials property predictions, npj Computational Materials **7**, 1 (2021), number: 1 Publisher: Nature Publishing Group.
[18] C. W. Park and C. Wolverton, Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery, Phys. Rev. Mater. **4**, 063801 (2020).
[19] P. R. Kaundinya, K. Choudhary, and S. R. Kalidindi, Prediction of the Electron Density of States for Crystalline Compounds with Atomistic Line Graph Neural Networks (ALIGNN), JOM **74**, 1395 (2022).
[20] R. Gurunathan, K. Choudhary, and F. Tavazza, Rapid prediction of phonon structure and properties using the atomistic line graph neural network (alignn), Phys. Rev. Mater. **7**, 023803 (2023).
[21] P. Reiser, M. Neubert, A. Eberhard, L. Torresi, C. Zhou, C. Shao, H. Metni, C. van Hoesel, H. Schopmans, T. Sommer, and P. Friederich, Graph neural networks for materials science and chemistry, Communications Materials **3**, 1 (2022), publisher: Nature Publishing Group.
[22] W. Han, Perspectives for spintronics in 2D materials, APL Materials **4**, 032401 (2016).
[23] N. S. Kiselev, A. N. Bogdanov, R. Schäfer, and U. K. Rößler, Chiral skyrmions in thin magnetic films: new objects for magnetic storage technologies?, Journal of Physics D: Applied Physics **44**, 392001 (2011).
[24] H. Y. Yuan, Y. Cao, A. Kamra, R. A. Duine, and P. Yan, Quantum magnonics: When magnon spintronics meets quantum information science, Physics Reports Quantum magnonics: When magnon spintronics meets quantum information science, **965**, 1 (2022).
[25] A. Hirohata, K. Yamada, Y. Nakatani, I.-L. Prejbeanu, B. Diény, P. Pirro, and B. Hillebrands, Review on spintronics: Principles and device applications, Journal of Magnetism and Magnetic Materials **509**, 166711 (2020).
[26] S. Lu, Q. Zhou, Y. Guo, Y. Zhang, Y. Wu, and J. Wang, Coupling a Crystal Graph Multilayer Descriptor to Active Learning for Rapid Discovery of 2D Ferromagnetic Semiconductors/Half-Metals/Metals, Advanced Materials **32**, 2002658 (2020), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202002658.
[27] C. M. Acosta, E. Ogoshi, J. A. Souza, and G. M. Dalpian, Machine Learning Study of the Magnetic Ordering in 2D Materials, ACS Applied Materials & Interfaces **14**, 9418 (2022), publisher: American Chemical Society.
[28] A. Elrashidy, J. Della-Giustina, and J.-A. Yan, Accelerated Data-Driven Discovery and Screening of Two-Dimensional Magnets Using Graph Neural Networks, The Journal of Physical Chemistry C 10.1021/acs.jpcc.3c07246 (2024), publisher: American Chemical Society.
[29] S. Lu, Q. Zhou, Y. Guo, and J. Wang, On-the-fly interpretable machine learning for rapid discovery of two-dimensional ferromagnets with high Curie temperature, Chem **8**, 769 (2022).
[30] H. A. Merker, H. Heiberger, L. Nguyen, T. Liu, Z. Chen, N. Andrejevic, N. C. Drucker, R. Okabe, S. E. Kim, Y. Wang, T. Smidt, and M. Li, Machine learning magnetism classifiers from atomic coordinates, iScience **25**, 105192 (2022).
[31] A. N. Mahmoodabadi, M. Modarresi, M. R. Roknabadi, and A. Mogulkoc, Efficient discovery of room temperature magnetic transition metal monolayers assisted by artificial neural network, Computational Materials Science **224**, 112166 (2023).
[32] T. D. Rhone, W. Chen, S. Desai, S. B. Torrisi, D. T. Larson, A. Yacoby, and E. Kaxiras, Data-driven studies of magnetic two-dimensional materials, Scientific Reports **10**, 10.1038/S41598-020-72811-Z (2020), publisher: Nature Research.
[33] P. Minch, R. Bhattarai, and T. D. Rhone, Data-driven study of magnetic anisotropy in transition metal dichalcogenide monolayers, Solid State Communications **371**, 115248 (2023).
[34] Y. Xie, G. A. Tritsaris, O. Grånäs, and T. D. Rhone, Data-Driven Studies of the Magnetic Anisotropy of Two-Dimensional Magnetic Materials, The Journal of Physical Chemistry Letters **12**, 12048 (2021).
[35] R. Bhattarai, P. Minch, and T. D. Rhone, Investigating magnetic van der Waals materials using data-driven approaches, Journal of Materials Chemistry C **11**, 5601 (2023), publisher: The Royal Society of Chemistry.
[36] I. Miyazato, Y. Tanaka, and K. Takahashi, Accelerating the discovery of hidden two-dimensional magnets using machine learning and first principle calculations, Journal of Physics: Condensed Matter **30**, 06LT01 (2018), publisher: IOP Publishing.
[37] A. Dutta and P. Sen, Machine learning assisted hierarchical filtering: a strategy for designing magnets with large moment and anisotropy energy, Journal of Materials Chemistry C **10**, 3404 (2022), publisher: The Royal Society of Chemistry.
[38] T. D. Rhone, R. Bhattarai, H. Gavras, B. Lusch, M. Salim, M. Mattheakis, D. T. Larson, Y. Krockenberger, and E. Kaxiras, Artificial Intelligence Guided Studies of van der Waals Magnets, Advanced Theory and Simulations **6**, 2300019 (2023), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.202300019.
[39] J. Sødequist and T. Olsen, Two-dimensional altermagnets from high throughput computational screening: Symmetry requirements, chiral magnons, and spin-orbit effects, Applied Physics Letters **124**, 182409 (2024), https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/5.0198285/19920273/182409_1_5.0198285.pdf.
[40] M.-C. Wang and C.-R. Chang, Goodenough-Kanamori-Anderson Rules in CrI3/MoTe2/CrI3 Van der Waals Heterostructure, Journal of The Electrochemical Society **169**, 053507 (2022), publisher: IOP Publishing.
[41] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).
[42] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Physical Review

Letters **77**, 3865 (1996), publisher: American Physical Society.

[43] T. E. Smidt, M. Geiger, and B. K. Miller, Finding symmetry breaking order parameters with euclidean neural networks, Phys. Rev. Res. **3**, L012002 (2021).

[44] K. Choudhary, B. DeCost, L. Major, K. Butler, J. Thiyagalingam, and F. Tavazza, Unified graph neural network force-field for the periodic table: solid state applications, Digital Discovery **2**, 346 (2023), publisher: RSC.

[45] See Supplemental Material at [URL by publisher] for further details regarding model training; information regarding the prediction of ground state spin configurations; analysis regarding the multicollinearity of input node features; and additional plots showing the predictions of traditional machine-learning models, the predictions of X and B site magnetic moments, and a comparison between the unit cell parameters obtained from the various surrogate relaxation models.

[46] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2019).

[47] R. D. De Veaux and L. H. Ungar, Multicollinearity: A tale of two nonparametric regressions, in *Selecting Models from Data*, edited by P. Cheeseman and R. W. Oldford (Springer New York, New York, NY, 1994) pp. 393–402.

[48] J. Y.-L. Chan, S. M. H. Leow, K. T. Bea, W. K. Cheng, S. W. Phoong, Z.-W. Hong, and Y.-L. Chen, Mitigating the multicollinearity problem and its machine learning approach: a review, Mathematics **10**, 1283 (2022).

[49] J. Schmidt, N. Hoffmann, H.-C. Wang, P. Borlido, P. J. M. A. Carriço, T. F. T. Cerqueira, S. Botti, and M. A. L. Marques, Machine-Learning-Assisted Determination of the Global Zero-Temperature Phase Diagram of Materials, Advanced Materials **35**, 2210788 (2023), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202210788.

[50] N. Sivadas, M. W. Daniels, R. H. Swendsen, S. Okamoto, and D. Xiao, Magnetic ground state of semiconducting transition-metal trichalcogenide monolayers, Phys. Rev. B **91**, 235425 (2015).

[51] J. Rodríguez-Carvajal and J. Villain, Magnetic structures, Comptes Rendus. Physique **20**, 770 (2019).

[52] H. A. Merker, H. Heiberger, L. Nguyen, T. Liu, Z. Chen, N. Andrejevic, N. C. Drucker, R. Okabe, S. E. Kim, Y. Wang, *et al.*, Machine learning magnetism classifiers from atomic coordinates, Iscience **25** (2022).