High-throughput search for magnetic and topological order in transition metal oxides

Nathan C. Frey¹,², Matthew K. Horton²,³, Jason M. Munro², Sinéad M. Griffin⁴,⁵, Kristin A. Persson²,³, Vivek B. Shenoy¹*

The discovery of intrinsic magnetic topological order in MnBi₂Te₄ has invigorated the search for materials with coexisting magnetic and topological phases. These multiorder quantum materials are expected to exhibit new topological phases that can be tuned with magnetic fields, but the search for such materials is stymied by difficulties in predicting magnetic structure and stability. Here, we compute more than 27,000 unique magnetic orderings for more than 3000 transition metal oxides in the Materials Project database to determine their magnetic ground states and estimate their effective exchange parameters and critical temperatures. We perform a high-throughput band topology analysis of centrosymmetric magnetic materials, calculate topological invariants, and identify 18 new candidate ferromagnetic topological semimetals, axion insulators, and antiferromagnetic topological insulators. To accelerate future efforts, machine learning classifiers are trained to predict both magnetic ground states and magnetic topological order without requiring first-principles calculations.

INTRODUCTION

Materials with coexisting quantum phases offer exciting opportunities for solid-state device applications and exploring new physics emerging from the interplay between effects including topology and magnetism (1). Intrinsic magnetic topological materials enable both explorations of fundamental condensed matter physics and next-generation technologies that rely on topological quantum states. Notable progress has been made on classifying and discovering topological materials (2–7). Separate efforts have recently enabled high-throughput classification of magnetic behavior (8). However, both the theoretical design and experimental realization of magnetic topological materials are confounded by the inherent difficulties in predicting and controlling magnetic order, often arising from strong electron correlations (9, 10). At the intersection of these two quantum orders, the focus has mainly been on taking prototypical topological materials like (Bi, Sb)₂Te₃ and introducing magnetic dopants (11, 12). This doping approach has a number of drawbacks, namely, that dopants are hard to control (13) and the critical temperature for observing exotic physics is low (below 2 K) (14). Important challenges remain in assessing the stability (synthetic accessibility) of proposed materials (15) and coupling topological property prediction to magnetism.

Recently, there has been a surge of interest in the field because of the experimental realization of intrinsic magnetic topological phases in the van der Waals material MnBi₂Te₄ (16, 17). The MnBi₂Te₄ family offers the first opportunity to possibly access the quantum anomalous Hall phase, topological axion states, and Majorana fermions in a single materials platform by controlling the interplay between the magnetic and topological orders (14). This demonstration of a true magnetic topological quantum material (MTQM) opens new avenues of research into the modeling, discovery, and characterization of magnetic topological materials. Many opportunities remain, in particular, the realization of new magnetic topological phases (10) and robust order in ambient conditions.

In this work, we develop and apply workflows to automate the calculation of magnetic exchange parameters, critical temperatures, and topological invariants to enable high-throughput discovery of MTQMs. Building on previous work to determine magnetic ground states with density functional theory (DFT) calculations (8), we apply the workflow to a subset of more than 3000 transition metal oxides (TMOs) in the Materials Project database (18). We focus on TMOs because of the large range of tunability that has already been demonstrated, both through ab initio calculations and molecular beam epitaxial growth of single phase and heterostructured oxide compounds, and the tantalizing possibility of incorporating them with oxide electronics. Several previous works have identified potential topological oxide candidates on a case-by-case basis (19, 20), although few have been successfully synthesized or measured to be topological (15). The workflow is used to identify candidate ferromagnetic topological semimetals (FMTSMs), axion insulators, and antiferromagnetic topological insulators (AFTIs), as well as layered magnetic and topological materials. Moreover, the computed magnetic orderings and a recently published dataset of predicted magnetic topological materials (10) are used to train machine learning (ML) classifiers to predict magnetic ground states and magnetic topological order, which may be used to accelerate the exploration of the remaining 31,000+ magnetic materials in the Materials Project. Future work will extend this modular workflow to explore diverse phenomena including other ferroic orders, exotic topological phases, and new materials systems.

RESULTS

Calculating magnetic and topological order

The workflow presented here is graphically summarized in Fig. 1. For any candidate magnetic material, the method previously developed by some of the coauthors (8) is used to generate likely collinear magnetic configurations based on symmetry considerations. The
method proceeds by generating up to eight candidate magnetic orderings and sorting them by symmetry [with ferromagnetic (FM) being the most symmetric]. If multiple orderings are found with equal symmetry at the eighth index, then the cutoff is increased, and up to 16 orderings are considered. The exact number of orderings considered depends on the size of the unit cell, the number of unique magnetic sublattices, and the number of different species of magnetic ions. Exhaustive DFT calculations are performed to compute the energies of each magnetic ordering and determine the ground state. Alternatively, the ML classifier discussed below can be used to predict the ground state ordering based solely on structural and elemental data to accelerate the ground state classification. The low energy orderings are then mapped to the Heisenberg model for classical spins, \( S \)

\[
H = -\sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j
\]

(1)

Solving the resulting system of equations yields the exchange parameters, \( J_{ij} \). The computed exchange parameters and magnetic moments provide all the necessary inputs to obtain the critical temperature through Monte Carlo simulations. The crystal is represented as a structure graph where nodes represent atoms and edges represent exchange interactions. The entire analysis has been implemented in the Python Materials Genomics (pymatgen) (21) code and an automated magnetism workflow is available in atomate (22). Monte Carlo calculations are enabled in the workflow by interfacing with the VAMPIRE atomistic simulations package (23). It should be noted that this method is only applicable for systems that are well described by the classical Heisenberg model, that is, systems with localized magnetic moments and reasonably high Curie or Néel temperatures (\( T_{CN} > 30 \) K), such that quantum effects can be neglected.

The second branch of the workflow diagnoses band topology. Topological invariants are determined using the vasp2trace (3) and ivsp (24) codes to compute irreducible representations of electronic states, as well as the hybrid Wannier function method in Z2Pack (25). Automated workflows to calculate topological invariants are implemented in the Python Topological Materials (pytopomat) code (26). By coordinating workflows, we are able to discover materials with coexisting quantum orders, like magnetic topological insulators, in a high-throughput context. The schematic in Fig. 1 shows one such example: a magnetic system exhibiting the quantum anomalous Hall effect.

**TMO database**

We restrict our search to the family of TMOs, which has the advantages of encompassing thousands of candidate magnetic materials and having standardized Hubbard \( U \) values based on experimental enthalpies of formation (27). A subset of 3153 TMOs were considered, encompassing more than 27,000 computed magnetic orderings, with any combination of Ti, V, Cr, Mn, Fe, Co, Ni, Cu, O, and any other non-\( f \) block elements. Because we have considered only oxides with magnetic ions and Hubbard \( U \) interactions that favor localized magnetic moments, most of the materials in the TMO dataset will exhibit magnetic ordering at 0 K. However, it is possible that some systems are nonmagnetic or weakly magnetic. Only stable and metastable phases within 200 meV of the convex hull are included in our database. For each TMO, up to 16 likely magnetic orderings were generated, yielding a total of 923 FM and 2230 antiferromagnetic (AFM) ground states. Seven structures relaxed into nonmagnetic ground states and were discarded. For simplicity, ferrimagnetic ground states were called AFM if they have an antiparallel spin configuration with a net magnetic moment less than 0.1 \( \mu_B \) per cell, and FM if the net magnetic moment in the system is greater than 0.1 \( \mu_B \) per cell. Our results on predicting AFM versus FM ground states are not sensitive to the choice of cutoff, as increasing the cutoff by an order of magnitude (to 1.0 \( \mu_B \) per cell) yields only 25 additional materials labeled AFM rather than FM.

A statistical summary of the dataset is presented in Fig. 2. Figure 2A shows a histogram of the crystal systems contained in the data. All seven crystal systems are represented, with monoclinic being the most prevalent and hexagonal the least. Similarly, there are a variety of space groups, compositions, and symmetries present in the materials considered. There are compounds with one, two, three, or more magnetic sublattices. As expected for TMOs, most compounds have an average coordination number of four or six. Considering the computed ground states, there is a large range of maximum magnetic moments per atom, with clustering observed around the integer values of 1, 2, 3, and 4 \( \mu_B \) per atom (Fig. 2B). This plot shows the maximum moment present in a unit cell rather than the net magnetic moment per cell, which is simply zero for AFM materials. The histogram of average nearest neighbor distance between two TM atoms in a compound is shown in Fig. 2C. There is again a large range of values, from 2 to 6 Å, with a peak around 3 Å. We also show the relative occurrence of 3d block transition metal species across FM and AFM compounds in Fig. 2D. Mn is the most common transition metal, occurring in 1016 compounds in the database, while Cu is the least prevalent, occurring in fewer than 100 compounds.

A wealth of information is available in the computed higher-energy orderings as well. We define the energy gap, \( \Delta E = E_0 - E_1 \),
where $E_0$ is the ground state energy and $E_1$ is the energy of the first excited state. When the low-energy orderings are successfully found, $\Delta E$ quantifies the robustness of the ground state ordering. The plot of $\Delta E$ in Fig. 3A shows the heavy-tailed distribution of the energy gaps. More than 600 compounds exhibit $\Delta E < 0.5$ meV per unit cell and may have correspondingly small $J$ and $T_{\text{CIN}}$ values. An effective $J$ parameter can be estimated from the energy gap as $|J_{\text{eff}}| = \Delta E / (NS^2)$, where $N$ is the number of magnetic atoms in the unit cell and $S$ is the magnitude of the average magnetic moment. From this crude estimate, the transition temperature is given in the mean field approximation as $T_{\text{CIN}}^{\text{MFT}} = 2J_{\text{eff}}/(3k_B)$, where $k_B$ is the Boltzmann constant. The plot in Fig. 3B shows the $J_{\text{eff}}$ values, more or less clustered by the integer values of the maximum magnetic moment (indicated by the light red ovals) in each material. One representative high $J_{\text{eff}}$ material is La$_2$NiO$_4$ (41.7 meV), which has an AFM ground state and an estimated $T_C$ of 323 K [measured value of 335 K (28)]. The mean field critical temperature estimates may be useful as an additional screening criterion for identifying materials that warrant further investigation, but they are not accurate enough to provide quantitative agreement with experiment.

We briefly highlight some promising material candidates that may be reduced into a two-dimensional (2D) form (29) with possible access to intriguing low-dimensional magnetic properties (30, 31) and materials with strong spin–orbit coupling (SOC). We apply the method from (32) to identify potentially layered TMOs, which yields 105 candidates. There are also 66 TMOs that contain either Bi or Hg and are therefore expected to exhibit strong SOC. At the intersection, we find three Bi-containing layered magnetic materials, Ba$_2$Mn$_2$Bi$_2$O, CoBiO$_3$, and CrBiO$_3$. It should be noted that in monolayers, the magnetic behavior and critical temperature is strongly dependent on magnetic anisotropy (33, 34). Therefore, the estimated exchange parameters and critical temperatures for the bulk phases will not correspond to the 2D limit.

The primary computational burden in generating this dataset is calculating the relaxed geometries and energies of all likely magnetic orderings, as we have no a priori way of determining the magnetic ground states. Further, it is useful to compute the spectrum of low-energy magnetic orderings to estimate the strength of exchange couplings, thereby determining the nature of magnetic interactions and critical temperatures. For simple compounds with small unit cells and a single type of magnetic ion, it is relatively easy to determine the ground state and only a few (<4) orderings need to be computed. However, there is a long tail to the number of orderings required for complex structures that may have many highly symmetric AFM orderings (8). For the TMO dataset, nine orderings per compound are required, on average, to find the ground state. Because of the computational cost, we are limited to collinear magnetic orderings in this combinatorial approach. Considering noncollinear ordering with SOC often increases the computation time by one to two orders of magnitude, because of the inclusion of the full spin-density matrix and the reduced symmetry (8). For this reason, a full noncollinear screening of the TMO database is not currently feasible (although all band topology calculations include SOC). Collinearity is a reasonable assumption, as many materials exhibit collinear magnetic ordering, although this assumption breaks down for triangular
and Kagomé lattices with frustrated antiferromagnetism and in systems with lanthanides and actinides (35). However, this work is an important and necessary first step toward determining possible noncollinear orderings. It is highly desirable to augment these laborious DFT calculations with computationally inexpensive, physics-informed models that can predict magnetic behavior.

**Magnetic ordering ML classifier**

The size of the TMO dataset and the number of easily available, physically relevant descriptors suggest that a physics-informed ML classifier may be able to predict magnetic ground states. Our goal is to use features based purely on structural and compositional information, without any DFT calculations, to predict magnetic orderings and prioritize calculations. With the matminer (36) package, we have access to thousands of descriptors that are potentially correlated with magnetic ordering. Drawing on physical and chemical intuition, this list was reduced to ~100 descriptors that are likely indicators of magnetic ordering, e.g., elemental d orbital filling, electronegativity, and tabulated atomic magnetic moments. We note that to explore the possible magnetic and topological ordering of a material, a composition and crystal structure is required. Crystal structures are already available in the Materials Project database and are used for generating features. The DFT geometry relaxations that yield reasonable crystal structures are not sufficient to determine magnetic and topological order. Our goal is to use these already available crystal structures and compositions to generate descriptors and predict magnetic properties. We use relaxed structures already available in the Materials Project database to generate additional features more specific to magnetic compounds, including the average nearest neighbor distance between TM atoms, TM-O-TM bond angle information, TM atom coordination number, and the number of magnetic sublattices. We have implemented these features in the “magnetism” module of pymatgen.

As expected, no features have Pearson correlation coefficients larger than 0.3 with respect to ground state ordering. There are no features with strong enough linear correlation to reliably predict magnetic behavior. To further reduce the feature space, we train a minimal model and discard features with extremely low impurity importance and then perform hierarchical clustering (37) of the features based on the Spearman rank correlation, removing a feature from each cluster. Hierarchical clustering with Spearman rank correlation is a standard method of feature selection for eliminating redundant features based on the dependence between the rankings of features rather than their linear correlation. This kind of feature reduction is useful to simplify the inputs to the model and remove redundancies in the feature set.

Next, using the reduced feature set of 14 features, we construct an ensemble of ML classifiers to predict the magnetic behavior. For simplicity and interpretability, a random forest classifier was used, although other techniques like Adaptive Boosting and Extra Trees perform similarly. The random forest achieves superior performance compared to other models considered through AutoML in Automatminer (38). Automatminer compiles model performance between many standard classifiers, including random forest, logistic regression, and support vector classification, to determine the optimal classifier. The random forest is an ensemble of decision trees made up of “leaves” like the one shown schematically in Fig. 4A. For each feature, the tree splits the dataset to enable classification. In the illustration in Fig. 4A, the simplified split illustrates that samples with more than one magnetic sublattice are more likely to be AFM than FM. To capture the complexity of the data, a full decision tree is more fleshed out, like the one shown in Fig. 4B. The random forest is an ensemble of many such trees, where the predictions of uncorrelated trees are averaged over to reduce overfitting. Ten percent of the data were held as a test set, hyperparameters were tuned through a grid search, and fivefold cross-validation was used for validation, following the conventions in Matbench and Automatminer (38).

Because of the class imbalance between FM and AFM ground states (30% of compounds are FM), the FM compounds are synthetically oversampled using SMOTE (39). SMOTE generates synthetic data from the minority class (ferromagnets) by randomly choosing a sample and one of its nearest neighbors (in feature space), taking the feature vector between the two and multiplying it by a random number between 0 and 1. In this way, synthetic data that resemble data in the minority class are generated without simply duplicating samples. Rather than using accuracy (number of correct predictions divided by the total number of predictions), which can be artificially inflated in imbalanced learning problems by always guessing the majority class, we use three more robust classification metrics to evaluate our models: F1 score, receiver operating characteristic (ROC) area under the curve (AUC), and precision–recall (PR) AUC. The classifier exhibits good and consistent performance in fivefold cross-validation, as seen in the mean and median F1 scores of 0.85 for both FM and AFM classes (table S1). The trained model achieved an F1 score of 0.85 (0.59) for AFM (FM) compounds on the test set, suggesting that the synthetic oversampling results in difficulties generalizing to new FM compounds, while AFM systems are well characterized. However, the classifier performs much better than either random guessing or naively always choosing the majority class, as can be seen in the ROC curve and PR curve for the test set in fig. S1. The F1 scores, ROC, and PR curves all show that the classifier outperforms random and majority class guessing. In addition, in fig. S1C, we show the confusion matrix for the test set predictions. For the test set of 315 samples (236 AFM and 79 FM), 198 AFM samples are correctly classified, and 50 FM samples are correctly classified. We also include in the Supplementary Materials a discussion of the classifier’s performance on a dataset of 405 experimental crystal structures from the Inorganic Crystal Structure Database (ICSD) (40) for which we have computed magnetic orderings.

The success of the classifier allows us to reexamine the input features and use the model feature importances to identify nontrivial predictors of magnetic behavior. Although the Gini impurity–based feature importance somewhat misleadingly shows equal contributions from many features, here, we use the permutation importance, which avoids bias toward numerical and high cardinality features. The most important features for classification are shown in fig. S2.
By far, the most important descriptor is the number of magnetic sublattices. Other features relate to space group symmetry, d electron counts, coordination number, and distances between TM atoms—features we expect to describe magnetism. Another important descriptor is the sine Coulomb matrix, which is a vectorized representation of the crystal structure that has been introduced and used in previous studies to predict formation energies of crystals (41). Last, the structural complexity (discussed in the Supplementary Materials) is observed to be 47% higher, on average, for AFM compounds than FM. The AFM systems exhibit an average structural complexity that is 17 bits per unit cell higher than the FM systems. This simple metric might indicate that more structurally complex materials are more likely to favor the more complex AFM orderings rather than simple FM configurations. This could be related to the most important descriptor, which is also a metric of magnetic lattice complexity. All features are defined in table S2. Unexpectedly, these complexity metrics along with simple TM-TM atom distances and the sine Coulomb matrix are much better predictors of magnetic ordering than bond angle information, as might be expected from the Goodenough-Kanamori rules. It is possible that more sophisticated features may do a better job at capturing the superexchange mechanisms that govern the magnetic behavior. We note that high model performance, measured by ROC AUC, was achieved for magnetic material classification in (42) using similar features, including d-orbital chemical descriptors.

It is clear that models like the one presented here are not guaranteed to generalize beyond the material types that comprise the training data. However, we expect that the physical insights related to feature engineering, as well as the tested methods, will be of use in future studies. The developed models are particularly useful in the context of high-throughput virtual screening, where tens of thousands of materials are potentially of interest, and it is highly desirable to quickly predict target properties to prioritize calculations. Even for smaller datasets of 100 to 1000 materials, models for property prediction can be used in active learning (43) to efficiently identify materials with desired properties. Further difficulties will be encountered when constructing ML models for critical temperature prediction, which is inherently a problem of outlier detection. Fortunately, this work provides both a set of promising materials to consider for further study and the framework to automate evaluation of exchange parameters and critical temperatures.

**Magnetic topological material identification**

Last, we discuss the search for nontrivial band topology in the magnetic TMOs. The zoo of available topological order is ever expanding. Fortunately, this work provides both a set of promising materials to consider for further study and the framework to automate evaluation of exchange parameters and critical temperatures.

Potential AFTIs were identified by screening the TMO database for AFM ground states with S symmetry, yielding 298 candidate materials. Of these, 46 are predicted to be layered antiferromagnets with centrosymmetric tetragonal structures, where ideal Weyl semimetal (WSM) features (Fig. 5E) may appear and where the magnetization direction can tune the band topology (45). These filters greatly simplify the screening, but recent work suggests that more than 30% of nonmagnetic (3, 46) and magnetic (10) materials exhibit nontrivial topology, so there are almost certainly many more MTQMs to uncover in the TMO dataset than we have considered here.

We classify potential $\mathbb{Z}_2$ phases in AFM systems by the set of indices

$$Z_2 = (v_1, v_2, v_3, v_4) \quad (2a)$$

$$v = \Delta(k_i = 0) + \Delta(k_i = 1/2) \mod 2 \quad (2b)$$

$$v_1 = \Delta(k_i = 1/2) \quad (2c)$$

where $\Delta(k_i)$ is the 2D topological invariant on the time-reversal invariant plane $k_i$ in the BZ and $k_i$ is in reduced coordinates. If $v = 1$, the system is a strong topological insulator, while a system with $v = 0$ and $v_1 = 1$ for any $v_i$ is a weak topological insulator (25, 47). FMTSMs are classified by the strong topological index $Z_4$ in terms of parity eigenvalues (9, 10), defined as

$$Z_4 = \sum_{n=1}^{8} \sum_{n=1}^{n_{occ}} 1 + \xi_\alpha(A_n) \mod 4 \quad (3)$$

where $\xi_\alpha$ are the eight TRIM points, $n$ is the band index, $n_{occ}$ is the number of occupied bands, and $\xi_\alpha(A_n)$ is the parity eigenvalue ($\pm 1$) of the $n$th band at $A_n$. $Z_4 = 1$ indicates a WSM phase with an odd number of Weyl points in half of the BZ, while $Z_4 = 2$ indicates an axion insulator phase with a quantized topological magnetoelectric response (48), or a WSM phase with an even number of Weyl points. $Z_4 = 0$ corresponds to a topologically trivial phase.

The TMO database was screened for materials with FM ground states, a tetragonal crystal structure, and inversion symmetry, resulting in 27 candidates. By computing the $Z_4$ indices for these materials, we identify eight materials with $Z_4 = 2$, indicating either a WSM phase with an even number of Weyl points in half of the BZ or an axion insulator phase. Recent work has shown that $Z_4 = 2$ can also indicate a 3D Chern insulator phase (49). Five materials have an odd number of Weyl points in half of the BZ, with $Z_4 = 1$. The candidate FMTSMs and axion insulators and their respective $Z_4$ indices are listed in Table 1. We also give the unique identifiers for the Materials Project database entries and the calculated energy above the convex hull. Here, we highlight the candidate FMTSM CuCr$_2$O$_4$ (Fig. 5C). CuCr$_2$O$_4$ has an FM ground state and $Z_4 = 1$. CuCr$_2$O$_4$ is a hausmannite-like spinel structure with the tetragonal $I4_1/amd$ space group. Cr atoms bond with O atoms to form CrO$_6$ octahedra that share corners with CuO$_2$ tetrahedra. Cr$^{3+}$ atoms occupy Wyckoff position $b$; Cu$^{2+}$ atoms occupy Wyckoff $4a$, and O$^{2-}$ atoms occupy Wyckoff 16h. We also draw special attention to the spinel CdNi$_2$O$_4$ (Fig. 5F), which is predicted to be an FM axion insulator with $Z_4 = 2$ and a bandgap $E_{gap} = 0.125$ eV. This material has not yet been successfully synthesized and represents one of many promising opportunities to grow new magnetic oxides and investigate their topology.

Potential AFTIs were identified by screening the TMO database for AFM ground states with $S$ symmetry, yielding 298 candidate materials. Of these, 46 are predicted to be layered antiferromagnets.
by at least one of the methods available in pymatgen. These layered systems are of special interest because of their unique and tunable topological and magnetic properties \((30, 16)\). Eight additional antiferromagnets with \(S\) symmetry exhibit small bandgaps (<0.5 eV) and are therefore likely candidates to exhibit band inversion. For each of these 54 materials, the \(\mathbb{Z}_2\) invariant is calculated using the hybrid Wannier function method in Z2Pack. Four layered AFTIs were identified: FeMoClO\(_4\), MnMoO\(_4\), Ca\(_2\)MnO\(_3\), and SrV\(_3\)O\(_7\). One small bandgap AFTI was also found: monoclinic Li\(_2\)TiVO\(_4\) in a \(P2/m\) phase. These systems and their \(\mathbb{Z}_2\) indices are listed in Table 2. We highlight the tetragonal \(I4/mmm\) phase of Ca\(_2\)MnO\(_3\) (Fig. 6A), which has a nontrivial \(\mathbb{Z}_2 = (1; 000)\). It is a caswellsilverite-like structure.
in which Ca\(^{2+}\) ions are bonded with O atoms to form CaO\(_6\) octahedra and Mn\(^{2+}\) ions bond to form MnO\(_6\) octahedra. In the primitive cell, Ca atoms occupy Wyckoff position 4e, Mn occupies Wyckoff 2a, and the O atoms occupy Wyckoff 2b and 4e. Because the topology of the AFTI phase is sensitive to the nature of the bandgap and the strength of electron correlations, we plot a phase diagram (Fig. 6B) for Ca\(_2\)MnO\(_3\) indicating the regions where the system is a strong AFTI or a trivial insulator. We find that the material is a strong AFTI under a wide range of Hubbard \(U\) values, although it is predicted to be topologically trivial at \(U = 4\) eV and for \(U > 6\) eV. Future work will identify the origin of this correlation-dependent change in topological order.

None of the identified candidate MTQMs was considered in previous efforts to screen the Materials Project for topological materials (3), because the correct magnetic orderings were not available (8). We have also highlighted theoretical materials, unique to the Materials Project database, which have not yet been experimentally synthesized and do not have experimental structures reported in the ICSD. Theoretical materials are labeled with an asterisk in Tables 1 and 2. Three materials (CuFe\(_2\)O\(_4\), VMg\(_2\)O\(_4\), and Li\(_2\)TiVO\(_4\)) relaxed into new phases not previously included in the Materials Project database after determining the magnetic ground states. Notably, all MTQM candidates are within 100 meV per atom of the convex hull, indicating that all candidate materials are thermodynamically stable or metastable and may be synthesizable (50). Candidates with experimentally determined structures and magnetic orderings are listed in table S3.

We have extended the ML approach discussed above to classify magnetic topological materials from a recently published dataset (10) of 403 magnetic structures containing 130 magnetic topological materials. Calculations were performed with \(U\) values of 0, 2, 4, and 6 eV for each material, and the 130 predicted magnetic topological materials are magnetic enforced semimetals or topological insulators for at least one tested value of the Hubbard \(U\) parameter. Using only structural and chemical information, we trained an ML classifier to identify magnetic materials that are topologically nontrivial for at least one \(U\) value versus materials that are trivial for all \(U\) values. The random forest model achieves a 0.74 \(F_1\) score on topological material classification in fivefold cross-validation (table S4), using 13 primarily symmetry- and orbital-based descriptors (listed in table S5), requiring no calculations. ROC and PR curves for the classifier are shown in fig. S3. The \(F_1\) scores, ROC, and PR curves indicate that the classifier does much better than random guessing and majority class guessing. The permutation feature importance graph is shown in fig. S4. Further details are presented in the Supplementary Materials.

Because of the modularity and interoperability of the workflows developed and applied here, it is straightforward to extend the search to other types of quantum orders. Here, we have provided a high-throughput, relatively coarse-grained method to identify promising MTQMs. The topological structure can be sensitive to the Hubbard \(U\) parameter value, noncollinear magnetic order and the resulting magnetic space group determination, and how the strength of SOC compares to the bandgap. Future work will involve detailed studies of candidate materials with the recently introduced Magnetic Topological Quantum Chemistry (MTQC) (10) formalism, better exchange-correlation functionals to more accurately compute bandgaps, applying the workflow to noncentrosymmetric materials, and careful determination of \(U\) values with the linear response approach.

**DISCUSSION**

We have developed and applied a high-throughput computational workflow to determine magnetic exchange couplings, critical temperatures, and topological invariants of electronic band structures in magnetic materials. By studying more than 3000 TMOs spanning all crystal systems, nearly all space groups, and a wide range of compositions, we have produced a dataset of materials rich in magnetic and topological physics. This enabled the training of an ML classifier to predict magnetic ground states and give insight into structural and chemical factors that contribute to magnetic ordering. We extended this ML approach to classify topological order in magnetic materials from a recently published dataset using only symmetry- and orbital-based descriptors. We identified 5 promising candidate AFTIs (e.g., tetragonal Ca\(_2\)MnO\(_3\)), including four layered materials, as well as 13 candidate FMTSMs (spinel CuCr\(_2\)O\(_4\)) and axion insulators (spinel CdNi\(_2\)O\(_4\)).

**METHODS**

DFT calculations were performed with the Vienna Ab Initio Simulation Package (VASP) (35, 51) and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional (52). Standard Materials Project input settings and Hubbard \(U\) values were used, as described in (8). Specifically, values were set for elements Co (3.32 eV), Cr (3.7 eV), Fe (5.3 eV), Mn (3.9 eV), Ni (6.2 eV), and V (3.25 eV) with the rotationally

---

**Table 2. Candidate antiferromagnetic topological insulators.**

<table>
<thead>
<tr>
<th>Material</th>
<th>Space group</th>
<th>Materials Project ID</th>
<th>Energy above hull (meV/atom)</th>
<th>(Z_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FeMoClO(_4)</td>
<td>P4/nmm</td>
<td>mp-23123</td>
<td>6 (1;100)</td>
<td></td>
</tr>
<tr>
<td>MnMoO(_4)</td>
<td>P2/c</td>
<td>mp-19455</td>
<td>5 (1;001)</td>
<td></td>
</tr>
<tr>
<td>Ca(_2)MnO(_3)*</td>
<td>I4/mmm</td>
<td>mp-1227324</td>
<td>27 (1;000)</td>
<td></td>
</tr>
<tr>
<td>SrV(_2)O(_7)</td>
<td>Pmnn</td>
<td>mp-510725</td>
<td>3 (1;010)</td>
<td></td>
</tr>
<tr>
<td>Li(_2)TiVO(_4)*</td>
<td>P2/m</td>
<td>mp-759813</td>
<td>19 (1;001)</td>
<td></td>
</tr>
</tbody>
</table>

*Theoretical materials that have not yet been experimentally synthesized.

**Fig. 6. The candidate AFTI, Ca\(_2\)MnO\(_3\).** (A) Crystal structure of Ca\(_2\)MnO\(_3\) in the tetragonal \(I4/mmm\) phase. (B) Phase diagram of bandgap versus the Hubbard \(U\) value for Mn showing the dependence of the band topology on the strength of Hubbard interactions.
invariant Hubbard correction. These values were determined by fitting to known binary formation enthalpies in TMOs (27). Maintaining consistent $U$ values with the Materials Project allows high-throughput calculations and integration within the generalized gradient approximation (GGA)/GGA + $U$ mixing scheme that enables the construction of phase diagrams. These $U$ values and the magnetic ordering workflow were shown to correctly predict ground state magnetic ordering in more than 60% of cases and non-FM ground states in 95% of 64 benchmark materials with experimentally determined nontrivial magnetic order (8). However, it is known that topological phase diagrams for magnetic materials can be strongly dependent on the strength of Hubbard interactions (10). Topological index calculations were implemented with pytopomat (26). To determine topological indices, pytopomat automates the necessary electronic structure calculations and then uses vasp2trace and irvsp to calculate symmetry eigenvalues from the wave function data. Symmetry eigenvalues follow from calculating the traces of symmetry operations in the Bloch basis

$$\text{Tr} \left[ \hat{S}_z \right] = \sum_{i} n_i$$

$$= \sum_{i} \langle \psi_k | \hat{S}_z | \psi_i \rangle$$

(4)

(5)

where $\hat{S}_z$ is a symmetry operation in the little group of the wave vector $k$ and $\psi_i$ is a normalized Bloch wave function of the $i$th band index and wave vector $k$. ML models were implemented with scikit-learn.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/50/eabd1076/DC1

**REFERENCES AND NOTES**

37. F. Nielsen, Hierarchical clustering, in Introduction to HPC with MPI for Data Science (Springer, 2016), pp. 195–211.

Acknowledgments
Funding: N.C.F. was supported by the U.S. Department of Defense through the National Defense Science and Engineering Graduate Fellowship program. V.B.S. acknowledges support from grants EFMA-542879 and CMMA-1727717 from the NSF and also W911NF-16-1-0447 from the Army Research Office. S.M.G. and integration with the Materials Project infrastructure was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231 (Materials Project program KC23MP). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under contract no. DE-AC02-05CH11231. Author contributions: N.C.F., M.K.H., J.M.M., S.M.G., K.A.P., and V.B.S. supervised the research. N.C.F., M.K.H., and J.M.M. designed the analyses. N.C.F. conducted the analyses and wrote the manuscript. K.A.P. and V.B.S. supervised the research. N.C.F., M.K.H., J.M.M., S.M.G., K.A.P., and V.B.S. analyzed the results and edited the manuscript. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The data used in this study are available at https://materialsproject.org. Machine learning models, data used in model training, and an example Jupyter notebook are available on GitHub at https://github.com/ncfrey/magnetic-topological-materials. Additional data related to this paper may be requested from the authors.

Submitted 2 June 2020
Accepted 26 October 2020
Published 9 December 2020
10.1126/sciadv.abb1076

Citation: N. C. Frey, M. K. Horton, J. M. Munro, S. M. Griffin, K. A. Persson, V. B. Shenoy, High-throughput search for magnetic and topological order in transition metal oxides. Sci. Adv. 6, eabb1076 (2020).
High-throughput search for magnetic and topological order in transition metal oxides
Nathan C. Frey, Matthew K. Horton, Jason M. Munro, Sinead M. Griffin, Kristin A. Persson and Vivek B. Shenoy

Sci Adv 6 (50), eabd1076.
DOI: 10.1126/sciadv.abd1076

ARTICLE TOOLS
http://advances.sciencemag.org/content/6/50/eabd1076

SUPPLEMENTARY MATERIALS
http://advances.sciencemag.org/content/suppl/2020/12/07/6.50.eabd1076.DC1

REFERENCES
This article cites 57 articles, 5 of which you can access for free
http://advances.sciencemag.org/content/6/50/eabd1076#BIBL

PERMISSIONS
http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title Science Advances is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).