



HAL
open science

Electronic Configurations of 3d Transition-Metal Compounds Using Local Structure and Neural Networks

Wenhao Zhang, David Berthebaud, Jean-François Halet, Takao Mori

► **To cite this version:**

Wenhao Zhang, David Berthebaud, Jean-François Halet, Takao Mori. Electronic Configurations of 3d Transition-Metal Compounds Using Local Structure and Neural Networks. *Journal of Physical Chemistry A*, 2022, 126 (40), pp.7373-7381. 10.1021/acs.jpca.2c03901 . hal-03816512

HAL Id: hal-03816512

<https://hal.science/hal-03816512>

Submitted on 16 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Electronic Configurations of 3d Transition-Metal Compounds Using Local Structure and Neural Networks

Wenhao Zhang,^{†,‡} David Berthebaud,[¶] Jean-François Halet,^{*,¶} and Takao Mori^{*,†,‡}

[†]*WPI-MANA, National Institute for Materials Science (NIMS), Namiki 1-1, Tsukuba 305-0044, Japan.*

[‡]*Graduate School of Pure and Applied Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba 305-8671, Japan*

[¶]*CNRS-Saint-Gobain-NIMS, IRL 3629, Laboratory for Innovative Key Materials and Structures (LINK), National Institute for Materials Science (NIMS), Namiki 1-1, Tsukuba 305-0044, Japan.*

E-mail: Jean-Francois.Halet@univ-rennes1.fr; MORI.Takao@nims.go.jp

Abstract

Machine learning (ML) methods extract statistical relationships between inputs and results. When the inputs are solid-state crystal structures, structure-property relationship can be obtained. In this work, we investigate whether a simple neural network is able to learn the $3d$ orbital occupations for the transition-metal (TM) centers in crystalline inorganic solid-state compounds using only the local structure around the transition-metal centers described by rotational invariant fingerprints based on spherical harmonics and one-hot elemental encoding. A multilayer neural network trained on density functional theory (DFT) results of about 1800 samples was developed showing good performance in predicting the TM orbital occupations (for both spin channels). We study in detail how the local structure affects the predictions of the local properties, and how they provide physical insights for the design of future machine learning model for materials chemistry. The proposed ML method is illustrated in practical application by predicting local magnetic moments of the transition-metal atoms in a full set of inorganic structures with large unit cells. Although less accurate compared to the experimental data, the ML results compared well with the DFT results, suggesting the feasibility of electronic property prediction based only on structure input.

INTRODUCTION

Although the physical phenomena for various properties of transition-metal compounds are well understood, it is still difficult today to suggest chemical compositions and crystal structures of transition-metal compounds with desired properties.¹⁻⁴ Given a structure, these properties are available recently thanks to advances in materials simulation via first-principles methods and various analysis tools.⁵⁻⁹ Unfortunately, this often incurs a high computational cost that limits effectiveness in exploring large search spaces and complex compounds with large unit cells. Nevertheless, it is clear that, due to the relatively localized behavior of the d orbitals, chemical interactions between $3d$ transition-metal elements and their neighboring atoms in these compounds are relatively easy to understand and play a central role in determining their electronic properties, such as band dispersion and density of states (DOS).¹⁰⁻¹²

On the other hand, machine learning (ML) statistical methods have been demonstrated as a powerful alternative to DFT calculations, offering the possibility to speed up materials' discovery by learning patterns within large data sets and proposing novel candidates that are statistically likely to have the desired properties.¹³⁻¹⁵ Indeed, ML models can be highly efficient, portable and able to make prediction for materials properties with good accuracy,¹⁶ assuming that a standard of "best practices" is followed.¹⁷

Chemical environments in transition-metal compounds offer a good playground where ML structure-property relationships can be examined with our chemical intuition and interpretation.^{10,11,18} Crystal-and-ligand field theory considering only local environments enabled simple theoretical analysis of the band dispersion and DOS in transition-metal compounds long before that first-principles calculations became a common practice.^{10,11,18} Therefore, the correspondingly formulated machine learning problem and its results should be relatively simple and easy to analysis with our chemical understanding.

Good ML prediction of materials properties strongly relies on both the learning algorithms and on the representation of the input data.¹⁹ For the latter, two main approaches can be used. One featurizes compounds by their composition and the individual properties of

the contained elements. The other featurizes compounds using only their structural arrangements. It is naturally desirable to use the latter approach to study the structure-property relationships directly, but encoding materials' atomic neighborhood environments efficiently is not a straightforward task. Nevertheless, some progress has been made in this direction over the last years, and we summarize here a few of the most important works published in this area. Bartok et al. pioneered the local description of chemical systems with different methods to represent atomic neighborhood environments.^{20,21} ML frameworks using these representations were recently developed to estimate electronic^{22,23} and phononic²⁴ density of states in condensed matter. Faber et al. compared various structure representations and ML models for getting properties of molecular systems.¹⁶ Schütt et al. tried to predict the electronic DOS at the Fermi level in crystalline materials using a simple radial distribution function and regression methods.²⁵ Later, the same authors developed a deep learning approach that enables spatially and chemically resolved insights into quantum-mechanical observables of organic molecules using purpose-designed deep neural networks that consider atoms and bonds as nodes and edges in a graph.^{26,27} Somewhat similarly, Xie and Grossman proposed a crystal graph convolutional neural network to directly learn material properties from the connection of atoms in the crystal, providing a universal and interpretable representation of crystalline materials. Their approach was successfully applied to the design of novel perovskites.²⁸ Equivariant tensor field neural networks were also proposed that further preserve rotational properties using symmetry group representations.²⁹

Although the recent graph-based networks²⁶⁻²⁸ are very powerful, they typically use atomic distances and structure connectivity as inputs, which does not describe well the local structure that is important for the transition-metal atoms. Following the insights from local crystal field and ligand field theory, we describe herein an ML model that is based on strictly local structure representations (fingerprints) developed by Uhrin³⁰ and trained to predict the electronic occupation of up and down spin channels of transition-metal $3d$ orbitals in solid-state compounds. We report the performance of the machine learning model. We also put

emphasis on a detailed analysis of the ML learning errors in complex chemical environments.

THEORY

We first provide a brief overview of the quantum chemistry theory of electronic bands formed from transition-metal orbitals to show that our ML problem mapping the local structure to the electronic configurations of the transition metals is straightforward.^{11,18}

In simple coordination environments, the formation of the energy bands from atomic orbitals can be considered in two steps: firstly, degenerated atomic $3d$ orbitals are perturbed by the presence of neighboring anions. Different local symmetries lead to different energy level splittings for the $3d$ orbitals, as shown for three cases in Figure 1 (b). In the second step, periodicity of the crystal is introduced and Bloch functions, modulated by the reciprocal vector k can be formed. The band structure is therefore given by the dispersion $E(k)$. The positions in energy of the $3d$ bands are determined by the energies of the MOs of the building motifs while the width of their dispersion is governed by the interaction between building motifs. If the interaction is small, then the bands formed from the $3d$ orbitals of the transition metals will be narrow or even localized. This is generally true if the metal–metal distance is relatively large, then, the DOS and occupation of the TM $3d$ orbitals are largely determined by its nearest neighbors (local environment), which consequently determine the electron occupations of the TM $3d$ orbitals.

The above qualitative analysis is limited in two ways: it is unrealistic to perform such a treatment for thousands of structures and, for more general cases, distortion from ideal symmetry or varying ligand field strength will be case-specific and will affect the energy of the MO levels quantitatively (see Figure 1 (b) for an elongated octahedron). ML can be expected to capture these quantitative effects and make accurate on-the-fly prediction for $3d$ orbital occupation on the metal centers. Since the theoretical background in our ML formulation is the analytic theory outlined above, we aim to achieve a similar prediction

accuracy in $3d$ orbital occupations using only crystal structures as input.

Nearest-neighbor effects are the most important ones for the electronic configurations of the transition metal atoms. Although less important, atoms that are outside the nearest-neighbor environment also affect the energies of the $3d$ orbital-derived bands. However, it should be noted that by taking more atoms into consideration (such as secondary neighbors), we sacrifice the simplicity of the structure-property relationship we want the ML model to comprehend. Therefore, it is important to study how many surrounding atoms, apart from the nearest neighbors, is necessary to take into account to give a good prediction of the electronic configurations of the transition-metal centers. We answer this question by comparing different choices as input for ML in this work.

Methods

Dataset

The entire workflow of training and application of ML method is shown in Figure 2 (a). The only input for this study is crystal structures, which were gathered from either Materials Project (MP) Database³¹ or from the Crystallography Open Database (COD).³² H, noble gases, and rare-earth elements were excluded when we searched for compounds in the databases. From all the available compounds containing transition-metal elements, we only selected those that contain $3d$ transition-metal element the nearest neighbors of which are main group elements in the p block of the Periodic Table. Such a consideration simplifies the transition-metal bonding that the ML needs to identify in this work by excluding strong metal-metal interactions, as discussed previously. Furthermore, B, C, N, O and F elements were also excluded in our study for the following two reasons: 1) the number of these compounds in the database dominates the dataset and could lead to biased training set, and 2) the correlation effects on transition metal are more important in the narrow d bands of nitrides or oxides and large errors are expected from a PBE treatment in DFT calculation.³³

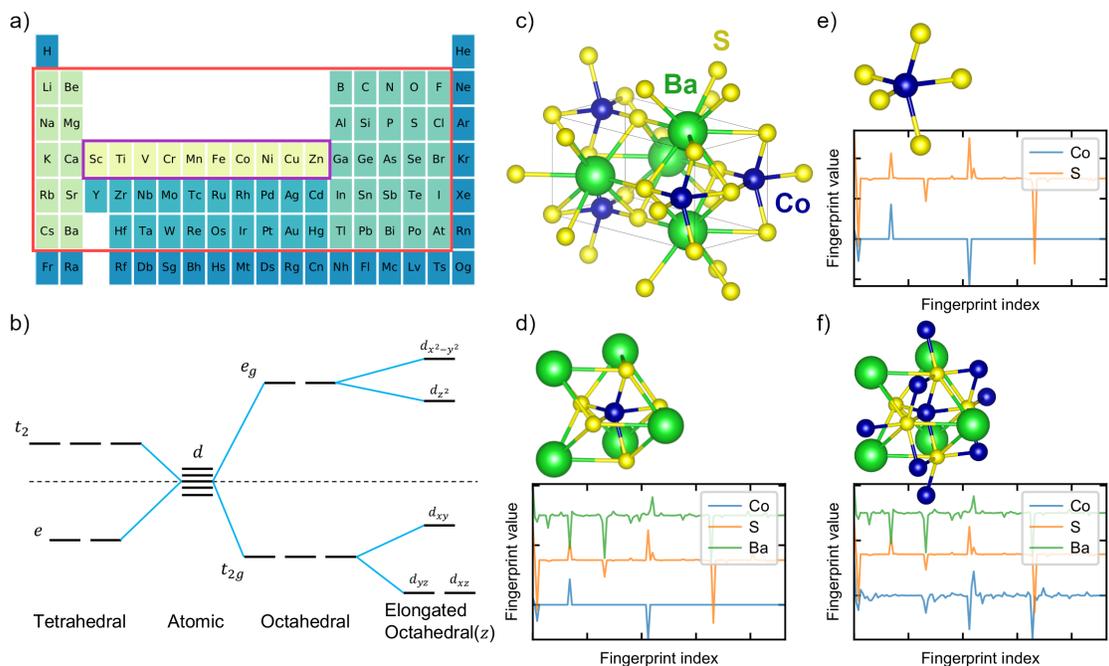


Figure 1: (a) Part of the Periodic Table showing the element set: transition metals (yellow), simple metals (light green) and main group anions (blue-green), 65 elements inside the red box are used for one-hot encoding. (b) Schematic illustration of how the symmetry of local polyhedron and distortion determine the energy level splitting of the transition-metal d states. (c) Crystal structure of the example compound BaCoS_2 (d)–(f) Different methods to represent the local structure of Co atoms in the example structure: (d) “Voronoi-bonded”, (e) “First-shell” and (f) “Cut-off” (5 Å) with their respective fingerprints.

Finally, we restricted the number of atoms per unit cell to be less than 30 to reduce data acquisition time by DFT calculations.

DFT results were obtained for a total of ca. 1800 compounds that satisfied the above criteria. In terms of elemental distribution in this training dataset, the number of compounds containing Cu is slightly larger than the number of compounds containing other transition-metal elements, which are distributed more or less evenly. For anions, S, Se and Te are the most frequent ones, followed by other anions such as Si, Ge, P, Sn and Cl. Overall, the dataset used does not bias heavily towards any particular metals or anions. A detailed description of how structures were gathered and their properties, including metal, anion distribution and their coordination, can be found in the *Supplementary Information* (SI).

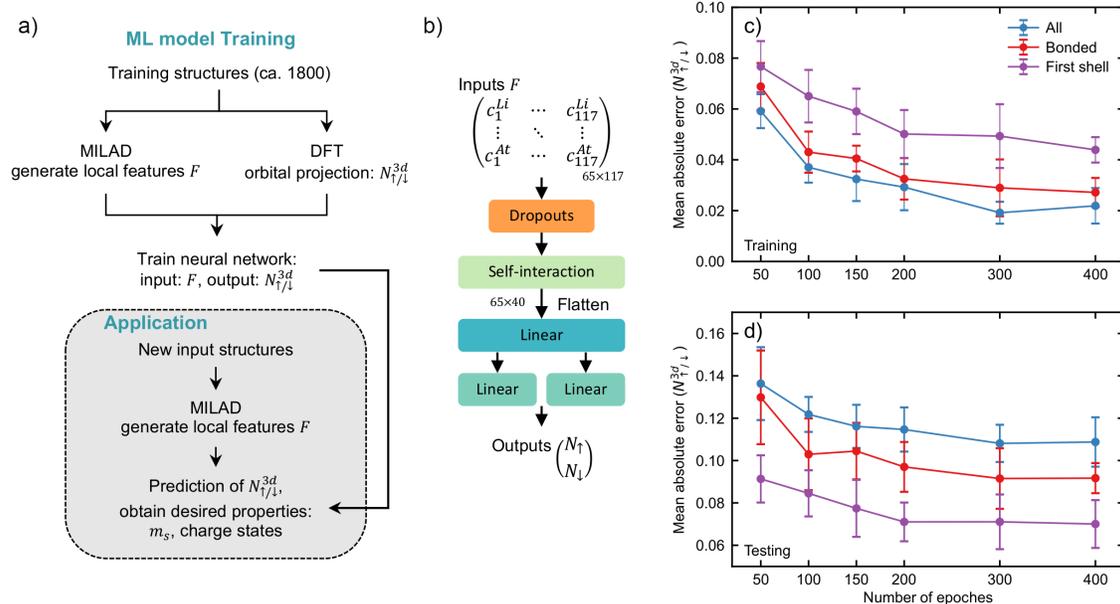


Figure 2: (a) Workflow divided into training and application stages. The input for the model is the local structure fingerprints F generated using MILAD package and one-hot encoding, and the target output is the DFT calculated $3d$ orbital occupations $N_{\uparrow/\downarrow}^{3d}$. (b) Detailed structure of the neural network model. Input matrices have the shape 65×117 , which is reduced to 65×40 dimension by self-interaction layers. The matrices are flattened and introduced into fully connected linear layers to give final outputs for the orbital occupation of spin-up and spin-down channels. Training (c) and testing (d) MAE of a 10-fold cross validation with different local environment descriptions as function of the number of epochs. The error bars show standard deviations.

DFT Calculations

As noted in previous studies,^{34,35} different calculation schemes and functionals can lead to errors ranging from 1% to 10% in calculated properties such as band gap and total magnetization, which is on the same magnitude of ML prediction errors.¹⁶ To have consistent results between model predictions and later DFT validation, we prepared training data from in-house spin polarized calculations, using the *Quantum Espresso* code^{36,37} with the GGA PBEsol functional³⁸ from Pseudo-Dojo³⁹ with suggested energy cutoffs. Hybrid functionals usually give more accurate results for the electronic structure of transition metal compounds, but they are forbiddingly expensive to use. On the other hand, although meta-GGA functionals improves upon GGA functionals in calculating band gap values,⁴⁰ it is not clear

whether they systematically outperform GGA functionals in predicting other properties, for example, magnetic moment.⁴¹ Structures for DFT calculations were used without relaxation. Comparing a few relaxed vs. non-relaxed structures, we observed that the effects on orbital occupations are relatively minor (see SI). Other details of calculations are provided in the corresponding sections in SI. The orbital occupations of the transition-metal d states were calculated using Löwdin population analysis and projection for both spin channels.⁴² Finally, it should be mentioned that although all DFT calculations were performed within a ferromagnetic configuration, we do not expect the spin moments to change dramatically imposing an antiferromagnetic configuration, since we assume that magnetic structure is determined by spin cooperative interactions, and spin moments are governed by the local coordination.⁴³

Structure Representation

Since we focus on transition metals in their local coordination, the structure representation should faithfully describe the position of the neighboring atoms.

In this work, a two-step procedure is used to generate the local structure representations, as shown in Figure 1 (c)–(f). First, we need to identify the local atoms surrounding the transition metal atom to be used as inputs for local structure representation. We tried three different strategies that we call “first-shell”, “Voronoi-bonded” and “cut-off”. The “first-shell” method considers only the nearest neighboring atoms in the covalent sphere while the “cut-off” method includes all atoms within a cut-off radius of 5 Å from the transition-metal atom, a distance beyond which chemical interactions between atoms should be negligible. The “Voronoi-bonded” approach sits in between the above two methods since it includes all the nearest neighbors and also atoms that share common boundaries upon partitioning the inter-atomic space, measured in terms of solid angle.^{44,45}

Next, to encode the structure arrangement and elemental information, we used the recently introduced structure representations by Uhrin,³⁰ which is a set of linear combinations of spectral coefficients that are rotational (rotation of the reference frame) and permutational

(shuffle of atomic index) invariants. It encodes local structure by projecting the atomic density function $f_c(\mathbf{x}) = \sum_i f(\mathbf{x} - \mathbf{r}_i^c)$ to a set of rotational invariant basis functions built from spherical harmonics. Here, \mathbf{r}_i^c is the position of the i th atom of element type c , function f in the summation can be a delta or Gaussian function providing the atomic contribution to the density function f_c . As shown in SI, a simple neural network can extract angular and distance distortions perfectly using the structure fingerprints generated from the distorted structure, showing that this representation encodes local geometric faithfully. This is in contrast to the graph-based networks^{26–28} typically built from only inter-atomic distances, which does not encode angular information. Finally, elemental information is encoded by one-hot encoding. The element set for one-hot encoding is shown in Figure 1 (a).

We take BaCoS₂ compound⁴⁶ as an example for generating local structure fingerprints with one-hot encoding in Figure 1: We first identify the non-equivalent transition-metal sites in the structure (one in this case) and its neighbors. Every Co atom is surrounded by a square pyramid of S atoms which constitute its first neighbors, Ba atoms as second neighbors and other Co atoms further out. Each kind of element is considered as a set of points for which structure fingerprints can be obtained separately. These fingerprints are put in rows in an input feature matrix (65×117, 117 fingerprints for each of the 65 elements taken from the Periodic Table). A row would be zeros if the element does not appear in the specific local structure. Although we did not include compounds containing B–F elements in the dataset, we included these elements in the one-hot encoding for uniformity.

The same procedures were carried out for all the non-equivalent transition-metal sites identified in the calculated structures in the preprocessing stage, which form the inputs for the ML model.

Model

We used a multilayer neural network^{47,48} to learn the electronic occupations from the above obtained descriptors, since neural networks are easy to build, flexible with respect to inputs

and are found to give good prediction accuracy in similar tasks.¹⁶ The design of the network is shown in Figure 2 (b).

As described above, the input to the network are rotational permutational invariant feature matrices. The first two layers consist of self-interaction for fingerprints that aims to extract element-specific structure features from the inputs and to reduce the number of total connections in the following layers. Linear layers are fully connected and diverge to give prediction of the occupation of the spin-up and spin-down channels. We used squared loss as loss function. Random drop-out and weight decay are used to improve the performance of the network⁴⁸ while training epochs are determined from 10-fold cross validation. Information on training parameters and further details of the neural network used in this work are provided in SI.

RESULTS

Evaluation of the Structure Description

To compare the different strategies of representing local structures, mean absolute errors (MAE) were evaluated using a 10-fold cross validation⁴⁷ for the models trained with different inputs. Cross validation is necessary to achieve consistent results because our data set is small and the random partitioning of the test set makes it difficult to compare the predictions systematically. The average error and the standard deviations of the training and testing sets over the models trained with each partition are shown in Figure 2 (c) and (d).

A clear trend can be observed from the errors in Figure 2, namely, the more atoms that are considered in the local structure representation, the better the fitting performance (training error) but the more it suffers from the over-fitting problem. In the end, the best generalization results were achieved by the network trained with only the information on the nearest neighboring atoms (“first-shell”), making better predictions on the unseen samples and leading to the lowest MAE in testing at 0.07-electron occupation.

The following conclusion can be drawn from the above result. Firstly, it again emphasizes that chemical and structure configurations of the nearest neighbor atoms play the dominant role in determining the electronic state of the transition-metal atoms, confirming the theory we outlined earlier, since using the nearest neighbor approach alone achieves a decent prediction. Secondly, it is easier for our neural network to learn the effect of the nearest neighbors of the transition-metal atoms than that of the higher order interactions from the secondary neighbors (over-fitting implies low learning quality). There can be multiple contributing effects: the configuration space of the secondary neighbors is much larger and more complex than that of the nearest neighbors, since the secondary neighbors can consist of simple ionic *s*-metals, transition metals in the neighboring cell, or even covalently bonded main group atoms in different coordinations. On the contrary, the chemical and structure configurations of the nearest neighbors are much simpler. As a result, the effects of the secondary neighbors are much more difficult to learn with a limited data set. It is also possible that the simple structure of our neural network used here limits the learning performance with more input information such as related to secondary neighbors, but it is not clear how this can be improved. This will be studied in future works.

Analysis of Errors

We dedicate this part to analyze the different sources of errors in our predictions. The errors reported in Figure 3 are the errors of all test sets in the 10-fold cross validation, each predicted by the ML trained with that set left out.

In Figure 3 (a) we show the prediction error grouped into three categories in terms of the chemical formula. We have three categories in increasing order of chemical complexity: 1) compounds containing only transition-metal atoms (TM) and main group anions (MG) are included, i.e., most of the binary compounds or compounds with mixed anions, 2) compounds that contain TM, MG and *s*-block metal atoms, and finally, 3) other compounds that can contain additional heavier *d*-block metals (Ag, Cd, Y, Zr, Nb and La being the common

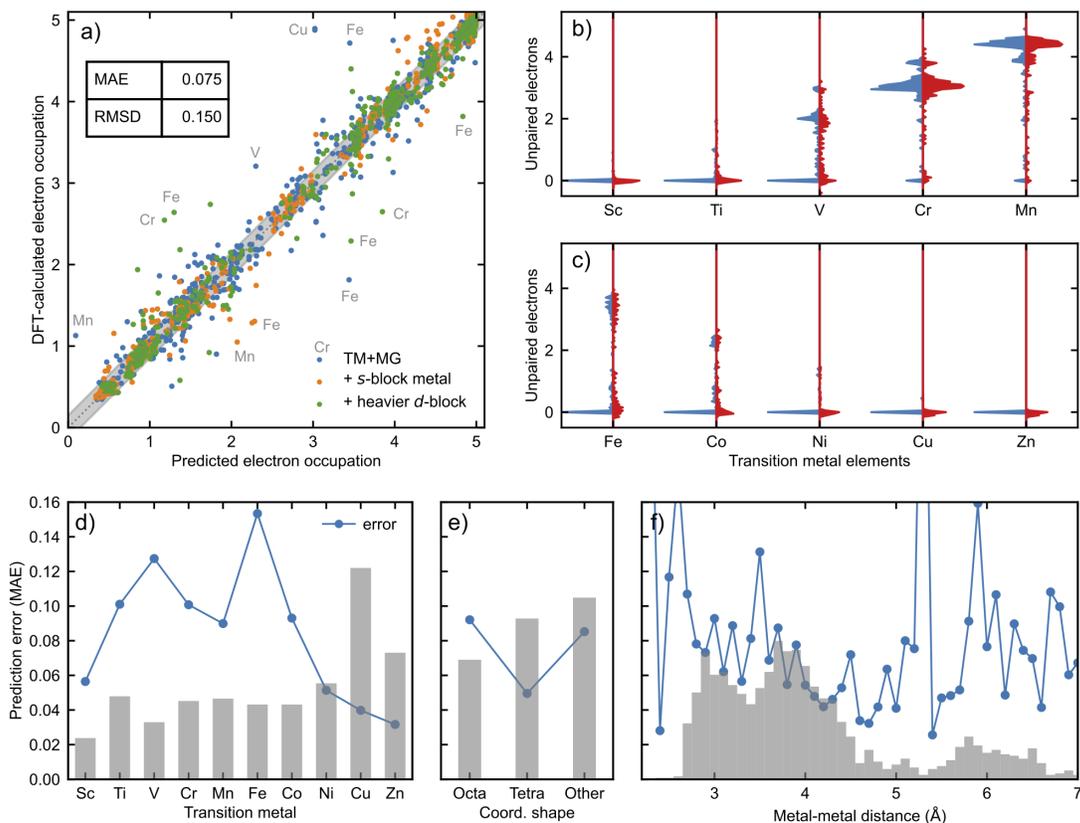


Figure 3: (a) Comparison between predicted and calculated values of the occupation of the TM $3d$ orbitals ($N_{\uparrow/\downarrow}^{3d}$) for both spin-up and spin-down channels in each compound of the training set (TM = transition metal, MG = main group). Spin-up and spin-down occupations are plotted together. The gray area refers to error of 0.15 electron. We highlighted several transition-metal type that deviate the most from the diagonal line. MAE and RMSD of all the test samples are shown in the table. (b) and (c) Histograms of calculated and predicted number of unpaired electron ($N_{\uparrow}^{3d} - N_{\downarrow}^{3d}$) for the complete training data (they are normalized in such a way that maximum values are the same for each elements). (d)–(f) Prediction error (MAE) for different categories of training samples (blue curve) and their relative distribution (gray bars). All errors refer to the y -axis in (d). Distribution in each plot sums up to 1. A decomposition of figure (e) and (f) for each transition metal element is given in SI.

additional elements). Overall, we find that predictions are most accurate when the formula contains TM, MG and s -block metal with an MAE of 0.066, while the other categories have slightly larger error and more mispredictions far away from the diagonal (MAE = 0.084 for TM + MG and MAE = 0.071 for others). It is possible that the s -block metals, the main effect of which is to donate electrons, play a minor and more predictive role in terms

of affecting the TM–anion interaction, and thus help to reduce prediction errors in these compounds.

In Figure 3 (b)–(d) we show errors related to each TM element. The error is relatively small for Sc, Cu and Zn, likely due to their fewer possible oxidation states. On the other hand, errors are relatively large for V and Fe. As it can be seen in the comparison between the calculated and predicted number of unpaired spins, these two elements have the most scattered spin polarization, leading to the more error-prone prediction results.

Finally, Figures 3 (e) and (f) show the error in terms of different coordination geometries and metal–metal bond distances. It is interesting to note that although both are well understood in terms of transition-metal chemistry, prediction for TM atoms in tetrahedral coordination is more accurate. For the metal–metal distance, we find a clear trend that with increasing the metal–metal distance up to 5 Å, the prediction errors decreased. This agrees well with the analysis made above which shows that the approach from local geometry is valid if long-range metal–metal interactions can be safely ignored. Above 5 Å, the prediction error increases again, which is likely due to complex crystal structures with a large unit cell.

Application

We apply our machine learning model to predict the local magnetic moments on the 3*d* transition metal elements given by $m_s = 2\mu_B\sqrt{S_p(S_p + 1)}$ from the predicted unpaired electron numbers $S_p = (N_{\uparrow}^{occ.} - N_{\downarrow}^{occ.})/2$.⁴⁹ Crystal structure parameters (for example, CIF files) are the only input that is necessary. Such an application can be useful, for example, to help assess the electronic configurations of the transition-metal compounds when access to DFT electronic structures is difficult (for experimental chemists) or when CPU time is of concern (screening). Long time (hours) needed for carrying out DFT calculations on structures of the data set can be reduced to within a few seconds to obtain magnetic moment prediction for hundreds of compounds on a normal laptop CPU, irrespective of the cell size. We use the “first-shell” method as the structure description and average the prediction

results among 10 separately trained models.⁴⁷

A total of 170 compounds were predicted and tested against DFT calculations. Each of the compounds contain from 30 to 100 atoms per unit cell. Therefore they are not included in the training data sets, and they were randomly chosen from the structure databases as long as they satisfy the requirement that $3d$ transition metal atoms are surrounded by main group anions. Prediction results in terms of unpaired electron numbers S_p are compared to the results of DFT calculations, and are shown in Figure 4. The entries were sorted according to the DFT calculated magnetizations.

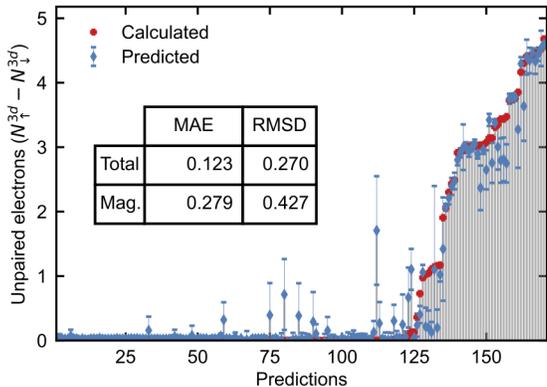


Figure 4: Predicted and calculated number of unpaired electrons for the application dataset. Only one site is predicted for a given compounds and results are sorted according to increasing polarization calculated using DFT method. The error bar shows the standard deviation in the predictions of the ensemble model. MAE and RMSD score are reported in the table for all prediction (Total) and prediction where number of unpaired electron is larger than 0.2 (Mag.)

It was found that the model shows good ability to distinguish magnetic compounds from the non-magnetic ones, although there are occasional overestimation of spin polarization. The most notable one is the compound Mn_4Si_7 , which is predicted to have an unpaired electron count around 2 but is in fact almost non-magnetic. In terms of values of predicted magnetization, a MAE value of 0.123 was obtained over all samples, which increased to 0.279 if only the magnetic ones were counted. From Figure 4, the main prediction errors in magnetization come from two regions where prediction underestimated the calculated values. We found that the first region was due to the mispredictions related to V and Cr,

while the second region of mispredictions concerns exclusively compounds containing Fe with a tetrahedral coordination. This agrees with our observation in the training that V and Fe show the largest error.

Table 1: Space Group (SG), d_{M-M} (Å), Metal Coordination (M coord.), ML Predicted (m_s^{pred}, μ_B) Spin-Only Magnetic Moments (\pm standard deviation) and Experimental Magnetic Moments (m_s^{exp}, μ_B) of ML Predicted Magnetic Compounds

Formula	SG	d_{M-M}	M coord.	$m_s^{pred.}$	$m_s^{exp.}$	ref.
Cr ₅ S ₈	$P\bar{1}$	2.96	Cr-octa	3.78 ± 0.07	4.00	50
Mn ₄ Si ₇	$P\bar{4}c2$	2.97	Mn-other	2.49 ± 0.92	0.01	51
Mn ₃ Ge ₅	$P\bar{4}$	3.05	Mn-octa	3.49 ± 0.44	1.00	52
Fe ₇ Se ₈	$P\bar{3}_1$	3.17	Fe-octa	3.02 ± 0.46	2.40	53
MnGa ₂ S ₄	$I\bar{4}$	5.45	Mn-tetra	5.32 ± 0.10	5.10	54
Ba ₃ FeS ₅	$Pnma$	6.27	Fe-tetra	3.62 ± 0.32	5.10	55
V ₄ S ₉ Br ₄	$P4/nmm$	2.98	V-other	1.68 ± 1.53	1.77	56
V ₂ P ₄ S ₁₃	$P\bar{1}$	3.71	V-octa	2.90 ± 0.05	2.80	57
CsMnInTe ₃	$C2/c$	4.31	Mn-tetra	5.34 ± 0.05	5.92	58
K ₁₀ Mn ₄ Sn ₄ S ₁₇	$R\bar{3}m$	3.91	Mn-tetra	5.21 ± 0.19	5.00	59
Ba ₃ FeS ₄ Br	$Pnma$	6.26	Fe-tetra	3.61 ± 0.31	3.85	60
Ba ₂ FeSbS ₅	$Pnma$	5.84	Fe-tetra	3.67 ± 0.35	5.09	61

Finally, we compared our model with a series of experimental measurements available, as shown in Table 1. The accuracy between prediction and experiments is worse than when comparing to DFT results, with a relatively large error observed for some of them such as Mn₃Ge₅, Mn₄Si₇, Fe₇Se₈, Ba₃FeS₅, or Ba₂FeSbS₅.

It should be pointed out that these compounds share some common features of disruption of long-range ordering that might influence their physical properties. For example, Mn₄Si₇ and Mn₃Ge₅ are examples of Nowotny chimney-ladder structures which consist of the interpenetration of [Si/Ge] and [Mn] subsystems with an incommensurate c parameter.^{52,62–65} Similarly, Fe₇Se₈ and its counterpart Fe₇S₈ are well known examples of compounds with a superstructure due to ordered intrinsic vacancies.^{66,67} The difference between ideal input structures and experimental ones must contribute to prediction deviations. Furthermore, some of these compounds may show a strong itinerant magnetic behavior, as exhibited by Mn₄Si₇, for instance, and therefore, local description of the magnetic moment may not be

valid. The last two compounds, namely $\text{Ba}_2\text{FeSbS}_5$ and Ba_3FeSe_5 correspond to the high-pressure and low pressure phases of Ba_3FeS_5 , respectively.^{55,68-71} For the latter, the experimental magnetic moment extracted from the magnetic susceptibility measurement seems to fit well with Fe^{4+} ions. However, a Mössbauer analysis rather showed Fe^{3+} ,⁵⁵ suggesting some complexity of the magnetic properties. It is also noteworthy that this compound can deviate from stoichiometry as part of the solid solution $\text{Ba}_3\text{Fe}_{1+x}\text{Se}_5$ ($0 \leq x \leq 1$).⁶⁸ Therefore, one important point that should not be overlooked is how the crystal and electronic structures of these compounds with a large unit cell in general are experimentally more complex than assumed in the training data sets, and this complexity contributes to the prediction errors when comparing to experimental results.

Other factors contributing to the deviations between prediction and experimental results include the choice of functionals, geometry relaxation and a more careful determination of the magnetic ground state, which are related to the DFT calculations which form the base of training data. It is known that GGA functionals systematically underestimate of the spin polarization due to the inaccuracy in the treatment of exchange interactions. However, hybrid functionals which could partly overcome this problem, are computationally too much expensive to perform in a high-throughput way. Geometry relaxation could also influence the value of the magnetic moment, but according to the few compounds we tested, its effect is minor in most cases. Finally, we note that DFT self-consistent calculations only determine a local minimum which may not be the true ground state. Exploration of different magnetic structures and carefully determination of the ground state could lead to more accurate DFT results and therefore better ML prediction, although this would come at the expense of CPU time.^{35,43}

Finally, it should be stressed again that the machine learning method developed here was trained on a dataset of solid-state compounds containing transition-metals surrounding by a specific type of neighbor atoms to predict their DFT-calculated properties. It is not then expected that it will perform well for oxides or nitrides for instance, due to the specific

choice of the training dataset. It will also not perform well for metals or covalently-bonded semiconductors due to the underlying theory of transition-metal–anion interactions.

CONCLUSION

Motivated by the use of local atomic structure to directly estimate the electronic structures of the transition-metal elements in inorganic chemistry, a neural network model using only local structure descriptions and trained from DFT-calculated electronic structures was developed to predict the electronic occupations of the 3*d* transition metal atoms in inorganic compounds including diverse intermetallic and semiconducting compounds. Trained on DFT calculations, this ML model is rather accurate in predicting the electron occupation of spin-up and spin-down channels for TM 3*d* orbitals. We also performed decomposition of prediction errors showing that they roughly agree with our chemical intuition. The application of this network model was illustrated to predict the magnetic moments of transition-metal atoms in compounds with large unit cells, that would be costly to calculate with the DFT method. Although the prediction results of the current model are not fully ideal when comparing to experimental observations, it constitutes a step toward building better ML models based on local structure features to predict electronic properties of inorganic solid-state compounds.

Data and Code Availability

The Python codes used in this work are organized as a package and can be found in Github repository <https://github.com/whzhangg/ElectronicConfiguration>. Crystal structures and calculated electronic occupations are provided along the code as well as examples. Complete data of DFT input and output can be obtained upon requests from the authors.

Acknowledgement

We acknowledge support from JST Mirai Program JPMJMI19A1 and helpful discussion with Dr. F. Pawula, Dr. I. Ohkubo and Dr. G. Lambard (NIMS, Tsukuba).

Supporting Information Available

Supplementary information can be found free of charge from the publisher website, including details on compound selection, DFT calculations, an overview of rotational invariant descriptor generation and neural network construction. It also contains a full list of compounds used for training and testing, with their properties such as metal–metal distance and coordination types.

References

- (1) Tokura, Y. Critical features of colossal magnetoresistive manganites. *Rep. Prog. Phys.* **2006**, *69*, 797–851.
- (2) Graf, T.; Felser, C.; Parkin, S. S. Simple rules for the understanding of Heusler compounds. *Prog. Solid State Chem.* **2011**, *39*, 1–50.
- (3) Hosono, H.; Tanabe, K.; Takayama-Muromachi, E.; Kageyama, H.; Yamanaka, S.; Kumakura, H.; Nohara, M.; Hiramatsu, H.; Fujitsu, S. Exploration of new superconductors and functional materials, and fabrication of superconducting tapes and wires of iron pnictides. *Sci. Technol. Adv. Mater.* **2015**, *16*, 033503.
- (4) Grin, Y. Inhomogeneity and anisotropy of chemical bonding and thermoelectric properties of materials. *J. Solid State Chem.* **2019**, *274*, 329–336.
- (5) Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.

- (6) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897–923.
- (7) Togo, A.; Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **2015**, *108*, 1–5.
- (8) Madsen, G. K.; Singh, D. J. BoltzTraP. A code for calculating band-structure dependent quantities. *Comput. Phys. Commun.* **2006**, *175*, 67–71.
- (9) Nelson, R.; Ertural, C.; George, J.; Deringer, V. L.; Hautier, G.; Dronskowski, R. LOBSTER: Local orbital projections, atomic charges, and chemical-bonding analysis from projector-augmented-wave-based density-functional theory. *J. Comput. Chem.* **2020**, *41*, 1931–1940.
- (10) Burdett, J. K. *Chemical bonding in solids*; Oxford University Press: New York, 1995.
- (11) Albright, T. A.; Burdett, J. K.; Whangbo, M.-H. *Orbital interactions in chemistry*, 2nd ed.; Wiley: Hoboken, New Jersey, 2013.
- (12) Housecroft, C. E.; Sharpe, A. G. *Inorganic chemistry*, 4th ed.; Pearson: Edinburgh Gate, Harlow, England, 2012.
- (13) Gaultois, M. W.; Oliynyk, A. O.; Mar, A.; Sparks, T. D.; Mulholland, G. J.; Meredig, B. Perspective: Web-based machine learning models for real-time screening of thermoelectric materials properties. *APL Mater.* **2016**, *4*, 053213.
- (14) Schmidt, J.; Marques, M. R. G.; Botti, S.; Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **2019**, *5*, 83.
- (15) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P. A Critical Review of Machine Learning of Energy Materials. *Adv. Energy Mater.* **2020**, *10*, 1903242.

- (16) Faber, F. A.; Hutchison, L.; Huang, B.; Gilmer, J.; Schoenholz, S. S.; Dahl, G. E.; Vinyals, O.; Kearnes, S.; Riley, P. F.; von Lilienfeld, O. A. Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error. *J. Chem. Theory Comput.* **2017**, *13*, 5255–5264.
- (17) Artrith, N.; Butler, K. T.; Coudert, F.-X.; Han, S.; Isayev, O.; Jain, A.; Walsh, A. Best practices in machine learning for chemistry. *Nat. Chem.* **2021**, *13*, 505–508.
- (18) Hoffmann, R. How Chemistry and Physics Meet in the Solid State. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 846–878.
- (19) Parsaeifard, B.; Sankar De, D.; Christensen, A. S.; Faber, F. A.; Kocer, E.; De, S.; Behler, J.; Anatole von Lilienfeld, O.; Goedecker, S. An assessment of the structural resolution of various fingerprints commonly used in machine learning. *Mach. Learn.: Sci. Technol.* **2021**, *2*, 015018.
- (20) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.
- (21) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B: Solid State* **2013**, *87*, 184115.
- (22) Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the electronic structure problem with machine learning. *npj Comput. Mater.* **2019**, *5*, 22.
- (23) Ben Mahmoud, C.; Anelli, A.; Csányi, G.; Ceriotti, M. Learning the electronic density of states in condensed matter. *Phys. Rev. B: Solid State* **2020**, *102*, 235130.
- (24) Chen, Z.; Andrejevic, N.; Smidt, T.; Ding, Z.; Xu, Q.; Chi, Y.-T.; Nguyen, Q. T.;

- Alatas, A.; Kong, J.; Li, M. Direct Prediction of Phonon Density of States With Euclidean Neural Networks. *Adv. Sci.* **2021**, *8*, 2004214.
- (25) Schütt, K. T.; Glawe, H.; Brockherde, F.; Sanna, A.; Müller, K. R.; Gross, E. K. U. How to represent crystal structures for machine learning: Towards fast prediction of electronic properties. *Phys. Rev. B: Solid State* **2014**, *89*, 205118.
- (26) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nat. Commun.* **2017**, *8*, 13890.
- (27) 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.* **2018**, *148*, 241722.
- (28) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.
- (29) Thomas, N.; Smidt, T.; Kearnes, S.; Yang, L.; Li, L.; Kohlhoff, K.; Riley, P. Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds. 2018; <https://arxiv.org/abs/1802.08219>, (accessed Nov 30, 2021).
- (30) Uhrin, M. Through the eyes of a descriptor: Constructing complete, invertible descriptions of atomic environments. *Phys. Rev. B: Solid State* **2021**, *104*, 144110.
- (31) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*, 011002.
- (32) Gražulis, S.; Daškevič, A.; Merkys, A.; Chateigner, D.; Lutterotti, L.; Quirós, M.; Serebryanaya, N. R.; Moeck, P.; Downs, R. T.; Le Bail, A. Crystallography Open

- Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Res.* **2012**, *40*, D420–D427.
- (33) Morales-Garcia, A.; Valero, R.; Illas, F. An Empirical, yet Practical Way To Predict the Band Gap in Solids by Using Density Functional Band Structure Calculations. *J. Phys. Chem. C* **2017**, *121*, 18862–18866.
- (34) Hegde, V. I.; Borg, C. K. H.; del Rosario, Z.; Kim, Y.; Hutchinson, M.; Antono, E.; Ling, J.; Saxe, P.; Saal, J. E.; Meredig, B. Reproducibility in high-throughput density functional theory: a comparison of AFLOW, Materials Project, and OQMD. 2020; <https://arxiv.org/abs/2007.01988>, (accessed Jan 21, 2022).
- (35) Noh, J.; Osman, O. I.; Aziz, S. G.; Winget, P.; Brédas, J.-L. A density functional theory investigation of the electronic structure and spin moments of magnetite. *Sci. Technol. Adv. Mater.* **2014**, *15*, 044202.
- (36) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; et al, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
- (37) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; et al, Advanced capabilities for materials modelling with QUANTUM ESPRESSO. *J. Phys.: Condens. Matter* **2017**, *29*, 465901.
- (38) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406.
- (39) van Setten, M.; Giantomassi, M.; Bousquet, E.; Verstraete, M.; Hamann, D.; Gonze, X.;

- Rignanese, G.-M. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Comput. Phys. Commun.* **2018**, *226*, 39–54.
- (40) Borlido, P.; Schmidt, J.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning. *npj Comput. Mater.* **2020**, *6*, 96.
- (41) Isaacs, E. B.; Wolverton, C. Performance of the strongly constrained and appropriately normed density functional for solid-state materials. *Phys. Rev. Materials* **2018**, *2*, 063801.
- (42) Ertural, C.; Steinberg, S.; Dronskowski, R. Development of a robust tool to extract Mulliken and Löwdin charges from plane waves and its application to solid-state materials. *RSC Adv.* **2019**, *9*, 29821–29830.
- (43) Horton, M. K.; Montoya, J. H.; Liu, M.; Persson, K. A. High-throughput prediction of the ground-state collinear magnetic order of inorganic materials using Density Functional Theory. *npj Comput. Mater.* **2019**, *5*, 64.
- (44) O’Keeffe, M. A proposed rigorous definition of coordination number. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1979**, *35*, 772–775.
- (45) Waroquiers, D.; George, J.; Horton, M.; Schenk, S.; Persson, K. A.; Rignanese, G.-M.; Gonze, X.; Hautier, G. ChemEnv: a fast and robust coordination environment identification tool. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2020**, *76*, 683–695.
- (46) Snyder, G.; Gelabert, M. C.; DiSalvo, F. Refined Structure and Properties of the Layered Mott Insulator BaCoS₂. *J. Solid State Chem.* **1994**, *113*, 355–361.
- (47) Bishop, C. M. *Pattern recognition and machine learning*; Information science and statistics; Springer: New York, 2006.

- (48) Goodfellow, I.; Bengio, Y.; Courville, A. *Deep learning*; Adaptive computation and machine learning; The MIT Press: Cambridge, Massachusetts, 2016.
- (49) Jiles, D. *Introduction to magnetism and magnetic materials*, third edition ed.; CRC Press, Taylor & Francis Group: Boca Raton FL, 2016.
- (50) Sleight, A. W.; Bither, T. A. Chromium chalcogenides prepared at high pressure and the crystal growth of chromium sesquisulfide. *Inorg. Chem.* **1969**, *8*, 566–569.
- (51) Gottlieb, U.; Sulpice, A.; Lambert-Andron, B.; Laborde, O. Magnetic properties of single crystalline Mn_4Si_7 . *J. Alloys Compd.* **2003**, *361*, 13–18.
- (52) Castillo, R.; Schnelle, W.; Bobnar, M.; Cardoso-Gil, R.; Schwarz, U.; Grin, Y. Structural, Magnetic and Thermoelectric Properties of hp- Mn_3Ge_5 . *Z. Anorg. Allg. Chem.* **2020**, *646*, 256–262.
- (53) Boumford, C.; Morrish, A. H. Magnetic properties of the iron selenide Fe_7Se_8 . *Phys. Status Solidi A* **1974**, *22*, 435–444.
- (54) Rimet, R.; Buder, R.; Schlenker, C.; Roques, R.; Zanchetta, J. New antiferromagnetic properties of manganese gallium sulfides $\text{MnGa}_2\text{S}_{4-\alpha}$ and $\text{MnGa}_2\text{S}_{4-\beta}$. *Solid State Commun.* **1981**, *37*, 693–697.
- (55) Lemley, J.; Jenks, J.; Hoggins, J.; Eliezer, Z.; Steinfink, H. $\text{Ba}_{15}\text{Fe}_7\text{S}_{25}$ and Ba_3FeS_5 : Crystal structures, Mössbauer, magnetic, and electrical behavior. *J. Solid State Chem.* **1976**, *16*, 117–128.
- (56) Mironov, Y. V.; Yarovoi, S. S.; Naumov, D. Y.; Kozlova, S. G.; Ikorsky, V. N.; Kremer, R. K.; Simon, A.; Fedorov, V. E. $\text{V}_4\text{S}_9\text{Br}_4$: A Novel High-Spin Vanadium Cluster Thiobromide with Square-Planar Metal Core. *J. Phys. Chem. B* **2005**, *109*, 23804–23807.

- (57) Evain, M.; Brec, R.; Ouvrard, G.; Rouxel, J. Synthesis and structure determination of a new layered compound: $V_2P_4S_{13}$. *J. Solid State Chem.* **1985**, *56*, 12–20.
- (58) Lin, H.; Shen, J.-N.; Chen, L.; Wu, L.-M. Quaternary Supertetrahedra-Layered Telluride $CsMnInTe_3$: Why Does This Type of Chalcogenide Tilt? *Inorg. Chem.* **2013**, *52*, 10726–10728.
- (59) Palchik, O.; Iyer, R. G.; Canlas, C. G.; Weliky, D. P.; Kanatzidis, M. G. $K_{10}M_4M'_4S_{17}$ (M = Mn, Fe, Co, Zn; M' = Sn, Ge) and $Cs_{10}Cd_4Sn_4S_{17}$: Compounds with a Discrete Supertetrahedral Cluster. *Z. Anorg. Allg. Chem.* **2004**, *630*, 2237–2247.
- (60) Li, X.; Li, C.; Kang, L.; Wu, Y.; Lin, Z.; Yao, J.; Wu, Y. Ba_3FeS_4Br : A 0D Iron-Based Chalcohalide with Unusual Magnetic Properties. *Eur. J. Inorg. Chem.* **2016**, *2016*, 1359–1363.
- (61) Maier, S.; Hebert, S.; Kabbour, H.; Pelloquin, D.; Perez, O.; Berthebaud, D.; Gascoin, F. Synthesis, electronic structure and physical properties of polycrystalline $Ba_2FePnSe_5$ (Pn=Sb, Bi). *Mater. Chem. Phys.* **2018**, *203*, 202–211.
- (62) Sato, N.; Ouchi, H.; Takagiwa, Y.; Kimura, K. Glass-like Lattice Thermal Conductivity and Thermoelectric Properties of Incommensurate Chimney-Ladder Compound $FeGe_\gamma$. *Chem. Mater.* **2016**, *28*, 529–533.
- (63) Le Tonquesse, S.; Hassam, C.; Michiue, Y.; Matsushita, Y.; Pasturel, M.; Mori, T.; Suzuki, T. S.; Berthebaud, D. Crystal structure and high temperature X-ray diffraction study of thermoelectric chimney-ladder $FeGe_\gamma$ ($\gamma \approx 1.52$). *J. Alloys Compd.* **2020**, *846*, 155696.
- (64) Le Tonquesse, S.; Joanny, L.; Guo, Q.; Elkaim, E.; Demange, V.; Berthebaud, D.; Mori, T.; Pasturel, M.; Prestipino, C. Influence of Stoichiometry and Aging at Operating Temperature on Thermoelectric Higher Manganese Silicides. *Chem. Mater.* **2020**, *32*, 10601–10609.

- (65) Guo, Q.; Zhang, W.; Liu, Z.; Fu, X.; Le Tonquesse, S.; Sato, N.; Son, H.-W.; Shimamura, K.; Berthebaud, D.; Mori, T. Thermoelectric Performance of Cr Doped and Cr–Fe Double-Doped Higher Manganese Silicides with Adjusted Carrier Concentration and Significant Electron–Phonon Interaction. *ACS Appl. Mater. Interfaces* **2021**, *13*, 8574–8583.
- (66) Li, G.; Zhang, B.; Baluyan, T.; Rao, J.; Wu, J.; Novakova, A. A.; Rudolf, P.; Blake, G. R.; de Groot, R. A.; Palstra, T. T. M. Metal–Insulator Transition Induced by Spin Reorientation in Fe₇Se₈ Grain Boundaries. *Inorg. Chem.* **2016**, *55*, 12912–12922.
- (67) Simon, J.; Guélou, G.; Srinivasan, B.; Berthebaud, D.; Mori, T.; Maignan, A. Exploring the thermoelectric behavior of spark plasma sintered Fe_{7-x}Co_xS₈ compounds. *J. Alloys Compd.* **2020**, *819*, 152999.
- (68) Cohen, S.; Rendon-Diazmiron, L.; Steinfink, H. Phases in the Ba₃Fe_{1+x}S₅ series: The structure of β-Ba₉Fe₄S₁₅ and its low-temperature α polymorph. *J. Solid State Chem.* **1978**, *25*, 179–187.
- (69) Berthebaud, D.; Preethi Meher, K.; Pelloquin, D.; Maignan, A. Synthesis, crystal structure and electronic properties of the new iron selenide Ba₉Fe₄Se₁₆. *J. Solid State Chem.* **2014**, *211*, 184–190.
- (70) Maier, S.; Gaultois, M. W.; Matsubara, N.; Surta, W.; Damay, F.; Hebert, S.; Hardy, V.; Berthebaud, D.; Gascoin, F. Sb-5 s lone pair dynamics and collinear magnetic ordering in Ba₂FeSbSe₅. *Phys. Rev. B: Solid State* **2021**, *103*, 054115.
- (71) Geng, L.; Cheng, W.-D.; Zhang, H.; Lin, C.-S.; Zhang, W.-L.; Li, Y.-Y.; He, Z.-Z. Syntheses, Crystal and Electronic Structures, and Characterizations of Quaternary Antiferromagnetic Sulfides: Ba₂MFeS₅ (M = Sb, Bi). *Inorg. Chem.* **2011**, *50*, 2378–2384.

TOC Graphic

