Complex spin Hamiltonian represented by an artificial neural network

Hongyu Yu, Changsong Xu, Xueyang Li, Feng Lou, L. Bellaiche, Zhenpeng Hu, Xingao Gong, and Hongjun Xiang

Complex Spin Hamiltonian Represented by Artificial Neural Network

Hongyu Yu,[a,b,c]* Changsong Xu,[a,b,d]* Xueyang Li,[a,b] Feng Lou,[a] L. Bellaiche,[d] Zhenpeng Hu,[c] Xingao Gong,[a,b] Hongjun Xiang[a,b,†]

[a]*Key Laboratory of Computational Physical Sciences (Ministry of Education), Institute of Computational Physical Sciences, and Department of Physics, Fudan University, Shanghai 200433, China*
[b]*Shanghai Qi Zhi Institute, Shanghai 200030, China*
[c]*School of Physics, Nankai University, Tianjin 300071, China*
[d]*Physics Department and Institute for Nanoscience and Engineering, University of Arkansas, Fayetteville, Arkansas 72701, USA*

Abstract: The effective spin Hamiltonian method is very useful for simulating and understanding the behavior of magnetism. However, it is not easy to construct an appropriate spin Hamiltonian for a magnetic system, especially for complex magnets such as itinerant topological magnets. Here, we put forward a machine learning (ML) approach, applying an artificial neural network (ANN) and a local spin descriptor to construct effective spin Hamiltonian for any magnetic systems. The obtained Hamiltonians include an explicit Heisenberg part and an implicit non-linear ANN part. Such a method successfully reproduces artificially constructed models and also accurately describes the itinerant magnetism of bulk Fe_3GeTe_2. Our work paves a new way for investigating complex magnetic phenomena (e.g., skyrmions) using ML techniques.

Main text

Magnetism covers a major area of condensed matter physics and is of both fundamental and technical importance. Novel magnetic phases, such as skyrmion states, magnetic Weyl semimetals, spin liquids, and spin glasses, have been attracting a lot of scientific attention in recent years [1–9]. These diverse spin textures and colorful physics are driven by magnetic interactions of various forms. Effective spin Hamiltonian containing necessary interactions is a low-energy approximation that is widely adopted to describe the magnetic properties of various systems [10–12]. A well-known example is the Heisenberg model, $H_{spin} = \sum_{i>j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$, where J_{ij} denotes the coupling strength between spins $\mathbf{S}_i$ and $\mathbf{S}_j$, and it can be obtained from *ab initio* calculations or fitting to experiments. Such a model is frequently adopted since it captures the basic physics of ferromagnetism and antiferromagnetism, as well as, frustrations in most systems. However, such a model becomes insufficient when novel mechanisms come into play. For example, novel states arise from Dzyaloshinskii–Moriya interaction (DMI) [13] and Kitaev interaction [14] that is related to spin-orbit coupling (SOC), and it is also challenging to describe systems in which many-body and higher-order interactions unrelated to SOC are crucial. For instance, (i) biquadratic interactions are reported to widely exist in 2D magnets [15,16], (ii) three/four-body interactions are found to stabilize some skyrmions [17], and (iii) topological orbital magnetism with even higher orders are predicted to play important roles in non-coplanar spin patterns [18]. There may exist numerous different many-body and higher-order interactions, which makes it almost impossible to determine the exact form and the interacting parameters of the effective spin Hamiltonian [19]. Therefore, a new method for constructing an effective spin Hamiltonian for *any* system is highly desired to accurately describe complex magnetism.

The fast-developing machine learning potential (MLP) method is a promising solution. MLP built with artificial neural networks (ANN) has demonstrated its significant advantages in accuracy and efficiency [20–26]. It has been proven to be a powerful tool to solve problems in simulations of non-magnetic materials [20,27–29]. A central concept in the MLP method is the descriptor which serves as the suitable input representation of atomistic systems [30]. To the best of our knowledge, almost all reported descriptors do not consider the intrinsic spin property, which dominates the

behavior of the magnetic systems. Descriptors without information of spin fail to capture the spin interactions in magnetic materials. Therefore, spin descriptors for describing the spin configurations of magnetic materials are necessary for machine learning modelling. The application of MLPs to spin systems has just started and is still in its infancy. Recent attempts include (i) the magnetic moment tensor potentials (mMTP) that well describe the single element magnetic systems [32] and (ii) the spin-dependent atom-centered symmetry functions (sACSF) [28]. However, both methods only deal with simple collinear magnetic orders, which prevent the study of complex noncollinear states (e.g., spirals, skyrmions, and bimerons). Hence, a general MLP method dedicated to spin systems with complex interactions and colorful spin patterns is still highly desired.

In this work, we present a general machine learning method to construct spin Hamiltonians based on ANN. We propose a new local spin descriptor to describe spin configurations where spins can freely rotate to any three-dimensional direction. Our spin Hamiltonian contains an explicit Heisenberg part and an implicit nonlinear part to take into account all other possible forms of spin interactions. Such machine learning method can not only reproduce any artificially created models with peculiar complex interactions but also well describe the complex magnetism of Fe_3GeTe_2, where ferromagnetism, diverse spirals and skyrmionic states are observed [31–33].

Machine learning spin Hamiltonian approach. In this work, we aim at building the machine learning spin Hamiltonian (MLSH) H for a given atomic structure, which can be subsequently used to obtain the thermodynamic and kinetic properties of the spin system by performing Monte-Carlo simulations or spin dynamic simulations. In principle, one can express the MLSH using solely ANN. However, we find that it is beneficial (see tests below) if the MLSH contains explicitly the isotropic Heisenberg spin Hamiltonian besides the ANN part: $H = H_{HB} + H_{NN}$，where the Heisenberg part $H_{HB} = \sum_{i>j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$ (J_{ij} are the exchange interactions and $\mathbf{S}_i$ are spin vectors) and H_{NN} is the remaining energy contribution as described by the ANN. The Heisenberg interaction is of the second-order two-body form and usually dominates in most magnets. However, many-body and higher-order spin interactions are important in many interesting systems such as itinerant metallic magnets. The ANN contribution

H_{NN} is intended to describe such complicated interactions. Such treatment of separating Hamiltonian into explicit Heisenberg and implicit ANN parts results in a good balance between efficiency and accuracy.

The whole process of our method is demonstrated below. First, the data of different spin configurations and corresponding energies are generated. The data can be calculated with non-collinear spin constrained density functional theory (DFT). Then, the spin configurations are transformed into spin descriptors. MLSH is built with the Heisenberg part and a neural network contribution whose input is the spin descriptors. With the trained MLSH, the Monte Carlo simulation is carried out for thermodynamic theories. More training details of the MLSH are presented in the Supplementary Materials (SM) [34] (see, also, references [35–40] therein).

Similar to the case of machine learning atomic potentials [20,41,42], the ANN part of the spin Hamiltonian can be written as a sum of contributions from individual central spins:

$$H_{NN} = \sum_i H_{NN,i}(\boldsymbol{D}_i), \qquad \text{Eq. (1)}$$

where $\boldsymbol{D}_i$ is the descriptor vector that describes the local spin environment of the i-th central spin, and $H_{NN,i}$ is a function expressed in terms of ANN. A natural way to choose central spins in H_{NN} is that each spin in the system acts as a central spin. The form (*i.e.*, ANN parameters) of $H_{NN,i}$ depends on the local atomic environment around the i-th central spin. As inequivalent Wyckoff positions have different local atomic environments, we can classify the central spin according to the Wyckoff position that it belongs to. If the crystal has n inequivalent magnetic Wyckoff positions, the ANN part of the Hamiltonian can be written as $H_{NN} = \sum_{\mu=1}^{n} \sum_{\gamma} H_{NN,\mu}(\boldsymbol{D}_{i=[\mu,\gamma]})$, where μ and γ are the spin class index and index of spins within the μ-th spin class, respectively.
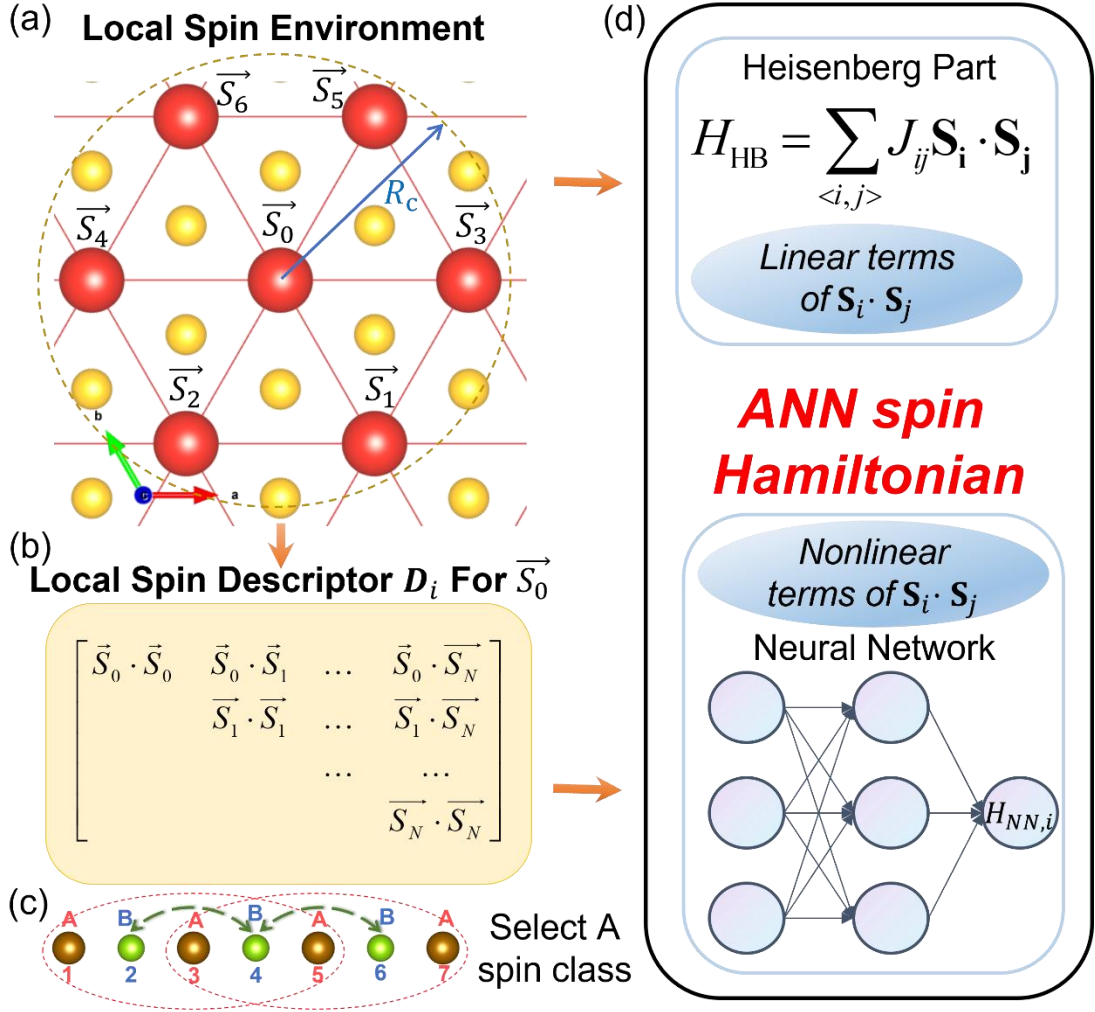
Figure 1. Workflow of the present ANN method to obtain spin Hamiltonians. (a) Illustration of the local structure centered on spin S_0 within a cutoff radius of R_c. The CuMO$_2$ structure is shown here as an example, where the red balls indicate magnetic sites of M. (b) The local spin descriptor centered on spin S_0, involving all neighboring spins within cutoff radius R_c. (c) Demonstration that one set of NN is enough for two different spin classes, as the descriptor of A spin class includes the spin interaction information of the B site. (d) The whole Hamiltonian that contains an explicit Heisenberg part and an implicit part, which includes all other forms of interactions. The total energy consists of energies from both parts.

Let us now concentrate on the spin descriptors D_i that depends not only on the central spin S_i itself, but also on the N_i neighboring spin vectors of S_i. The reason why we need to transform the spin vectors into descriptors instead of directly using the spin

vectors as inputs to the ANN is that the symmetry properties of the spin Hamiltonian can be enforced strictly through the descriptors. In this work, we mainly focus on the usual case where the spin-orbit coupling (SOC) effects can be neglected. To consider the SOC effects within our framework (*e.g.*, DMI, Kitaev interaction and/or single ion anisotropy terms), we can add these terms into the explicit Heisenberg part. The ANN part of the Hamiltonian H_{NN} is intended to deal with the high-order spin interactions in the absence of SOC effects. In the absence of SOC effects, the spin interaction energy is invariant under any simultaneous global rotation of all the spin vectors. In addition, the spin interaction energy is invariant under the time-reversal operation that reverses all the spin directions. To satisfy these conditions, each element of the descriptor vector is set to a dot product between two spin vectors $\mathbf{S}_p \cdot \mathbf{S}_q$. To be more specific, let us consider the spin descriptor $\boldsymbol{D}_0$ of the central spin 0 with N_0 spin neighbors within the cutoff distance R_c (see Fig. 1). We define the descriptor vector $\boldsymbol{D}_0$ with the elements taken from the upper triangular matrix of $(\mathbf{S}_p \cdot \mathbf{S}_q)_{N_0 \times N_0}$: $\boldsymbol{D}_0 = (\mathbf{S}_0 \cdot \mathbf{S}_0, \ldots, \mathbf{S}_0 \cdot \mathbf{S}_{N_0}, \mathbf{S}_1 \cdot \mathbf{S}_1, \ldots, \mathbf{S}_1 \cdot \mathbf{S}_{N_0}, \ldots, \mathbf{S}_{N_0} \cdot \mathbf{S}_{N_0})^T$. There is another symmetry that the descriptor must be able to account for. If the Wyckoff position that the central spin belongs to displays the point group G symmetry with m operations, then any operation R of G on the spin configuration $(\mathbf{S}_0, \mathbf{S}_1, \ldots, \mathbf{S}_{N_0})$ will leave the spin interaction energy invariant. Note that R will not only change the spin directions but may also result in a permutation of the spin sites. To take the point group symmetry into account, we first generate all m spin configurations by applying m operations to a given spin configuration, then we select a representative spin configuration among the m spin configurations according to some predefined criteria (see SM [34]). The representative spin configuration is used to construct the spin descriptor $\boldsymbol{D}_0$. In this way, these m spin configurations will be guaranteed to have the same spin interaction energy. As the symmetry analysis is performed automatically, the construction of descriptors is an automatic process without any manual intervention.

To better demonstrate the process of constructing the spin descriptor, a simple example is presented. Here, we take $CuMO_2$ (M is a magnetic ion) with the $CuAlO_2$-type structure as an example to show how the spin descriptor is constructed. For simplicity, the cutoff radius is chosen to be 3 Å including only the nearest neighbor magnetic sites. We consider the descriptor of site 0 shown in Fig. 2. First, we analyze

the symmetry of atoms in the unit cell. As only site M in $CuMO_2$ has magnetization and all M sites are equivalent, the neighbor information for each M site is the same. The M ion at site 0 in Fig. 2 has 6 neighbor magnetic atoms. The Wyckoff position that the site 0 belongs to displays the point group D_{3d} symmetry with 12 operations including $\{e, (s_6^1, s_6^5), (c_3^1, c_3^2), i, 3c_2^{'}, 3\sigma_d\}$. Then, we construct a local spin descriptor for site 0. For illustration purposes, we take the spin configuration in Fig. 2(a) as the input. With the operation R of D_{3d} on the spin configuration, we get 12 different spin configurations and descriptors $(\boldsymbol{D}_0^1, \dots, \boldsymbol{D}_0^{12})$ corresponding to 12 point group operations. Two spin configurations for different point group operations are shown in Fig. 2. More precisely, Fig. 2(a) is the input spin structure, while Fig. 2(b) is the spin configurations corresponding to the operation s_6^1. Their corresponding descriptors are:

$$
(a): \begin{bmatrix} 1 & 1 & -1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (b): \begin{bmatrix} 1 & -1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

These 12 descriptors should have the same energies: $H_{NN}(\boldsymbol{D}_0^i) = H_{NN}(\boldsymbol{D}_0^j)$ with $i,j=1,\dots,12$ due to the local point group symmetry. To satisfy these conditions, we select a representative descriptor among all 12 descriptors so that $H_{NN}(\boldsymbol{D}_0^i) = H_{NN}(\boldsymbol{D}_0^{rep})$ with $i=1,\dots,12$. In principle, any ordering scheme of the descriptors should work. Here, we adopt the following ordering scheme. The descriptor with the smallest $\mathbf{S}_0 \cdot \mathbf{S}_1$ is set as the final descriptor. If the values of $\mathbf{S}_0 \cdot \mathbf{S}_1$ are the same, other elements in the matrix are used to compare from $\mathbf{S}_0 \cdot \mathbf{S}_2$ to $\mathbf{S}_0 \cdot \mathbf{S}_n$. The descriptor of (b) is selected as the final descriptor.
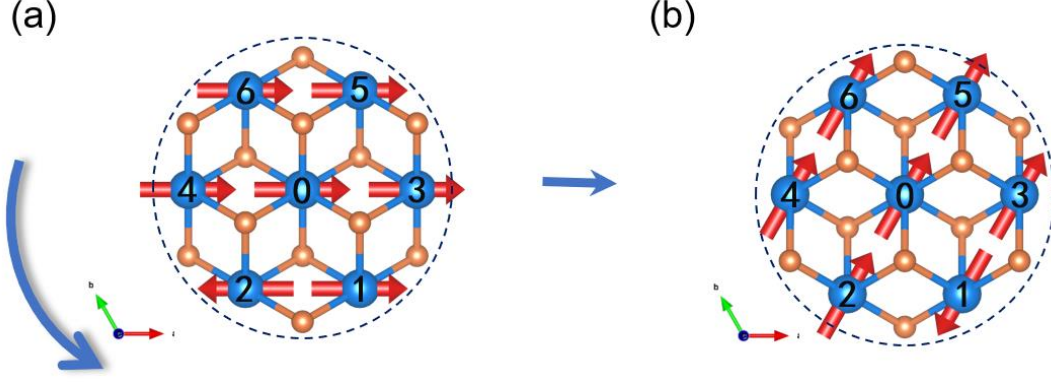
Figure 2. Schematization of the local structure that is involved in a descriptor. (a) The input spin structure under e point group operation. (b) The spin structure under s_6^1 point group operation with site 0 as the center. The descriptor with the lower value of $\mathbf{S}_0 \cdot \mathbf{S}_1$ are selected as the final descriptor. Blue balls represent the magnetic sites with the red vectors denoting the spin vectors. The order of the spins that appear in the descriptor is denoted by the number imposed on the magnetic sites.

In our above discussion on the ANN part of the spin Hamiltonians (see Eq. 1), every spin in the system acts as a central spin. We find (see tests below) that the efficiency can be significantly improved by reducing the number of central spins in the case of multiple inequivalent magnetic Wyckoff positions (i.e., $n > 1$). Consider the crystal [see Fig. 1(c)] with two inequivalent magnetic Wyckoff positions A and B, one can select A as the central spin but leave out B. This is because the spin interactions involving the B spin are included in that related to the neighboring central spins (A spins here) within a cutoff distance R_c. Reducing the number of central spins not only reduces the number of ANN parameters and thus the risk of overfitting but also speeds up the evaluation of the total spin interaction energies. We note that this idea of using a reduced number of central sites can also be adopted to construct the highly efficient machine learning atomic potential for simulating atomic process (*e.g.*, conduction of heat, ferroelectric phase transition) that maintains the basic framework of the crystal (see SM [34]).
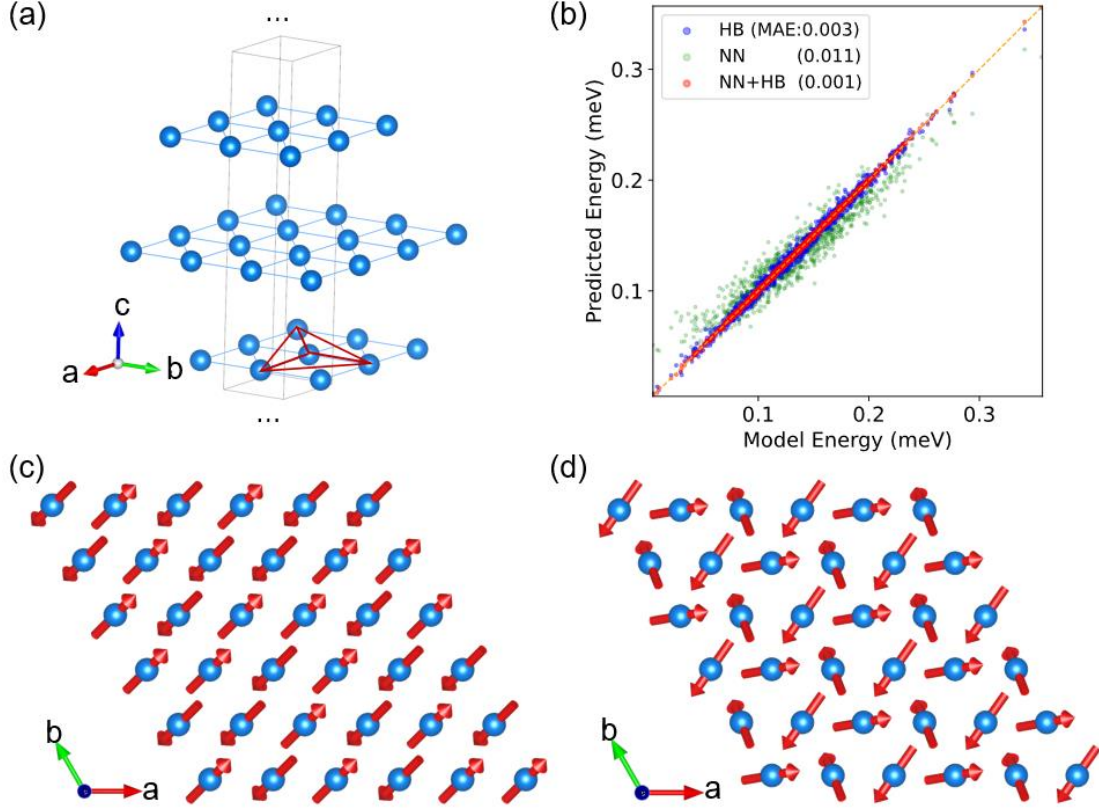
Figure 3. (a) Schematization of the CuMO$_2$-type structure, where only the magnetic M sites are shown. The atoms connected by red lines form an example of the four-body cluster, for which the fourth-order interaction is considered in Eq. (2). (b) Comparison of the energies among three different fitting methods – only Heisenberg terms, only NN and both Heisenberg and NN. The Mean Absolute Error (MAE) of each fitting method is shown in the legend. (c) The (↑↑↓↓) AFM magnetic ground state of the exact model Hamiltonian is reproduced by the HB+NN method. (d) The frustration induced a short period spiral, which is the ground state predicted by the pure Heisenberg method.

An artificial model. To validate our machine learning approach for constructing spin Hamiltonian, we employ a simple model Hamiltonian. This model is intended to describe the spin interactions between magnetic ions that occupy the Al sites of the CuAlO$_2$-type [43] structure. The model spin Hamiltonian contains both intralayer interactions within the triangular lattice and interlayer couplings:

$$H = \sum_{<ij>_n}^{n=1,\dots,5} J_n \, \mathbf{S}_i \cdot \mathbf{S}_j + \sum_{<ij>_n}^{n=1,5,7} K_n \left(\mathbf{S}_i \cdot \mathbf{S}_j \right)^2 + \sum_{<ijkl>_1} L_1 \left(\mathbf{S}_i \cdot \mathbf{S}_j \right)\left(\mathbf{S}_k \cdot \mathbf{S}_l \right). \qquad (2)$$

Here (i) J_n are the 2nd-order Heisenberg spin interactions, where $J_1 - J_3$ are AFM intralayer couplings and are thus strongly frustrated, while J_4 are relatively weak interlayer couplings and J_5 is set to zero but included while fitting for testing purposes; (ii) K_n are the 4th-order biquadratic interactions for the 1st, 5th and 7th nearest neighboring pairs; and (iii) L_1 represents the four-body fourth-order interaction for the triangular cluster with the fourth site in the center of the triangle as shown in Fig. 3(a). The values of J, K, and L as listed in Table I, are set in such a way that leads to a complex spin model with strong frustrations and high-order interactions. Our MC simulations using the exact model Hamiltonian suggest that our model with the parameters from Table I results in a nontrivial (↑↑↓↓) antiferromagnetic (AFM) ground state, as shown in Fig. 3(c).

We now try to reproduce the results from the model Hamiltonian by building our MLSH. The data set of 10,000 fully random spin configurations is generated based on a $4 \times 4 \times 2$ supercell with 96 magnetic sites (*i.e.*, six layers of 4×4 triangular lattices), and their energies are obtained from Eq. (2). Our MLSH contains five Heisenberg terms and a multi-layer fully connected neural network. During the fitting, 90% of the data are used for training and the remaining 10% for testing. To check whether the explicit inclusion of the Heisenberg interaction is crucial to the accuracy of the MLSH, we consider three different forms of the spin Hamiltonian: (i) HB Hamiltonian which only includes the Heisenberg interactions (five terms for this system) (ii) NN Hamiltonian which includes the ANN part, and (iii) NN+HB full Hamiltonian which includes both Heisenberg and ANN parts.

As shown in Fig. 3(b), the comprehensive NN+HB Hamiltonian is found to display a much smaller (1.03×10^{-5} meV/site) mean absolute error (MAE, defined as $\frac{\sum |f(x_i) - y_i|}{n}$) in total energy than that (3.4×10^{-5} meV/site) of the HB Hamiltonian and that (1.1×10^{-4} meV/site) of the NN Hamiltonian. Further, by performing Monte Carlo (MC) simulations, we find that the NN+HB Hamiltonian can correctly reproduce the (↑↑↓↓) AFM ground state [see Fig. 3(c)], while the HB Hamiltonian predicts incorrectly a non-collinear ground state [see Fig. 3(d)]. The Néel temperature obtained with the NN-HB Hamiltonian is also in agreement with that obtained with the exact model Hamiltonian. Note that the MLSH can also well describe non-collinear exotic spin states. For example, by tuning J_1 from 3.659 meV to 18.295 meV, the ground state of the exact

model becomes a noncollinear order spiral state with q-vector as (1/3, 0) in triangular lattice plane, which can be well reproduced by the MLSH approach.

Table I. Original and fitted magnetic coefficients for the artificial model in Eq. (1) and the real compound of Fe_3GeTe_2 (Energy unit meV). The subscript i, j of bulk Fe_3GeTe_2 part represents the atom index of the two-body Heisenberg interaction $J_{i,j}$ in Fig. 1a. Note that "HB" denotes the results from pure Heisenberg model, while "NN+HB" refers to the model with Heisenberg part and implicit part treated by ANN separately.

| | Artificial model Eq. (1) | | | | Bulk Fe_3GeTe_2 | |
	Model	NN+HB	HB		NN+HB	HB
J_1	3.659	3.714	3.659	$J_{1,3}$	-0.486	-0.474
J_2	1.117	0.988	1.114	$J_{1,2}$	-0.207	-0.199
J_3	1.728	1.779	1.729	$J_{2,2}$	0.069	0.065
J_4	1.003	0.995	0.998	$J_{1,1}$	0.034	0.023
J_5	0.000	-0.002	0.003	$J_{1,3(2)}$	-0.078	-0.069
K_1	-1.386	\	\	$J_{1,6}$	0.007	0.010
K_5	-0.376	\	\	$J_{2,6}$	-0.008	-0.006
K_7	-0.143	\	\	$J_{2,5}$	-0.006	-0.007
L_1	0.316	\	\			
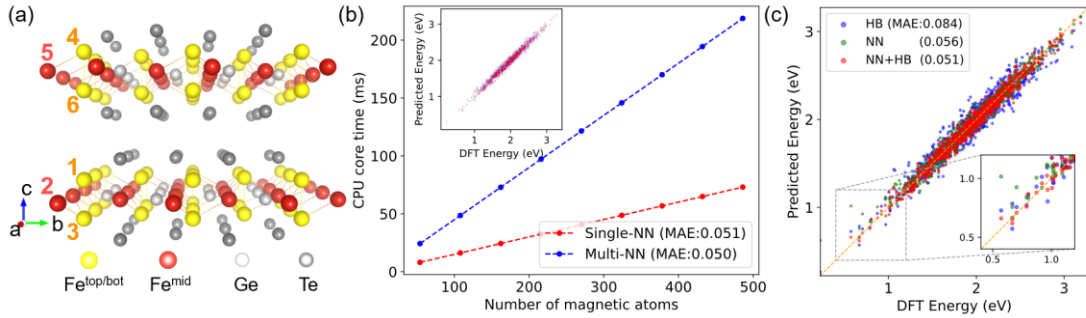
Figure 4. (a) Crystal structure of bulk Fe_3GeTe_2. The unit cell consists of two Van der Waals layers with six Fe atoms, that are categorized into two types, two Fe^{mid} and four Fe^{top}/Fe^{bot}. **(b)** Comparison of the CPU time between the single-NN method and two-NN method, as a function of magnetic atom number. Both sets of simulations are performed on Intel Xeon Gold 6244. The inset indicates similar accuracy of the two methods. **(c)** Comparison among energies from DFT and three different fitting strategies. The MAE of each method is shown in the legend.

Metallic ferromagnet Fe₃GeTe₂. Fe$_3$GeTe$_2$ (FGT) is a metallic Van der Waals layered ferromagnetic (FM) material [see crystal structure in Fig. 4(a)] for which the Curie temperature is experimentally determined to be 150~230 K [44,45] and can further be tuned to near room-temperature with gate voltage [31]. Recently, FGT thin films of about a few tens of nanometers in thickness are intensively studied and skyrmions and domain walls are widely observed there [32,33,44,46,47]. These interesting observations suggest that complex high-order interactions are essential besides second-order Heisenberg interactions [48–50]. We thus apply the presently developed ANN method to explore the magnetism of bulk FGT.

As shown in Fig. 4(a), each Van de Waals layer of bulk FGT contains three layers of Fe ions, where the top and bottom Fe ions are equivalent but are different from the middle Fe ion. In another word, there are two kinds of inequivalent magnetic Wyckoff positions, *i.e.*, $n = 2$. To see whether the strategy of reducing the number of central spins works, we consider two cases: (i) use all Fe sites as central spins (refers to as multi-NN strategy) *versus* (ii) only use the middle Fe sites as central spins (refers to as single-NN strategy). For our NN-HB Hamiltonian, the Heisenberg part includes both intralayer and interlayer couplings interactions up to 18 Å, while the ANN part only considers intralayer interactions with $R_c = 9$ Å. The data set consists of 8000 random spin configurations within a $3 \times 3 \times 1$ supercell, which contains 54 Fe, and their energies are computed with DFT and see details in Methods of SM [34]. Again, 90% of the data are used for training and 10% for testing. We find that the MLSH obtained with the single-NN strategy is as accurate as that with the multi-NN strategy [see insets of Fig. 4(c)]. However, the single-NN strategy is much more efficient than the multi-NN strategy for two reasons. In the training process, it takes much less time as the number of parameters in the single-NN strategy is one half of that in the multi-NN strategy. When using the MLSH, the calculations of the spin interaction energies in the case of single-NN strategy costs only 1/3 computational time of that in the case of the multi-NN strategy because the number of central spins in the formal case is 1/3 of that in the latter case [see Fig. 4(c)]. Hereafter, we will thus adopt the more efficient single-NN strategy for FGT.

Similar to the case of the model spin Hamiltonian, we find that the NN+HB

Hamiltonian is more accurate than the HB Hamiltonian and NN Hamiltonian: the MAE in these three cases are 0.94 meV/Fe, 1.56 meV/Fe, and 1.04 meV/Fe, respectively; in addition, the errors in the predicted energies for some low-energy spin configurations are rather large when using the NN Hamiltonian [see insets of Fig. 4(b)]. These tests suggest that the NN+HB Hamiltonian is desirable in general cases. We also compare the performance of another newly developed Machine Learning Method for Constructing Hamiltonian (MLMCH) [10] considering up to four-body and four-order magnetic interactions. In MLMCH, the model consists of 44 terms including the Heisenberg interaction and two-body fourth-order interaction up to the first 20 neighbors, three-body fourth-order up to the first 11 neighbors, and four-body fourth-order up to the first 4 neighbors. It gives a worse MAE as 1.07 meV/Fe in the test dataset which shows the priority of our MLSH method. It also proves that our new MLSH methods can take into account many-body and high-order magnetic interaction beyond the above-mentioned magnetic interactions.

The finite-temperature properties of FGT are then simulated with parallel tempering Monte Carlo (PTMC) [51,52] (see PTMC details in SM [34]). The ANN model from NN+HB scheme (iii) results in a Curie temperature of 554 K. Such value is based on the data from DFT with no Hubbard U correction, while further calculations (See details in SM [34]) indicate that finite U correction may lead to Curie temperature that compares well with experiment [45].

In summary, we develop a machine learning method to construct spin Hamiltonians. Such a method involves a local spin descriptor and adopts an artificial neural network to describe the complex spin interactions of any form. During the fitting, the Heisenberg part and other interactions are treated separately, which leads to improvement in accuracy. The effectiveness and accuracy of such a method are verified with an artificial complex Hamiltonian and also with the interesting itinerant magnet Fe_3GeTe_2. The presently developed method could give thermodynamic properties of any magnets in an automatic way and could be adopted to carry on the spin dynamics in the future. We expect that our method may find wide use in the study of magnetism.

* Contributed equally to this work.

† hxiang@fudan.edu.cn

Reference

[1] U. K. Rößler, A. N. Bogdanov, and C. Pfleiderer, *Spontaneous Skyrmion Ground States in Magnetic Metals*, Nature **442**, 7104 (2006).

[2] C. Pappas, E. Lelièvre-Berna, P. Falus, P. M. Bentley, E. Moskvin, S. Grigoriev, P. Fouquet, and B. Farago, *Chiral Paramagnetic Skyrmion-like Phase in MnSi*, Phys. Rev. Lett. **102**, 197202 (2009).

[3] A. Neubauer, C. Pfleiderer, B. Binz, A. Rosch, R. Ritz, P. G. Niklowitz, and P. Böni, *Topological Hall Effect in the A Phase of MnSi*, Phys. Rev. Lett. **102**, 186602 (2009).

[4] T.-H. Han, J. S. Helton, S. Chu, D. G. Nocera, J. A. Rodriguez-Rivera, C. Broholm, and Y. S. Lee, *Fractionalized Excitations in the Spin-Liquid State of a Kagome-Lattice Antiferromagnet*, Nature **492**, 7429 (2012).

[5] S. Yan, D. A. Huse, and S. R. White, *Spin-Liquid Ground State of the S = 1/2 Kagome Heisenberg Antiferromagnet*, Science **332**, 1173 (2011).

[6] C. Castelnovo, R. Moessner, and S. L. Sondhi, *Magnetic Monopoles in Spin Ice*, Nature **451**, 7174 (2008).

[7] S. T. Bramwell and M. J. P. Gingras, *Spin Ice State in Frustrated Magnetic Pyrochlore Materials*, Science **294**, 1495 (2001).

[8] K. Binder and A. P. Young, *Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions*, Rev. Mod. Phys. **58**, 801 (1986).

[9] E. Liu et al., *Giant Anomalous Hall Effect in a Ferromagnetic Kagome-Lattice Semimetal*, Nat. Phys. **14**, 1125 (2018).

[10] X. Li, F. Lou, X. Gong, and H. Xiang, *Constructing Realistic Effective Spin Hamiltonians with Machine Learning Approaches*, New J. Phys. **22**, 053036 (2020).

[11] X. Li, H. Yu, F. Lou, J. Feng, M.-H. Whangbo, and H. Xiang, *Spin Hamiltonians in Magnets: Theories and Computations*, Molecules **26**, 4 (2021).

[12] J. Li, J. Feng, P. Wang, E. Kan, and H. Xiang, *Nature of Spin-Lattice Coupling in Two-Dimensional CrI3 and CrGeTe3*, Sci. China Phys. Mech. Astron. **64**, 286811 (2021).

[13] T. Moriya, *New Mechanism of Anisotropic Superexchange Interaction*, Phys. Rev. Lett. **4**, 228 (1960).

[14] C. Xu, J. Feng, M. Kawamura, Y. Yamaji, Y. Nahas, S. Prokhorenko, Y. Qi, H. Xiang, and L. Bellaiche, *Possible Kitaev Quantum Spin Liquid State in 2D Materials with S=3/2*, Phys. Rev. Lett. **124**, 087205 (2020).

[15] A. Kartsev, M. Augustin, R. F. L. Evans, K. S. Novoselov, and E. J. G. Santos, *Biquadratic Exchange Interactions in Two-Dimensional Magnets*, Npj Comput. Mater. **6**, 150 (2020).

[16] J. Y. Ni, *Giant Biquadratic Exchange in 2D Magnets and Its Role in Stabilizing*

Ferromagnetism of NiCl2 Monolayer, Phys. Rev. Lett. **127**, 247204 (2021).

[17] S. Paul, S. Haldar, S. von Malottki, and S. Heinze, *Role of Higher-Order Exchange Interactions for Skyrmion Stability*, Nat. Commun. **11**, 4756 (2020).

[18] S. Grytsiuk, J.-P. Hanke, M. Hoffmann, J. Bouaziz, O. Gomonay, G. Bihlmayer, S. Lounis, Y. Mokrousov, and S. Blügel, *Topological–Chiral Magnetic Interactions Driven by Emergent Orbital Magnetism*, Nat. Commun. **11**, 511 (2020).

[19] E. Svanidze et al., *An Itinerant Antiferromagnetic Metal without Magnetic Constituents*, Nat. Commun. **6**, 7701 (2015).

[20] H. Wang, L. Zhang, J. Han, and W. E, *DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics*, Comput. Phys. Commun. **228**, 178 (2018).

[21] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons*, Phys. Rev. Lett. **104**, 136403 (2010).

[22] R. M. Balabin and E. I. Lomakina, *Support Vector Machine Regression (LS-SVM)—an Alternative to Artificial Neural Networks (ANNs) for the Analysis of Quantum Chemistry Data?*, Phys. Chem. Chem. Phys. **13**, 11710 (2011).

[23] A. P. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles, and G. J. Tucker, *Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials*, J. Comput. Phys. **285**, 316 (2015).

[24] A. V. Shapeev, *Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials*, Multiscale Model. Simul. **14**, 1153 (2016).

[25] J. S. Smith, O. Isayev, and A. E. Roitberg, *ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost*, Chem. Sci. **8**, 3192 (2017).

[26] R. Drautz, *Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials*, Phys. Rev. B **99**, 014104 (2019).

[27] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Machine Learning a General-Purpose Interatomic Potential for Silicon*, Phys. Rev. X **8**, 041048 (2018).

[28] V. L. Deringer, N. Bernstein, G. Csányi, C. Ben Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, and S. R. Elliott, *Origins of Structural and Electronic Transitions in Disordered Silicon*, Nature **589**, 7840 (2021).

[29] L. Hu, B. Huang, and F. Liu, *Atomistic Mechanism Underlying the Si(111)-(7×7) Surface Reconstruction Revealed by Artificial Neural-Network Potential*, Phys. Rev. Lett. **126**, 176101 (2021).

[30] L. Himanen, M. O. J. Jäger, E. V. Morooka, F. Federici Canova, Y. S. Ranawat, D. Z. Gao, P. Rinke, and A. S. Foster, *DScribe: Library of Descriptors for Machine Learning in Materials Science*, Comput. Phys. Commun. **247**, 106949 (2020).

[31] Y. Deng et al., *Gate-Tunable Room-Temperature Ferromagnetism in Two-Dimensional Fe3GeTe2*, Nature **563**, 94 (2018).

[32] B. Ding, Z. Li, G. Xu, H. Li, Z. Hou, E. Liu, X. Xi, F. Xu, Y. Yao, and W. Wang, *Observation of Magnetic Skyrmion Bubbles in a van Der Waals Ferromagnet Fe3GeTe2*, Nano Lett. **20**, 868 (2020).

[33] Y. Wu, S. Zhang, J. Zhang, W. Wang, Y. L. Zhu, J. Hu, G. Yin, K. Wong, C. Fang,

C. Wan, X. Han, Q. Shao, T. Taniguchi, K. Watanabe, J. Zang, Z. Mao, X. Zhang, and K. L. Wang, *Néel-Type Skyrmion in WTe 2 /Fe 3 GeTe 2 van Der Waals Heterostructure*, Nat. Commun. **11**, 3860 (2020).

[34] See Supplemental Material at [URL will be inserted by the production group] for details of computation, ML methods, and training strategy, which includes Refs. [35-39].

[35] G. Kresse and J. Furthmüller, *Efficient Iterative Schemes for* Ab Initio *Total-Energy Calculations Using a Plane-Wave Basis Set*, Phys. Rev. B **54**, 11169 (1996).

[36] L. Prechelt, *Automatic Early Stopping Using Cross Validation: Quantifying the Criteria*, Neural Netw. **11**, 761 (1998).

[37] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, *Electron-Energy-Loss Spectra and the Structural Stability of Nickel Oxide: An LSDA+U Study*, Phys. Rev. B **57**, 1505 (1998).

[38] G. Kresse and D. Joubert, *From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method*, Phys. Rev. B **59**, 1758 (1999).

[39] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77**, 3865 (1996).

[40] J. Hubbard and B. H. Flowers, *Electron Correlations in Narrow Energy Bands*, Proc. R. Soc. Lond. Ser. Math. Phys. Sci. **276**, 238 (1963).

[41] H. Wang, X. Guo, L. Zhang, H. Wang, and J. Xue, *Deep Learning Inter-Atomic Potential Model for Accurate Irradiation Damage Simulations*, Appl. Phys. Lett. **114**, 244101 (2019).

[42] L. Zhang, J. Han, H. Wang, R. Car, and W. E, *Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics*, Phys. Rev. Lett. **120**, 143001 (2018).

[43] T. Ishiguro, A. Kitazawa, N. Mizutani, and M. Kato, *Single-Crystal Growth and Crystal Structure Refinement of CuAlO2*, J. Solid State Chem. **40**, 170 (1981).

[44] A. F. May, S. Calder, C. Cantoni, H. Cao, and M. A. McGuire, *Magnetic Structure and Phase Stability of the van Der Waals Bonded Ferromagnet Fe3−xGeTe2*, Phys. Rev. B **93**, 014411 (2016).

[45] H.-J. Deiseroth, K. Aleksandrov, C. Reiner, L. Kienle, and R. K. Kremer, *Fe3GeTe2 and Ni3GeTe2 – Two New Layered Transition-Metal Compounds: Crystal Structures, HRTEM Investigations, and Magnetic and Electrical Properties*, Eur. J. Inorg. Chem. **2006**, 1561 (2006).

[46] B. Chen, J. Yang, H. Wang, M. Imai, H. Ohta, C. Michioka, K. Yoshimura, and M. Fang, *Magnetic Properties of Layered Itinerant Electron Ferromagnet Fe3GeTe2*, J. Phys. Soc. Jpn. **82**, 124711 (2013).

[47] T.-E. Park et al., *Néel-Type Skyrmions and Their Current-Induced Motion in van Der Waals Ferromagnet-Based Heterostructures*, Phys. Rev. B **103**, 104410 (2021).

[48] B. Huang et al., *Layer-Dependent Ferromagnetism in a van Der Waals Crystal down to the Monolayer Limit*, Nature **546**, 270 (2017).

[49] M. Yang et al., *Creation of Skyrmions in van Der Waals Ferromagnet Fe3GeTe2 on (Co/Pd)n Superlattice*, Sci. Adv. **6**, eabb5157 (2020).

[50] C. Xu, X. Li, P. Chen, Y. Zhang, H. Xiang, and L. Bellaiche, *Assembling Diverse*

Skyrmionic Phases in Fe3GeTe2 Monolayers, Adv. Mater. **34**, 2107779 (2022).

[51] F. Lou, X. Y. Li, J. Y. Ji, H. Yu, J. S. Feng, X. G. Gong, and H. J. Xiang, *PASP: Property Analysis and Simulation Package for Materials*, J. Chem. Phys. **154**, 114103 (2021).

[52] D. J. Earl and M. W. Deem, *Parallel Tempering: Theory, Applications, and New Perspectives*, Phys. Chem. Chem. Phys. **7**, 3910 (2005).