ARTICLE   OPEN

Check for updates

Machine learning assisted derivation of minimal low-energy models for metallic magnets

Vikram Sharma [1✉], Zhentao Wang [1,2,4] and Cristian D. Batista [1,3]

We consider the problem of extracting a low-energy spin Hamiltonian from a triangular Kondo Lattice Model (KLM). The non-analytic dependence of the effective spin-spin interactions on the Kondo exchange excludes the use of perturbation theory beyond the second order. We then introduce a Machine Learning (ML) assisted protocol to extract effective two- and four-spin interactions. The resulting spin model reproduces the phase diagram of the original KLM as a function of magnetic field and single-ion anisotropy and reveals the effective four-spin interactions that stabilize the field-induced skyrmion crystal phase. Moreover, this model enables the computation of static and dynamical properties with a much lower numerical cost relative to the original KLM. A comparison of the dynamical spin structure factor in the fully polarized phase computed with both models reveals a good agreement for the magnon dispersion even though this information was not included in the training data set.

INTRODUCTION

Lattice models of fermions interacting with classical fields encompass different areas of knowledge, including quantum chemistry, condensed matter, and high-energy physics. This broad class of models poses some notoriously difficult numerical challenges. On one hand, Monte Carlo (MC) sampling of the classical field requires repeated diagonalization of the single-particle fermion matrix. On the other hand, it is difficult to eliminate size effects when the effective interactions between the classical degrees of freedom (CDOF) are orders of magnitude smaller than the bare interaction between the fermions and the CDOF[1]. While the first problem has been addressed by different approaches that reduce the numerical cost of the simulations[2–5], the second problem, which arises in the weak-coupling to intermediate-coupling regimes, is more difficult to solve. By "weak-coupling" we mean that the value of the interaction J between the classical and fermionic degrees of freedom is small in comparison to the characteristic energy scale of the fermions (e.g., the dominant hopping amplitude t of the conduction electrons of a KLM). Since this is the relevant regime for different incarnations of this class of models, such as 4f-electron materials described by KLMs, it is necessary to develop new approaches that can address this problem.

For the weak to intermediate-coupling regimes, the traditional approach for this class of models is perturbation theory. Effective interactions between the CDOF are obtained by expanding them in powers of J/t. For instance, the application of perturbation theory to the KLM leads to the celebrated RKKY interaction between localized moments[6–8]. Unfortunately, perturbation theory can not be extended beyond the second order because the corresponding diagrams diverge at $T = 0$[9–12], suggesting that the coefficients of n-spin interactions with $n \geq 4$ are non-analytic functions of the coupling constant. In general, the problem of integrating out fermionic degrees of freedom in presence of a Fermi surface is highly non-trivial because the effective interactions are expected to be non-analytic functions of the coupling

constant. However, phenomenological approaches based on minimal models that only include 2 and 4-spin interactions to explain the emergence of multi-Q magnetic orderings have demonstrated to be effective[10,12–16]. Based on the success of the RKKY model[6–8] and on the extended use of these minimal models, here we will conjecture that, for weak-enough coupling constant, one can still neglect n-spin interactions with $n > 4$ and approximate the effective two- and four-spin interactions by analytic functions. In other words, we will assume that, despite the non-analytic dependence of the coefficients of n-spin interactions on the coupling constant J, there is still a hierarchy of interactions, i.e., there is a regime where six and higher-spin interactions can be neglected in comparison to the terms including two and four-spin interactions. Moreover, we will also assume that non-analytic behavior in momentum space, caused by the long-range nature of the real space interactions, can be approximated by a sequence of analytic functions obtained by systematically increasing the range of the interactions. As we will discuss in this work, the verification of these conjectures is particularly relevant for addressing situations where the RKKY Hamiltonian is frustrated in the sense that the exchange interaction in momentum space $\tilde{\mathcal{J}}(\boldsymbol{q})$ is minimized by multiple symmetry related wave vectors $\boldsymbol{Q}_v$.

The problem of frustrated magnetic metals has multiple attractive aspects. For instance, four-spin interactions can stabilize non-coplanar orderings that induce nonzero Berry curvature of the reconstructed bands. This momentum space Berry curvature can in turn lead to a large topological Hall effect below the magnetic ordering temperature T_N[14,17–22]. Since T_N can be comparable to or even higher than room temperature, multiple experimental efforts are trying to achieve this goal[23]. An outstanding example is the search for skyrmion crystals (SkXs) in f-electron magnets. Field-induced SkXs with large topological Hall effect have been recently observed in the rare earth-based centrosymmetric materials Gd_2PdSi_3 and $Gd_3Ru_4Al_{12}$[24–28] that can be modeled by a KLM.

[1]Department of Physics and Astronomy, The University of Tennessee, Knoxville, TN 37996, USA. [2]School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA. [3]Quantum Condensed Matter Division and Shull-Wollan Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [4]Present address: Center for Correlated Matter and School of Physics, Zhejiang University, Hangzhou 310058, China. ✉email: vsharma7@vols.utk.edu

In a recent work[29], we have demonstrated that SkXs emerging from triangular lattice RKKY models with easy-axis anisotropy can naturally lead to values of the Hall conductivity that are comparable to the quantized value (e^2/h). The key observation is that the magnitude $Q \equiv |\boldsymbol{Q}_\nu|$ of magnetic ordering wave vector is dictated by the Fermi wave vector $Q \simeq 2k_F$. In a subsequent work[1], we used a systematic variational study of the triangular KLM to show that SkXs are ubiquitous phases of centrosymmetric metals with localized magnetic moments. This variational study, which is crucially important to eliminate undesirable finite-size effects, reveals mesoscale field-induced SkXs, whose stability range depends on the coupling strength J/t (t is the nearest-neighbor hopping of the KLM). These results are consistent with an increasing amount of numerical and experimental evidence in favor of the emergence of multi-$\boldsymbol{Q}$ orderings, including SkXs, in metallic six-fold symmetric layered materials comprising localized magnetic moments coupled via exchange interaction to conduction electrons[24–28].

The emergence of multi-$\boldsymbol{Q}$ orderings in magnetic systems with localized moments has stimulated different groups to propose effective four-spin interactions which can account for this phenomenon[12–16,30]. In view of the lack of a controlled analytical procedure to derive these effective interactions, the proposed effective models are mostly phenomenological and ad-hoc because they do not consider all the symmetry-allowed four-spin interactions. For instance, Hayami et al.[12,16] have proposed a phenomenological bilinear biquadratic model based on a trend that they observe in the divergent terms of the perturbative expansion. While these phenomenological approaches can offer useful insights due to their simplicity, they severely limit the predictive power of the original high-energy model. Similar considerations apply to the low-energy excitations (magnons) of each magnetically ordered state. It is then relevant to ask if there is an alternative, less biased, method that can output all symmetry-allowed effective 4-spin interactions while preserving the simplicity of phenomenological spin models and the predictive power of the original KLM. A method with these characteristics can be used to understand the origin of the different magnetic orderings induced by magnetic field and/or single-ion anisotropy, as well as to compute the low-energy magnon spectrum of each magnetic phase. Similarly to the case of Landau-Ginzburg expansions of the free energy near a critical point, the signs of the different symmetry allowed quartic terms offer valuable insights into the effective interactions among distinct modes. As we will see in this work, another important advantage is that simulations of the effective spin model turn out to be a few orders of magnitude faster than simulations of the original KLM.

The alternative approach that we propose here is inspired by a recent proposal for constructing effective low-energy Hamiltonians by supervised learning on energy[31]. In their work, Fujita et al. used supervised learning to derive an effective spin-1/2 Hamiltonian in the strong-coupling limit of a half-filled Hubbard model. In contrast to the scenario under our current consideration, this problem admits a well-behaved perturbative expansion, which can be used to corroborate the success of the supervised learning algorithm. The classical nature of the spin degrees of freedom in the KLM that we consider here introduces another important difference because the spectrum of a classical spin model is always continuous, while the spectrum of a quantum spin model is discrete on finite lattices. In other words, each product of coherent spin states is an eigenstate in the classical ($S \to \infty$) limit of the KLM, while this is, of course, not true for the quantum case. While one can always fit the lowest energy M eigenstates of the discrete spectrum of the high-energy model on a finite lattice and determine the optimal parameters of the low-energy *quantum* spin model, this is not possible for classical spins due to the continuous character of the spectrum. As we will see in the next

sections, it is necessary to introduce an iterative protocol to sample from the continuous set of low-energy states of the KLM. Indeed, finding an adequate set of classical spin states to feed the ML process and find an adequate low-energy effective model is the most challenging part of the protocol because a big advantage of the classical limit is that the cost function is a convex function globally (i.e. it has no extra local minima), which vastly reduces the cost of the optimization procedure after each iteration.

From a pure mathematical standpoint, the challenge is then to generate a weighted set of low and high-energy classical spin states in order to learn what is the low-energy spin model $\tilde{\mathcal{H}}$ that better reproduces the low-energy spectrum of the original high-energy model $\mathcal{H}$. Finding a good balance between the relative weights assigned to low and high-energy states of the original high-energy model $\mathcal{H}$ is critical. Another challenge is to find an algorithm that converges relatively fast to the final version of $\tilde{\mathcal{H}}$, and to implement another optimization algorithm based on L1 regularization[31] that minimizes the number of interactions to obtain the *minimal* effective model that still reproduces the zero-temperature phase diagram of the high-energy model. As we will see in the following sections, the algorithm that we are proposing meets both challenges. The main limiting factor for the efficiency of the algorithm is the time associated with the generation of the initial training data set.

The discovery of magnetic SkX in chiral magnets, such as MnSi, $Fe_{1−x}Co_xSi$, FeGe, and Cu_2OSeO_3[32–36] spawned efforts for identifying stabilization mechanisms of SkX in different classes of materials. These efforts are revealing that new stabilization mechanisms are typically accompanied by novel physical properties. For instance, while the vector chirality is fixed in the magnetic skyrmions of chiral magnets such as the above-mentioned B20 compounds, it is a degree of freedom in the SkX of centrosymmetric materials such as $BaFe_{12−x−0.05}Sc_xMg_{0.05}O_{19}$, $La_{2−2x}Sr_{1+2x}Mn_2O_7$, Gd_2PdSi_3, and $Gd_3Ru_4Al_{12}$[24–28,37,38]. In the former case, the underlying spiral structure emerges from the competition between ferromagnetic exchange and the Dzyaloshinskii-Moriya (DM) interaction[39,40]. In contrast, the spiral ordering of centrosymmetric materials arises from frustration, i.e., from the competition between different exchange or dipolar interactions[11,41–44].

Most of the known magnetic SkXs have been reported for metallic materials, where the interplay between magnetic moments and conduction electrons leads to response functions that are of both fundamental and applied interests, such as the well-known topological Hall effect (THE)[18,19,21,22] and the current-induced skyrmion motion[45–48]. The THE is a direct consequence of the Berry curvature acquired by the reconstructed electronic bands. In the adiabatic limit, the momentum space Berry curvature is controlled by a real space Berry curvature, that is proportional to the skyrmion density in the absence of spin-orbit interaction[49]: each skyrmion produces an effective flux equal to the flux quantum Φ_0. Consequently, Hall conductivities comparable to the quantized value (e^2/h) can in principle be achieved if the ordering wave vector of the SkX is comparable to the Fermi wave vector k_F. This condition can be naturally fulfilled in f-electron systems where the interaction between magnetic moments is mediated by conduction electrons[29]. Indeed, field-induced SkXs with large topological Hall effect has been recently observed in the rare earth-based centrosymmetric materials Gd_2PdSi_3 and $Gd_3Ru_4Al_{12}$[24–28], which are in principle described by a KLM including multiple bands of conduction electrons coupled to localized f-magnetic moments. The combination of state-of-the-art band structure calculations with standard degenerate perturbation theory seems to be a promising route for deriving these KLMs from first principle calculations[50]. Consequently, it is important to develop efficient methodologies for

computing the quantum ($T = 0$) phase diagram of this class of models.

To understand the origin of the field-induced SkX phases, we will consider a simple triangular KLM with *classical* local magnetic moments:

$$
\mathcal{H} = -t \sum_{\langle \boldsymbol{r}, \boldsymbol{r}' \rangle, \sigma} \left(c_{\boldsymbol{r}\sigma}^{\dagger} c_{\boldsymbol{r}'\sigma} + h.c. \right) + J \sum_{\boldsymbol{r}, \alpha\beta} c_{\boldsymbol{r}\alpha}^{\dagger} \boldsymbol{\sigma}_{\alpha\beta} c_{\boldsymbol{r}\beta} \cdot \boldsymbol{S}_{\boldsymbol{r}}
$$
$$
- h \sum_{\boldsymbol{r}} S_{\boldsymbol{r}}^{z} + D \sum_{\boldsymbol{r}} \left(S_{\boldsymbol{r}}^{z} \right)^{2}, \tag{1}
$$

where the first term corresponds to a tight-binding model with the nearest-neighbor hopping amplitude t ($\langle \boldsymbol{r}, \boldsymbol{r}' \rangle$ denotes bonds of nearest-neighbor sites). The operator $c_{\boldsymbol{r}\sigma}^{\dagger}$ ($c_{\boldsymbol{r}\sigma}$) creates (annihilates) an itinerant electron with spin σ on site $\boldsymbol{r}$. J is the Kondo exchange interaction between the local magnetic moments $\boldsymbol{S}_{\boldsymbol{r}}$ and the conduction electrons ($\boldsymbol{\sigma}$ is the vector of the Pauli matrices), and the localized moments are described by the normalized classical vector field $\boldsymbol{S}_{\boldsymbol{r}}$ ($|\boldsymbol{S}_{\boldsymbol{r}}| = S$). (In this work we simply set $S = 1$.) The last two terms represent a Zeeman coupling to an external field H ($h = g\mu_B H$) and an easy-axis single-ion anisotropy ($D < 0$).

The coupling of localized spins with itinerant electrons leads to effective spin-spin interactions which can potentially stabilize multiple-$\boldsymbol{Q}$ magnetic orderings. Modeling the system with the "high-energy" KLM is numerically challenging because of a combination of reasons. The numerical cost of computing the energy for a lattice of N sites is of the order of $\mathcal{O}(N^3)$. While this cost can in principle be reduced to a linear function of N by implementing approximated numerical schemes, such as the Kernel Polynomial Method (KPM)[5,51,52], finite-size effects remain relevant in general even for very large lattices[1]. This important limitation was overcome by a variational calculation for the case of periodic structures with relatively small periods, which are fixed by carefully tuning the band-filling fraction. While these approaches can be used to obtain zero-temperature phases diagrams, like the ones that we will discuss in Sec. II C, they cannot be used to study dynamical response functions or the finite-temperature phase diagram. Moreover, it is difficult to extract stabilization mechanisms of multi-$\boldsymbol{Q}$ orderings from numerical solutions of the KLM. Since these orderings can be determined by n-spin interactions with $n \geq 4$, the derivation of a low-energy model beyond the RKKY level is crucial to understanding different aspects of the competition between single and multi-$\boldsymbol{Q}$ orderings.

In the weak-coupling limit, $J\eta(\epsilon_F) \ll 1$, where $\eta(\epsilon_F)$ is the density of states at Fermi energy ϵ_F, the KLM can be approximated by an effective RKKY model[6–8]:

$$
\mathcal{H}_{\mathrm{RKKY}} = \sum_{\boldsymbol{k}} \tilde{\mathcal{J}}(\boldsymbol{k}) \boldsymbol{S}_{\boldsymbol{k}} \cdot \boldsymbol{S}_{-\boldsymbol{k}}, \tag{2}
$$

where

$$
\boldsymbol{S}_{\boldsymbol{k}} = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{r}} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{S}_{\boldsymbol{r}} \tag{3}
$$

is the Fourier transform of the spins, and N is the total number of lattice sites. In momentum space, the effective RKKY interaction is proportional to the magnetic susceptibility $\chi_{\boldsymbol{k}}$ of the conduction electrons:

$$
\tilde{\mathcal{J}}(\boldsymbol{k}) = -J^2 \chi_{\boldsymbol{k}}, \tag{4}
$$

where

$$
\chi_{\boldsymbol{k}} = -\int \frac{d\boldsymbol{q}}{\mathcal{A}_B} \frac{f(\epsilon_{\boldsymbol{q}+\boldsymbol{k}}) - f(\epsilon_{\boldsymbol{q}})}{\epsilon_{\boldsymbol{q}+\boldsymbol{k}} - \epsilon_{\boldsymbol{q}}}. \tag{5}
$$

Here $f(\epsilon)$ is the Fermi distribution function, $\mathcal{A}_B$ is the area of the first Brillouin zone, and $\epsilon_{\boldsymbol{k}}$ is the bare dispersion of the electrons. Since $\boldsymbol{S}_{\boldsymbol{k}} = \boldsymbol{S}_{-\boldsymbol{k}}^{*}$ ($\boldsymbol{S}_{\boldsymbol{r}}$ is a real vector field) implies that $\boldsymbol{S}_{\boldsymbol{k}} \cdot \boldsymbol{S}_{-\boldsymbol{k}} \geq 0$ and $\sum_{\boldsymbol{k}} \boldsymbol{S}_{\boldsymbol{k}} \cdot \boldsymbol{S}_{-\boldsymbol{k}} = NS^2$, the RKKY interaction favors a helical spin ordering with an ordering wave vector that minimizes $\tilde{\mathcal{J}}(\boldsymbol{k})$

(maximizing $\chi_{\boldsymbol{k}}$). If a wave vector $\boldsymbol{Q}$ that optimizes the RKKY model (2) is not invariant under the symmetry group $\mathcal{G}$ of transformations that leave the KLM invariant, there are other symmetry-related wave vectors $\mathcal{G}\boldsymbol{Q}$ that also minimize the energy. These wave vectors will be denoted with the index v: $\{\boldsymbol{Q}_v\}$. The spatial inversion operation is not included in the symmetry group because, as we mentioned above, the $\boldsymbol{Q}_v$ and $-\boldsymbol{Q}_v$ components of the vector field are not independent ($\boldsymbol{S}_{\boldsymbol{Q}_v} = \boldsymbol{S}_{-\boldsymbol{Q}_v}^{*}$). Thus, there is a degeneracy for the helical orderings of symmetry-related wave vectors. The real space version of the RKKY model is given by Fourier transforming Eq. (2):

$$
\mathcal{H}_{\mathrm{RKKY}} = \frac{1}{2} \sum_{\boldsymbol{r} \neq \boldsymbol{r}'} \mathcal{J}(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{S}_{\boldsymbol{r}} \cdot \boldsymbol{S}_{\boldsymbol{r}'}. \tag{6}
$$

Away from the weak-coupling limit, four-spin and higher-order interactions are naturally generated from the KLM, which are ignored in the RKKY model. Since the RKKY part alone is not sufficient to stabilize a field-induced SkX phase, four-spin interactions play a crucial role in the absence of single-ion anisotropy. The most general effective Hamiltonian including two and four-spin interactions is

$$
\mathcal{H}_{\mathrm{eff}} = \sum_{\boldsymbol{k}} \tilde{\mathcal{J}}(\boldsymbol{k}) \boldsymbol{S}_{\boldsymbol{k}} \cdot \boldsymbol{S}_{\overline{\boldsymbol{k}}}
$$
$$
+ \sum_{\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3} \frac{g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)}{N} (\boldsymbol{S}_{-\boldsymbol{K}} \cdot \boldsymbol{S}_{\boldsymbol{k}_1})(\boldsymbol{S}_{\boldsymbol{k}_2} \cdot \boldsymbol{S}_{\boldsymbol{k}_3}), \tag{7}
$$

where $\boldsymbol{K} \equiv \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$. The four-spin interaction term helps to lift the massive ground state degeneracy of the RKKY model. Suppose the term $(\boldsymbol{S}_{\boldsymbol{Q}_i} \cdot \boldsymbol{S}_{-\boldsymbol{Q}_i})(\boldsymbol{S}_{\boldsymbol{Q}_i} \cdot \boldsymbol{S}_{-\boldsymbol{Q}_i})$ in the Hamiltonian has a large positive coefficient $g(-\boldsymbol{Q}_i, \boldsymbol{Q}_i, -\boldsymbol{Q}_i)$, then the single-$\boldsymbol{Q}_i$ ordering will be heavily penalized. On the other hand, a large negative coefficient $g(\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3)$ favors three different $\boldsymbol{Q}$ components ($\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3$) being finite in the presence of a finite uniform magnetization $\boldsymbol{S}_0$. In other words, this term favors a triple-$\boldsymbol{Q}$ ordering, such as SkXs, if a finite magnetization is induced by an external magnetic field. Thus, the function $g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ plays a critical role in the identification of the ground state. As we demonstrate in this paper, ML turns out to be a valuable tool in extracting this important piece of information.

Since $g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ is an unknown function of 3 continuous multidimensional variables, we need an efficient scheme to reduce the number of model parameters. To obtain a valid effective low-energy theory, the scheme should provide a good approximation near the most relevant wave vectors ($\boldsymbol{k} \simeq \boldsymbol{Q}_v$ and $\boldsymbol{k} \simeq \boldsymbol{0}$). To achieve this goal, we will first derive a real space version of the low-energy effective Hamiltonian, where all symmetry-allowed interactions are included up to a certain distance. After these real space interactions are determined with ML, the expressions of $\tilde{\mathcal{J}}(\boldsymbol{k})$ and $g(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ can be obtained by a simple Fourier transform of the effective real space interactions.

To avert the risk of introducing manual selection bias in the two- and four-spin interactions, we sought to implement a generalized parameter for their selection. Consequently, we decided to retain all interactions within a specific distance range. For the RKKY contribution (6), we cutoff the interactions beyond $|\boldsymbol{r} - \boldsymbol{r}'| = 2\sqrt{3}a$, which leads to 6 inequivalent exchange parameters $\{\mathcal{J}_1, \dots, \mathcal{J}_6\}$. Similarly, the four-spin contribution can be expressed as:

$$
\mathcal{H}_{4-\mathrm{spin}} = \sum_{\langle \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4 \rangle} f(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4)(\boldsymbol{S}_{\boldsymbol{r}_1} \cdot \boldsymbol{S}_{\boldsymbol{r}_2})(\boldsymbol{S}_{\boldsymbol{r}_3} \cdot \boldsymbol{S}_{\boldsymbol{r}_4}), \tag{8}
$$

where the notation $\langle \boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4 \rangle$ indicates that each set of four sites ($\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4$) is counted only once. Here we restrict $|\boldsymbol{r}_i - \boldsymbol{r}_j| \leq \sqrt{3}a$, which leads to 7 inequivalent exchange parameters $\{f_1, \dots, f_7\}$ (see Fig. 1). The primary source of computational cost arises from the calculations pertaining to the four-spin interactions. To minimize computational cost, we selected the
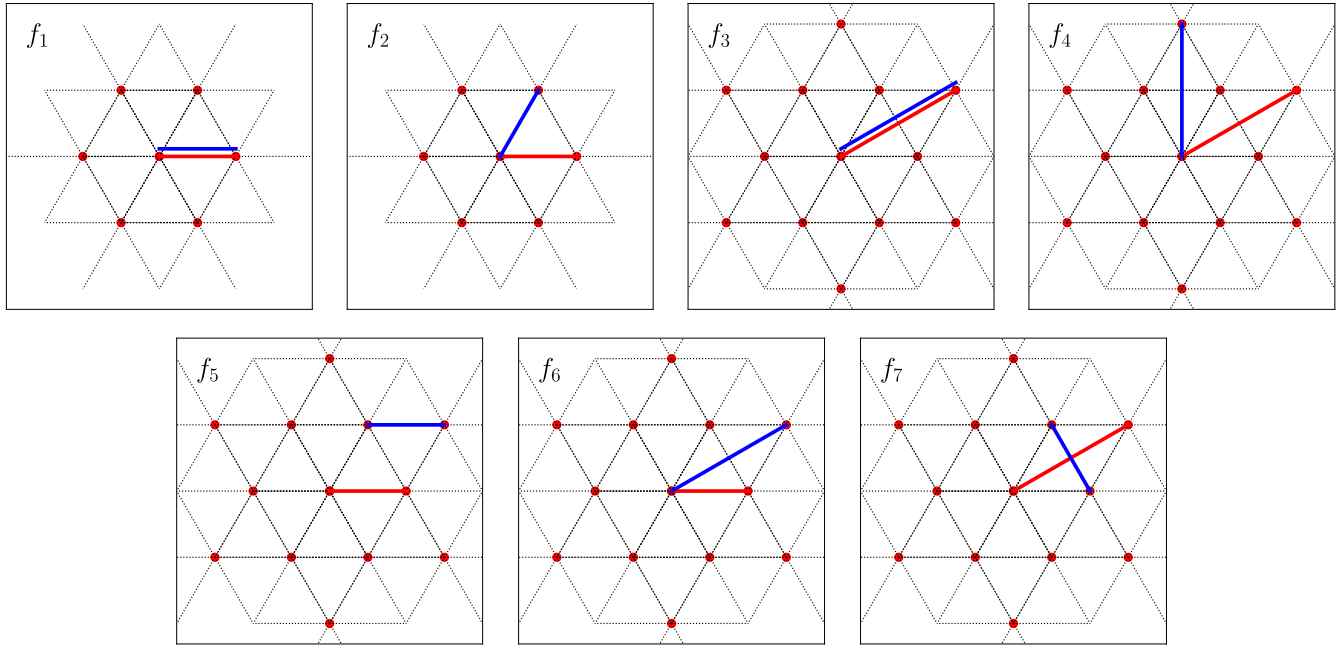
Fig. 1 Four-spin interactions with cutoff $|r_i - r_j| \leq \sqrt{3}a$ on a triangular lattice. The red and blue lines represent $(S_{r_1} \cdot S_{r_2})$ and $(S_{r_3} \cdot S_{r_4})$ terms, respectively.

distance range to incorporate only the minimum number of four-spin interactions necessary for producing an adequate phase diagram. For the RKKY interaction, we incorporated all interactions within the length scale of the magnetic unit cell discussed in section II A.

By implementing the ML algorithm described in section II B, we compute these real space parameters and subsequently evaluate $\tilde{\mathcal{J}}(\mathbf{k})$ and $g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$. The aim is to gain a comprehensive understanding of the low-energy physics. The significance of the coefficients associated with ordering wave vectors and their higher harmonics is of critical importance, as these play a dominant role in shaping the low-energy states. Provided the ground state weight is not excessively concentrated at the higher harmonics, we could in principle study the system by just considering the terms corresponding to $\mathbf{k}$ values in the set $(\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{0})$. The energy resulting from four-spin interactions is confined to a specific subset of possible combinations of $g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ by restricting the available wave vectors to the set $(\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{0})$ and it can be written as:

identifying which interactions between modes play a role in stabilizing multi-$\mathbf{Q}$ spiral orderings. Additionally, this minimal model can be used to differentiate between four-spin contributions involving S_0, that become relevant for finite magnetic field, from those that are present in absence of uniform magnetization S_0. By isolating these factors, we can further examine the impact of magnetic field on the stability of multi-$\mathbf{Q}$ states. As we shall observe in subsequent sections, the insights derived from this minimal model are illuminating and provide us with a wealth of information about the system.

RESULTS

Data from high energy model - KLM calculations

When the Kondo exchange interaction is comparable to the nearest-neighbor hopping ($J \lesssim t$), the effective spin-spin interactions are orders of magnitude smaller than the bare Hamiltonian parameters. This leads to large finite-size effects that can not be ignored—even for very large system sizes, the relative stability of

$$
\begin{aligned}
NE_{4-\text{spin}} = \; & \tilde{g}_0(S_0 \cdot S_0)^2 + \tilde{g}_1 \sum_v (S_0 \cdot S_0)(S_{Q_v} \cdot S_{-Q_v}) + \tilde{g}_2 \sum_v (S_0 \cdot S_{Q_v})(S_0 \cdot S_{-Q_v}) \\
& + \tilde{g}_3[(S_0 \cdot S_{Q_1})(S_{Q_2} \cdot S_{Q_3}) + (S_0 \cdot S_{-Q_1})(S_{-Q_2} \cdot S_{-Q_3})] \\
& + \tilde{g}_3[(S_0 \cdot S_{Q_2})(S_{Q_1} \cdot S_{Q_3}) + (S_0 \cdot S_{-Q_2})(S_{-Q_1} \cdot S_{-Q_3})] \\
& + \tilde{g}_3[(S_0 \cdot S_{Q_3})(S_{Q_1} \cdot S_{Q_2}) + (S_0 \cdot S_{-Q_3})(S_{-Q_1} \cdot S_{-Q_2})] \\
& + \tilde{g}_4[(S_{Q_1} \cdot S_{-Q_2})(S_{-Q_1} \cdot S_{Q_2}) + (S_{Q_2} \cdot S_{-Q_3})(S_{-Q_2} \cdot S_{Q_3}) + (S_{Q_3} \cdot S_{-Q_1})(S_{-Q_3} \cdot S_{Q_1})] \\
& + \tilde{g}_5[(S_{Q_1} \cdot S_{Q_2})(S_{-Q_1} \cdot S_{-Q_2}) + (S_{Q_2} \cdot S_{Q_3})(S_{-Q_2} \cdot S_{-Q_3}) + (S_{Q_3} \cdot S_{Q_1})(S_{-Q_3} \cdot S_{-Q_1})] \\
& + \tilde{g}_6 \sum_v (S_{Q_v} \cdot S_{Q_v})(S_{-Q_v} \cdot S_{-Q_v}) + \tilde{g}_7 \sum_v (S_{Q_v} \cdot S_{-Q_v})(S_{Q_v} \cdot S_{-Q_v}) \\
& + \tilde{g}_8[(S_{Q_1} \cdot S_{-Q_1})(S_{Q_2} \cdot S_{-Q_2}) + (S_{Q_2} \cdot S_{-Q_2})(S_{Q_3} \cdot S_{-Q_3}) + (S_{Q_3} \cdot S_{-Q_3})(S_{Q_1} \cdot S_{-Q_1})],
\end{aligned}
\tag{9}
$$

where the expressions of $\{\tilde{g}_i\}$ are given in the Supplementary Information. By utilizing this absolute minimal model in momentum space, we can gain a deeper understanding of the system by

two competing states can switch in comparison to the thermodynamic limit. To avoid these undesirable effects, it is imperative to work in the thermodynamic limit. Thus, we implemented a

variational approach on a fixed magnetic unit cell[1]. In the following, we assume that the magnetic unit cell is spanned by the basis $\{L\boldsymbol{a}_1, L\boldsymbol{a}_2\}$ where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the primitive vectors of the lattice. We label different sites in the magnetic unit cell by $\boldsymbol{R}$ and the different magnetic unitcells by $\tilde{\boldsymbol{r}}$. Coordinates of each site can be expressed as $\boldsymbol{r} = \tilde{\boldsymbol{r}} + \boldsymbol{R}$. By using the translation symmetry of commensurate states, we can write the Fourier transform as

$$c_{\tilde{\boldsymbol{r}}, \boldsymbol{R}, \sigma} = \sqrt{\frac{L^2}{N}} \sum_{\tilde{\boldsymbol{k}}} e^{i\tilde{\boldsymbol{k}} \cdot \tilde{\boldsymbol{r}}} c_{\tilde{\boldsymbol{k}}, \boldsymbol{R}, \sigma}, \tag{10}$$

where $\tilde{\boldsymbol{k}}$ labels are the allowed momenta in the reduced Brillouin zone $\mathcal{B}_r$ and N is the total number of sites. The KLM Hamiltonian is block-diagonal in momentum space, $\mathcal{H} = \sum_{\tilde{\boldsymbol{k}}} \mathcal{H}_{\tilde{\boldsymbol{k}}}$, where

$$\mathcal{H}_{\tilde{\boldsymbol{k}}} = \sum_{\boldsymbol{R}} \left[-t \sum_{\eta} \sum_{\sigma} c^{\dagger}_{\tilde{\boldsymbol{k}}, \boldsymbol{R}, \sigma} c_{\tilde{\boldsymbol{k}}, \boldsymbol{R}+\boldsymbol{r}_{\eta}, \sigma} \right.$$
$$\left. + J \sum_{\alpha\beta} c^{\dagger}_{\tilde{\boldsymbol{k}}, \boldsymbol{R}, \alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{\tilde{\boldsymbol{k}}, \boldsymbol{R}, \beta} \cdot \boldsymbol{S}_{\boldsymbol{R}} - h S^z_{\boldsymbol{R}} + D\left(S^z_{\boldsymbol{R}}\right)^2 \right]. \tag{11}$$

Here, $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_6\}$ denote the relative position of the nearest-neighbor sites. Note that Bloch's theorem is implied here:

$$c_{\tilde{\boldsymbol{k}}, \boldsymbol{R}+\tilde{\boldsymbol{r}}, \sigma} \equiv e^{i\boldsymbol{k} \cdot \tilde{\boldsymbol{r}}} c_{\tilde{\boldsymbol{k}}, \boldsymbol{R}, \sigma}. \tag{12}$$

A $2L^2 \times 2L^2$ block matrix is diagonalized for each $\tilde{\boldsymbol{k}}$ to obtain single-particle eigenstates. The $T=0$ energy density is then computed as:

$$\frac{E}{N} = \frac{1}{N} \sum_{\tilde{\boldsymbol{k}}} \sum_{n=1}^{2L^2} \Theta\left(\mu - \epsilon_{n,\tilde{\boldsymbol{k}}}\right) \epsilon_{n,\tilde{\boldsymbol{k}}} + \frac{1}{L^2} \sum_{\boldsymbol{R}} \left[-h S^z_{\boldsymbol{R}} + D\left(S^z_{\boldsymbol{R}}\right)^2 \right], \tag{13}$$

where Θ is the step function which selects the energy levels below the chemical potential μ, and $\epsilon_{n,\tilde{\boldsymbol{k}}}$ represent the eigenvalues of the block matrix. In order to accurately identify the ground state, we take the thermodynamic limit by converting the discrete sum $\frac{1}{N}\sum_{\tilde{\boldsymbol{k}}}$ into an integral:

$$\frac{E}{N} = \frac{1}{L^2} \int_{\mathcal{B}_r} \frac{d\tilde{\boldsymbol{k}}}{\mathcal{A}_{\mathcal{B}_r}} \sum_{n=1}^{2L^2} \Theta\left(\mu - \epsilon_{n,\tilde{\boldsymbol{k}}}\right) \epsilon_{n,\tilde{\boldsymbol{k}}} + \frac{1}{L^2} \sum_{\boldsymbol{R}} \left[-h S^z_{\boldsymbol{R}} + D\left(S^z_{\boldsymbol{R}}\right)^2 \right], \tag{14}$$

where $\mathcal{A}_{\mathcal{B}_r}$ represents the area of the reduced Brillouin zone.

In the minimization process, the chemical potential μ was determined self consistently from the filling fraction:

$$n_c = \frac{1}{2L^2} \int_{\mathcal{B}_r} \frac{d\tilde{\boldsymbol{k}}}{\mathcal{A}_{\mathcal{B}_r}} \sum_{n=1}^{2L^2} \Theta\left(\mu - \epsilon_{n,\tilde{\boldsymbol{k}}}\right). \tag{15}$$

The various phases of the phase diagram were then obtained by minimizing the ground state energy over all the possible spin structures for the fixed magnetic unit cell. For each set of parameters, we typically performed at least 20 independent minimization runs with different initial spin configurations to avoid metastable local minima[1].

The immediate challenge is to find the right size of the magnetic unit cell, which is a requirement to validate the variational scheme. In the RKKY limit ($J/t \to 0$), ordering wave vectors $\boldsymbol{Q}_\nu$ ($\nu = 1, 2, 3$) are located along the high-symmetry Γ-M directions with magnitude $|\boldsymbol{Q}_\nu| = 2k_F$[29]. However, a finite value of J/t leads to a shift of the wave vectors[53]. To find the correct values, we simulated the KLM on a 96×96 lattice using stochastic Landau-Lifshitz (SLL) dynamics. We employed the KPM to obtain the free energy and local forces[5,51,52]. Even though a typical finite lattice is not adequate to accurately compute the relative energies of competing states, it is sufficient to determine the size of the magnetic unit cell. KPM-SLL represents a completely unbiased approach to find the period of the optimal ground state ordering. However, since this method is not effective in handling zero temperature, here we introduced a very small temperature T. For $J/t = 0.5$, $T = 10^{-5}J^2/t$, $D = 0$, and $h = 0$, we integrated the
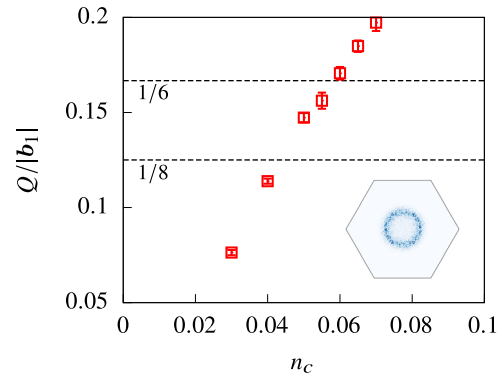

Fig. 2 Ordering wave number as a function of the filling fraction for the triangular KLM on a 96×96 lattice with $h = D = 0$, $J/t = 0.5$, and $T = 10^{-5}J^2/t$, obtained from KPM-SLL simulation. The error bars were estimated by 6 independent runs with different random number seeds. The inset shows a snapshot of the static spin structure factor $\mathcal{S}(\boldsymbol{q})$ in the first Brillouin zone with $n_c = 0.06$. This figure is reproduced from ref. [1].

dimensionless SLL dynamics with a unit damping parameter using the Heun-projected scheme for a total of 45000 steps of duration $\Delta\tau = \frac{0.5}{(J^2/t)}$. We used the gradient-based probing method with $S = 256$ colors and $M = 1000$ for the order of the Chebyshev polynomial expansion[52]. We discarded the first 30,000 steps for equilibration and used the rest 15000 steps for measurements. Final results were averaged over 6 independent runs to estimate the error bars. To get a magnetic unit cell of size 6×6, the results yielded a filling fraction $n_c \approx 0.0586$ (see Fig. 2).

Machine learning algorithm
The first step to extract the spin Hamiltonian from the KLM (or any other high-energy model) is generating an initial high-energy data set on a grid of magnetic field strength (h) and easy-axis anisotropy (D). This grid need not be very dense. In particular, our grid for the ML algorithm had 47 points, while the grid for producing the final phase diagram using the low-energy model used more than 3000 points. For each value of h and D on the grid, we generate a random spin configuration. Starting from this random configuration, we find the local minimum using a gradient-based method of the original KLM (calculation details are given in section II A). All the spin configurations generated in the minimization process are stored along with their corresponding energies and values of h and D in a data set[1]. The next step is to set up a trial Hamiltonian:

$$\mathcal{H} = \sum_{j=1}^{M} c_j \mathcal{H}_j \qquad (j = 1, \ldots, M). \tag{16}$$

where $\mathcal{H}_j$ consists of a constant term, real space RKKY exchange interactions, and four-spin interactions (described in the Introduction). The effective low-energy model is expected to reproduce the low-energy spectrum of the original KLM. Correspondingly, to formulate the search for this effective model as an optimization problem, we introduce the cost function:

$$\text{Cost}(\{c_j\}) = \frac{1}{\tilde{N}} \sum_{D,h} \sum_{i=0}^{\tilde{N}-1} \delta_i^{D,h}\left(E_i^{D,h} - \tilde{E}_i^{D,h}\right)^2, \tag{17}$$

$$\delta_i^{D,h} = \frac{1}{\sqrt{E_i^{D,h} - E_{\min}^{D,h} + \epsilon}}, \tag{18}$$

where $\tilde{N}$ is the total number of states in the data set, $E_i^{D,h}$ is the energy obtained by the high-energy model for the i-th spin configuration for a particular point on grid (D, h), $\tilde{E}_i^{D,h}$ is the

predicted energy from the trial Hamiltonian, ϵ is a very small hyperparameter ensuring that the weight function $\delta_i^{D,h}$ does not diverge and $E_{\min}^{D,h}$ is the lowest energy in the training data set for a particular value of (D, h).

The data set comprises of considerable number of states stretching across the full energy spectrum. At first glance, it may seem reasonable to only consider low-energy states in the training phase of the ML model and disregard high-energy states. However, this strategy is problematic for several reasons. Firstly, the low-energy spectrum changes in response to variations in magnetic field and anisotropy. The ground state and low-energy states are contingent upon the strength of magnetic field and easy-axis anisotropy, meaning that low energy states for one set of parameters (D, h) may not coincide with those for a different set of parameters. Furthermore, in absence of high-energy states in the training data set, the training can lead to a model that assigns high-energy states with lower energies even if the energies of the low-energy states are accurately predicted. The accuracy of the energy prediction for the low-energy states must surpass the energy differences between competing states, and high-energy states must maintain a clear separation in the energy spectrum. Given these factors, we opted to utilize all states in the training process. To prioritize the lower energy segment while accommodating the high-energy states, we implemented a weight factor that is guided by the energies of these states. The lower the energy, the greater the weight, and the higher the energy, the lower the weight. The weight factor $\delta_i^{D,h}$ not only determines the relative weight of low and high-energy states in the original model, but also keeps track of the magnetic field and anisotropy for which the state was generated. To optimize the weight factor, we implemented our algorithm with three different choices:

$$\delta_i^{D,h} = \left\{ \frac{1}{\sqrt{E_i^{D,h} - E_{\min}^{D,h} + \epsilon}}, 1, \frac{1}{|E_i^{D,h} - E_{\min}^{D,h} + \epsilon|} \right\}. \quad (19)$$

For $\delta_i^{D,h} = 1$, the precision in the energy prediction for low-energy states was comparable to the energy scale of the low-energy excitations and hence, it was not adequate to establish the appropriate energy order of the competing low-energy states. For the choice

$$\delta_i^{D,h} = \frac{1}{|E_i^{D,h} - E_{\min}^{D,h} + \epsilon|},$$

there was an excessive weight on the low-energy states and, as a consequence, the trained model had exceedingly depleted precision for other states. The choice

$$\delta_i^{D,h} = \frac{1}{\sqrt{E_i^{D,h} - E_{\min}^{D,h} + \epsilon}}$$

provided an appropriate balance between the accuracy for the low-energy and high-energy states. After setting up the cost function and initializing the coupling constants, we perform the following three steps iteratively:

1. Update parameters $\{c_j\}$ iteratively using the gradient descent method: $c_j \rightarrow c_j - a \frac{\partial \text{Cost}(\{c_j\})}{\partial c_j}$ until cost function reaches a minimum (a is the learning rate). Keeping the learning rate a small, the gradient can be analytically calculated.

$$c_j \rightarrow c_j - a \sum_{i=0}^{\bar{N}-1} \frac{\partial \text{Cost}(\{c_j\})}{\partial E_i(\{c_j\})} \frac{\partial E_i(\{c_j\})}{\partial c_j} \quad (20)$$

$$\Rightarrow c_j \rightarrow c_j - a \sum_{i=0}^{\bar{N}-1} \frac{\partial \text{Cost}(\{c_j\})}{\partial E_i(\{c_j\})} \langle \psi_i | H_j | \psi_i \rangle \quad (21)$$

One of the advantages of the search for a low-energy *classical model* is that multivariate weighted linear regression can be used to find the optimal model parameters. In contrast, a non-linear regression is required for quantum mechanical low-energy models because the eigenstates change with the model parameters[31]. In the classical limit, each direct product of coherent spin states is an eigenstate, a feature not shared by quantum models. Consequently, every eigenstate of a low-energy model remains an eigenstate of the high-energy model, even when the low-energy model parameters are altered.

2. Produce the zero-temperature phase diagram of the "ML model" (with the parameters that minimize the cost function) by energy minimization via a gradient descent method. For local minimization algorithms, the converged results are usually metastable local minima, i.e., different initial spin configurations can lead to different final states. Correspondingly, for each h and D, we typically performed 60 independent runs with different random initial spin configurations to find the global minima.

3. For the global minimum energy states of the ML model calculate the KLM energies and add them to the training data set. Then update the weights $\delta_i^{D,h}$ for the new variational space.

The iterative process stops when the required precision is reached for the minimized states, i.e., when the "ML model" produces the same lowest energy states as in the previous iteration up to a certain decimal place. When the model converges, the error for the low-energy states, $|E - \tilde{E}|$, is lower than the energy difference $|E_A^{0,0} - E_B^{0,0}|$ between competing states A and B near the phase boundaries. This three-step iterative approach addresses the challenge of choosing states from the continuous spectrum by avoiding arbitrary determinations of which states are important. The models that are produced at each intermediate step of the iterative process are used to generate new low-energy states, which are added to the training data set. The intermediate models continually evolve, progressively enriching the training data set with relevant states until the model is converged.

To obtain a minimal model that includes the minimum number of interaction terms, we apply an $L1$ norm regularization of the cost function[31,54]. The details of this procedure are described in Sec. IV. Once the $L1$ regularization eliminates the irrelevant terms from the Hamiltonian, we optimize the model again without the regularization term. In this way, the powerful technique of $L1$ regularization outputs a minimal, yet accurate, low-energy model Hamiltonian.

Phase diagram of the ML model

Our algorithm took four iterations to converge for $J/t = 0.5$. Table 1 shows the real space Hamiltonian parameters for each iteration and the final minimal model. Note that all the exchange parameters are about three orders of magnitude smaller than the bare energy scales (t, J) in the original model, which again explains the difficulty of directly solving the KLM numerically. The first neighbor RKKY interaction ($\mathcal{J}_1$) is an order of magnitude larger than the next biggest interaction, and the magnitude of four-spin interactions is comparable to the other RKKY interactions, confirming that the KLM parameters are in the intermediate-coupling regime.

To obtain the phase diagram of the ML model, we generated random spin configurations for each h and D and used the gradient descent method to minimize the energy. To find the global minimum for each point, we performed multiple independent runs with different initial random spin configurations. The phase diagrams obtained from variational calculations of the original KLM and the ML models are given in Fig. 3. These diagrams include seven different phases: the vertical spiral (VS), vertical spiral with in-plane modulation along one direction (VS″), vertical spiral with in-plane modulation (VS′), 2**Q**-conical spiral (2**Q**-CS), 2**Q**-conical spiral

with unequal in-plane structure factor intensities($2Q$-CS′), SkX and the fully polarized (FP) phase, where the real space spin configurations of all the phases are given in refs. [1,29]. The ML models reproduce the phase boundaries with remarkable accuracy. Even the simplest minimal model in momentum space (Table 2) is able to predict the correct phase boundaries. We should note,

however, that this might not be always the case for the minimal model in momentum space if the ground state magnetic orderings have sizable weight in higher harmonics. For the phase transitions: VS″ to VS′ and VS′ to VS, the phase boundaries predicted by the ML models are qualitatively accurate but quantitatively they do not reproduce the KLM phase boundaries. This shortcoming of the low-energy ML model can be attributed to the extremely small energy differences between the three competing phases that are approximately equal to a single-Q vertical spiral. Magnitudes of the extra Fourier components, that differentiate between VS, VS′ and VS″ states are two orders of magnitude smaller than the magnitude of major Q-component. Consequently, the exceedingly small energy difference between these phases cannot be captured by the limited number of spin interaction terms included in the ML Hamiltonian.

The next question that we can ask is: what can we learn from the effective low-energy model derived with the ML protocol? To answer this question, we will consider the problem of understanding the stabilization mechanism of the SkX phase. As it has been shown in a recent work[29], a finite easy-axis anisotropy is required to stabilize the field-induced SkX phase of the RKKY model. Since the phase diagram of the KLM for $J/t = 0.5$ includes a field-induced SkX phase for $D = 0$, it is clear that the effective four-spin interactions are responsible for its stabilization in absence of single-ion anisotropy. It is then interesting to inquire about the nature of the four-spin interactions that stabilize the SkX phase. The low-energy effective spin model given in Eq. (9) includes two types of four-spin interactions: those involving the uniform spin component S_0 and those that only involve the finite-Q Fourier components S_{Q_ν} with $\nu = 1, 2, 3$. The field-induced SkX phase has been traditionally attributed to the $\tilde{g}_3$ interaction[55], which is only

Table 1. Real space Hamiltonian parameters for Eqs. (6) and (8) at each iteration (units: $10^{-3}t$).

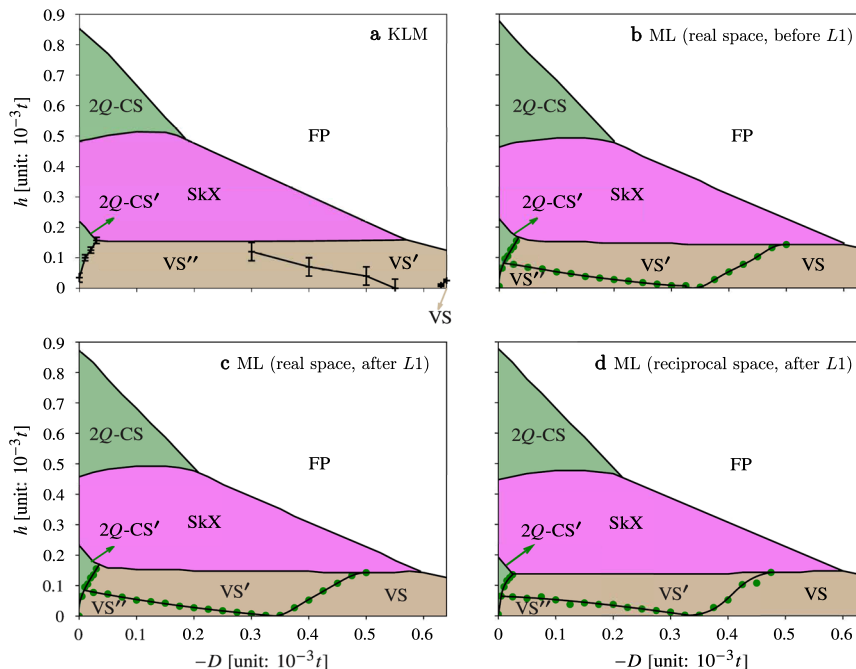|	Iteration 1	Iteration 2	Iteration 3	Iteration 4	After $L1$
$\mathcal{J}_1$	−2.35165	−2.64772	−2.64827	−2.64978	−2.68884
$\mathcal{J}_2$	−0.39479	−0.25397	−0.25489	−0.25327	−0.21033
$\mathcal{J}_3$	−0.07018	0.01546	0.01673	0.01721	/
$\mathcal{J}_4$	0.22517	0.25812	0.25730	0.25776	0.26126
$\mathcal{J}_5$	0.08376	0.15388	0.15375	0.15422	0.15312
$\mathcal{J}_6$	0.21353	0.07107	0.07125	0.07081	0.05770
f_1	−0.07113	0.01635	0.02298	0.01984	0.02927
f_2	−0.75221	−0.33654	−0.33809	−0.33496	−0.27487
f_3	−0.13638	−0.12970	−0.13573	−0.13209	−0.14971
f_4	−0.10768	−0.09602	−0.09395	−0.09633	−0.08659
f_5	0.28331	0.04871	0.04790	0.04722	/
f_6	0.12508	0.16483	0.16582	0.16523	0.18360
f_7	0.84930	0.31596	0.31372	0.31393	0.22350

Fig. 3 $T = 0$ phase diagrams of the KLM at $J/t = 0.5$ and $n_c = 0.0586$. **a** phase diagram obtained via high-energy model (KLM) [reproduced from Fig. 2b of ref. [1]. The error bars of phase boundaries at low field indicate the limited numerical accuracy due to quasi-degenerate states. **b** phase diagram of the real space spin model obtained via ML. **c** phase diagram corresponding to real space ML model after $L1$ regularization, and **d** phase diagram corresponding to the minimal momentum space model obtained from $L1$ regularized real space ML model. The phase boundaries at low field in **b**–**d** are denoted by the filled circles, where the corresponding lines are guides to the eye.

Table 2. Momentum space Hamiltonian parameters for Eq. (7) in units of $10^{-3}t$.

$\tilde{\mathcal{J}}(\mathbf{Q})$	$\tilde{\mathcal{J}}(\mathbf{0})$	$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{g}_4$	$\tilde{g}_5$	$\tilde{g}_6$	$\tilde{g}_7$	$\tilde{g}_8$	$\tilde{g}_0$
−6.245024	−6.497512	−0.020120	1.627376	−0.347720	1.592957	−2.084520	1.183781	−1.380299	−3.438664	0.343609

present in three-fold symmetric systems, such as the one under consideration, and that involves the uniform field-induced component S_0. The simple reason is that this contribution is finite only for spin configurations that have finite Fourier components S_{Q_1}, S_{Q_2} and S_{Q_3}. However, the negative sign of $\tilde{g}_3$ that we are obtaining from the ML model (see Table 2) indicates that this term actually increases the energy of the SkX phase relative to the competing single-Q and double-Q orderings. The prevalent interactions that assist in the stabilization of SkXs are $\tilde{g}_8$ and $\tilde{g}_5$ – They have the biggest magnitude in the minimal model and the large negative coefficients correspond to an attractive interaction between pairs of different modes that lowers the energy for triple-Q and double-Q orderings relative to the single-Q vertical spiral ordering. The magnitude of their negative contribution to the SkX energy is higher than that of double-Q spiral and thus they help stabilize the SkX phase. We note that the single-Q vertical spiral ordering is stable at low enough fields because it is the only phase that has zero magnetization, implying that the weight $||S_{Q_1}||^2$ is higher than the sum of the three weights $||S_{Q_1}||^2 + ||S_{Q_2}||^2 + ||S_{Q_3}||^2$ of the SkX and the double-Q phases. As we will discuss in Section III, this observation has important consequences for the stabilization of SkX phases in tetragonal materials[56,57].

Dynamics

In view of the robustness demonstrated by our ML-derived models in predicting the ground state physics of the system, it is only logical to inquire about their ability to predict low-energy dynamics. Given that the prediction of low-energy dynamics was not a primary objective during the development of the model, a successful demonstration of its proficiency in this regard would be surprising. Thus, to further check the validity of the ML models, in this section, we compare the underlying spin dynamics to the one obtained from the original KLM.

For concreteness, in this section, we choose a magnetic field $h = 0.01t$ that is higher than the saturation field, i.e., strong enough to fully polarize the local spins, and $D = 0$. At very low temperatures (linear regime), the semi-classical spin wave dispersion can be obtained by solving the Landau-Lifshitz (LL) equations of motion:

$$\frac{dS_i}{dt} = -S_i \times \frac{dE}{dS_i}, \tag{22}$$

where E is the internal energy of the system.

The dynamical spin structure factor can be obtained by Fourier transforming the spin configurations both in space and time:

$$\mathcal{S}^{ab}(k, \omega) \approx \frac{\omega}{T} \langle S_k^a(\omega) S_{-k}^b(-\omega) \rangle, \tag{23}$$

where the prefactor $\omega/T \gg 1$ accounts for the classical-quantum correspondence factor required to obtain the quantum mechanical result (linear spin wave theory) from the classical one[58], and

$$S_k(\omega) \equiv \frac{1}{\sqrt{T_S}} \int_0^{T_S} dt e^{i\omega t} \frac{1}{\sqrt{N}} \sum_i e^{-ik \cdot r_i} S_i(t). \tag{24}$$

The above Eqs. (22)–(24) can be applied both to the original KLM and to the effective ML models. The main difference is that the cost of computing the local forces $-dE/dS_i$ from the KLM is much higher in comparison with the effective low-energy spin models. In this work, the forces of the KLM are always computed with the KPM[5,51,52,59].

To obtain $\mathcal{S}(k, \omega)$ from the KLM, we initialized the spins from the fully polarized state on a 216×216 triangular lattice. We then applied the KPM-SLL method (same as in Sec. II A) to equilibrate the system at $T = 10^{-5}t$ with parameters $n_c = 0.0586$, $t = 1$, $J = 0.5$, $D = 0$, and $h = 0.01$. We used the Heun-projected scheme with a unit damping parameter and a total of 14000 steps of duration $\Delta\tau = 2.5/(J^2/t)$. In addition, we adopted $S = 324$ colors for the gradient-based probing. The order of the Chebyshev expansion was set to $M = 2000$.

The spin configuration at the last step was used to seed the LL dynamics (22), where the damping is set to zero. Once again, the
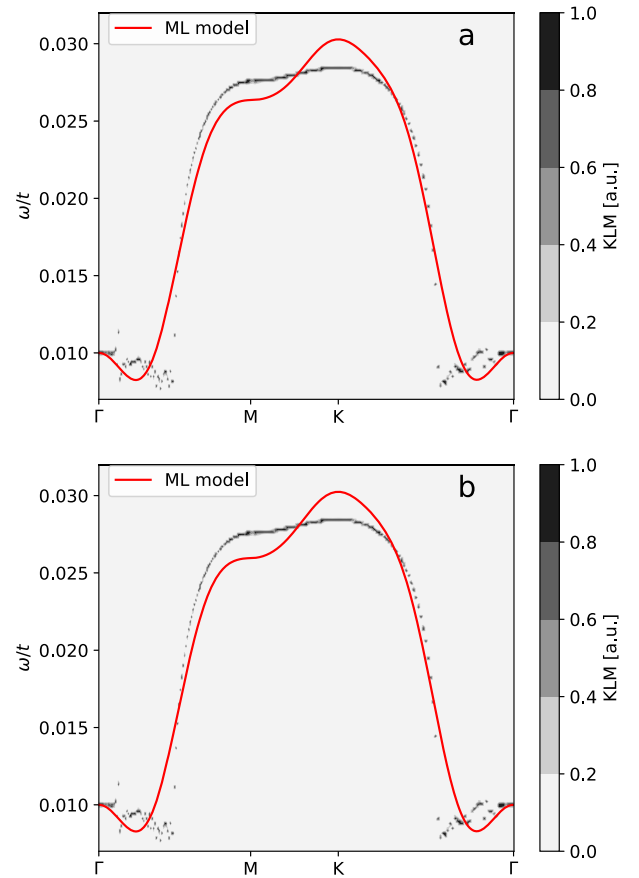

Fig. 4 **Magnon dispersion in the fully polarized phase.** **a** Comparison between linear spin wave results for the ML model before $L1$ regularization and $\mathcal{S}(k, \omega)$ obtained with the original KLM model. **b** Comparison between the ML model after $L1$ regularization and $\mathcal{S}(k, \omega)$ obtained with the original KLM model.

local forces were evaluated using the KPM with $M = 2000$ and $S = 324$. For convergence, we applied the Heun-projected scheme with a total of 40000 steps of size $\Delta\tau = 0.25/(J^2/t)$ [$T_S = 10^4/(J^2/t)$ in Eq. (24)]. Finally, we used 10 independent runs to compute the average of $\mathcal{S}(k, \omega)$, which is presented in Fig. 4.

For the ML models, we can calculate the magnon dispersion analytically in the fully polarized state by implementing the usual Holstein-Primakoff transformation (linear spin waves):

$$\omega_k = \Delta E + \sum_r t_r \cos(k \cdot r), \tag{25}$$

where

$$\Delta E = h + D(1 - 2S) - 6(\mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + 2\mathcal{J}_4 + \mathcal{J}_5 + \mathcal{J}_6)S$$
$$- 12(f_1 + f_3)S^3 - 24(f_2 + f_4 + 2f_6)S^3$$
$$- 12(2f_5 + f_7)S^3, \tag{26}$$

and t_r depends only on $|r|$, with $t_r = t_1$ for nearest-neighbors, $t_r = t_2$ second-nearest-neighbors, and so on and so forth:

$$t_1 = \mathcal{J}_1 S + (2f_1 + 4f_2 + 4f_6 + 4f_5 + f_7)S^3, \tag{27}$$

$$t_2 = \mathcal{J}_2 S + (2f_3 + 4f_4 + 4f_6 + f_7)S^3, \tag{28}$$

$$t_v = \mathcal{J}_v S \quad \{v = 3, 4, 5, 6\}. \tag{29}$$

Despite not being explicitly designed to reproduce dynamical properties, the ML model's ability to capture the main features of

the magnon dispersion (as demonstrated in Fig. 4) is quite remarkable. The weight factor $\delta^{D,h}$ indirectly encodes the information about low-lying excitations in the model. The main quantitative discrepancy arises because the magnon dispersion predicted by the ML model is an analytical function of $\boldsymbol{k}$. In contrast, the magnon dispersion obtained from the KLM seems to have a discontinuous gradient near the band bottom. We must note, however, that this singular behavior originated from the long-range nature of the effective spin-spin interactions. Since we are cutting off the bilinear spin interactions beyond sixth neighbors in the ML model, the resulting magnon dispersion is necessarily an analytic function of $\boldsymbol{k}$. We then expect that the agreement between the magnon dispersions obtained from the high and low-energy models will systematically improve upon including longer-range interactions in the effective spin Hamiltonian. Comparisons of the excitation spectrum of both models can then be used as a criterion to fix the cutoff length scale of the spin-spin interactions included in the low-energy model.

DISCUSSION

Our work unravels a pathway to derive spin Hamiltonians when conventional methods, like perturbation theory, fail. One such example is a triangular KLM which gives rise to effective four-spin interactions that are non-analytic functions of J/t. We demonstrate that the effective low-energy model derived with the ML-assisted protocol accurately predicts the main phase boundaries of the phase diagram obtained by directly simulating the KLM. Remarkably, over 98 percent of the points in the phase diagram represented regions where the model was not trained, yet it accurately predicted the ground states, highlighting the generalizability of our model. While simulations of the KLM are numerically expensive, simulations of the effective spin model are roughly two orders of magnitude faster for the magnetic unit cell of 36 spins that we used in this work. The protocol presented in this work can then be used as an efficient tool to compute the phase diagram of the original KLM. The most expensive part is the generation of the training data set that is much smaller than the set required to build the phase diagram of the original KLM.

While numerical efficiency is an important aspect to consider, the most relevant aspect of our contribution is the use of ML to gain new insights into the underlying physical mechanisms. While this is always a main motivation behind the derivation of simple effective low-energy models, we are not aware of any previous attempt of using ML for this purpose in situations where the coefficients of the effective model are non-analytic functions of the coupling constant. This latter aspect represents the main innovation of our work relative the the pioneering work by Fujita et al.[31], who used supervised learning to derive an effective spin-1/2 Hamiltonian in the strong-coupling limit of a half-filled Hubbard model, i.e., in a situation where the effective low-energy model can be derived from degenerate perturbation theory. As we explained in detail, this qualitative difference introduces additional challenges to the problem, making it more intricate. Our approach is also qualitatively different from the more standard applications of ML, where the effective spin model works as a "black box" that can infer the energy of given spin configuration without revealing the nature of the effective spin interactions[60–62]. While these approaches tend to be numerically more accurate, their opaque nature can severely limit our understanding of the physical problem.

The simplicity of the effective low-energy model derived with our ML-assisted protocol provides insights into the stabilization mechanism of the different field-induced multi-$\boldsymbol{Q}$ magnetic orderings. Previous studies of the RKKY model showed that a finite easy-axis anisotropy is required to stabilize a magnetic field-induced triple-$\boldsymbol{Q}$ SkX phase[29]. In contrast, as shown in Fig. 3a, the zero-temperature phase diagram of the KLM includes a field-induced SkX phase even in absence of easy-axis anisotropy[1]. The

source of stabilization of SkX was speculated to be four-spin interactions mediated by itinerant electrons, which are expected to become significant when the ratio J/t is not much smaller than one. The procedure that is described in this work allows us to quantify the magnitude of *all symmetry allowed* four-spin interactions without introducing any bias, other than their limited range, in the types of interactions that can emerge at low energies. The good agreement between the phase diagrams of the original KLM and the effective low-energy model suggests that the field-induced SkX phase is indeed stabilized by effective four-spin interactions that turn out to be attractive between pairs of modes with different wave vector $\boldsymbol{Q}_\nu$. While this attractive channel, which is already present at zero magnetic field, was conjectured to be present in previous studies of zero field multi-$\boldsymbol{Q}$ orderings in metallic systems[9,13,14], it was not clear until now if it is also responsible for the field-induced SkX phases that were reported in more recent studies of centrosymmetric metallic systems[1]. In general, this attractive interaction between pairs of modes is expected to be present whenever the ordering wave vectors $\boldsymbol{Q}_\nu$ connect different regions of the Fermi surface. On a more intuitive level, the multi-$\boldsymbol{Q}$ ordering gaps out a bigger region of the Fermi surface in comparison to the single-$\boldsymbol{Q}$ phase.

As we discussed in Sec. II C, the low-energy spin model derived with the ML-assisted protocol provides insights into the stabilization mechanism of the SkX phase in KLMs. After recognizing that the stabilization arises from an attractive interaction between pairs of different modes $\{\boldsymbol{Q}_\nu, \boldsymbol{Q}_{\nu'}\}$ represented by the $\tilde{g}_8$ and $\tilde{g}_5$ terms of Eq. (9), and that the $\tilde{g}_3$ contribution, which is usually invoked as the most natural driver of the SkX phase[42,55], is actually increasing the energy of the SkX relative to the competing orderings, we can infer that field-induced square SkX phases should also arise in KLMs with tetragonal symmetry. The key observation is that the $\tilde{g}_3$ term is no longer relevant because the square SkX is a double-$\boldsymbol{Q}$ ($\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$) ordering with $\boldsymbol{Q}_1 \cdot \boldsymbol{Q}_2 = \boldsymbol{0}$. This means that the square SkX phase will still benefit from the attractive interaction between the pair of different modes, while the penalization from a $\tilde{g}_3$-like term with $\boldsymbol{Q}_3 = -\boldsymbol{Q}_1 - \boldsymbol{Q}_2$ will be significantly smaller because $\boldsymbol{Q}_3$ is no longer a fundamental wave vector, but a second harmonic ($\|\boldsymbol{S}_{\boldsymbol{Q}_3}\| \ll \|\boldsymbol{S}_{\boldsymbol{Q}_1}\|, \|\boldsymbol{S}_{\boldsymbol{Q}_2}\|$). The only remaining obstacle for the stabilization of the square SkX is the RKKY energy cost of the second harmonic component $\|\boldsymbol{S}_{\boldsymbol{Q}_1+\boldsymbol{Q}_2}\|$. In the perturbative regime, $|J| \ll |t|$, this energy cost of order $J^2 \|\boldsymbol{S}_{\boldsymbol{Q}_1+\boldsymbol{Q}_2}\|^2 (\chi_{\boldsymbol{Q}_1} - \chi_{\boldsymbol{Q}_1+\boldsymbol{Q}_2})$ can be reduced by choosing tetragonal materials/models with square-like Fermi surfaces. In other words, square SkXs should emerge for $J^2 \|\boldsymbol{S}_{\boldsymbol{Q}_1+\boldsymbol{Q}_2}\|^2 (\chi_{\boldsymbol{Q}_1} - \chi_{\boldsymbol{Q}_1+\boldsymbol{Q}_2}) N \lesssim \tilde{g}_8 \|\boldsymbol{S}_{\boldsymbol{Q}_1}\|^2 \|\boldsymbol{S}_{\boldsymbol{Q}_2}\|^2$.

Finally, it is interesting to note that the $\tilde{g}_7$ biquadratic interaction, which has been adopted in previous works as the only 4-spin interaction generated by the KLM[12,16], is revealed by our ML protocol to be a subdominant term (see Table 2). Since this biquadratic interaction was selected based on a trend observed in the divergent terms of the perturbative expansion in J/t, we conclude that such procedure is not reliable to quantify the relative strength of the different symmetry allowed four-spin interactions.

These simple examples illustrate how ML-assisted protocols can be used to extract guiding principles. While it is natural to assume that the $\tilde{g}_3$ term is responsible for the stabilization of SkXs in hexagonal lattices[55], this "intuitive" argument is the result of a cognitive bias which, in absence of additional information, selects the simplest explanation over those that involve a larger number of assumptions and variables. While Occam's razor is a very useful principle to guide our understanding of complex systems, it can also lead to oversimplifications caused by lack of validation of the implicit assumptions. The unbiased nature of the ML-assisted protocol not only allows us to correct these assumptions, but also gives us enough information to infer new guiding principles. In other words, besides providing an efficient tool to accelerate the computation of phase diagrams, the ML approach is also an efficient *learning tool* that can be used to understand mechanisms and accelerate discovery.
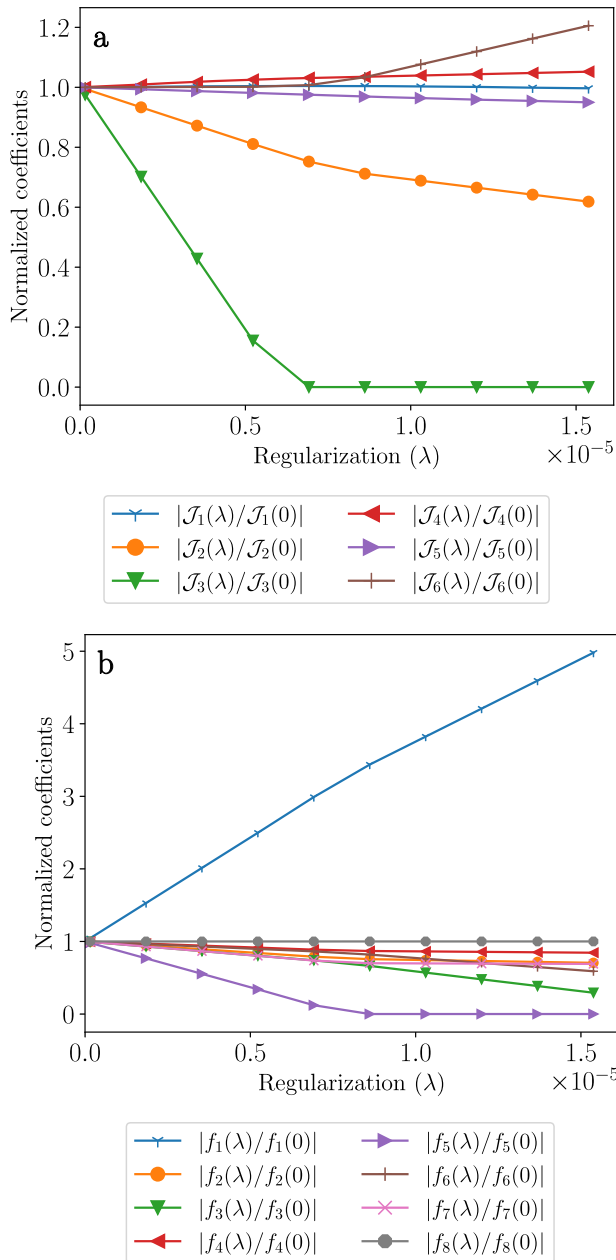
function to eliminate the least important interactions:

$$\text{Cost}(\{c_j\}) = \frac{1}{\bar{N}} \sum_{D,h} \sum_{i=0}^{\bar{N}-1} \delta_i^{D,h} \left(E_i^{D,h} - \tilde{E}_i^{D,h} \right)^2 + \lambda \sum_j |c_j|. \quad (30)$$

The added second term in the cost function, with a positive λ, penalizes large values of the coupling constants. We could have also used $L2$ regularization which adds "squared magnitude" of the coefficients as penalty term to the loss function instead of the "absolute value of magnitude" of coefficients in $L1$ regularization. Both $L1$ and $L2$ regularizations avoid over-fitting, but the key difference is that while $L2$ pushes the coefficients to become small, $L1$ regularization gives sparse estimates (in a high dimensional space it shrinks the less important features' coefficients to zero). Since our main purpose is to obtain the simplest model that reproduces the low-energy physics of the original high-energy model, $L1$ regularization is the preferred option. The optimal solution is then obtained by minimizing the new cost function with the gradient descent method or equivalently with multivariate weighted regression with $L1$ penalty. Once this procedure eliminates the least important interactions, we get an ansatz for the minimal Hamiltonian. The new ansatz is then optimized without the regularization term to get the actual estimation of the minimal model.

Determining the range for hyperparameter λ to deduce the most important features can be challenging. We first calculated the weighted mean square error of the full data (weight for each data entry is given by $\delta_i^{D,h}$) and then selected a range of λ so that the contribution from the regularization term was in the range from zero to 100% of the calculated weighted mean square error. By using this range, we avoided the problem of too high regularization penalty. As the value of λ is gradually increased, J_3 and f_5 go to zero (see Fig. 5). After eliminating these two interactions, we select the final ansatz i.e., the remaining interactions and apply the last step of our algorithm to get the minimal model (Table 1). In principle, we could have eliminated more interactions by increasing the value of λ further, but that would have resulted in a less accurate low-energy model.

DATA AVAILABILITY
Data supporting the findings of this study can be obtained from the corresponding author upon a reasonable request.

CODE AVAILABILITY
The computer codes supporting the findings of this study can be obtained from the corresponding author upon a reasonable request.

Fig. 5 Evolution of parameters as the strength of $L1$ regularization is increased. Each parameter is normalized by dividing its values at various λs with its value at $\lambda = 0$.

METHODS

Sparse modeling
In setting up the trial Hamiltonian, it is usually not clear how to determine the number of terms to be included in the initial Hamiltonian. Too many terms can lead to over-fitting, reducing the generalization capability of the acquired model. Furthermore, a *minimal* effective Hamiltonian must include the minimum number of symmetry allowed terms necessary to reproduce the phase diagram and the values of relevant physical observables that are predicted by the high-energy model. After beginning with an initial Hamiltonian comprising of all the RKKY interaction up to sixth neighbor and all the four-spin interactions up to second neighbor, we implement the $L1$ norm regularization of the cost

REFERENCES
1. Wang, Z. & Batista, C. D. Skyrmion crystals in the triangular kondo lattice model. https://arxiv.org/abs/2111.13976 (2022).
2. Alonso, J. L., Capitán, J. A., Fernández, L. A., Guinea, F. & Martín-Mayor, V. Monte Carlo determination of the phase diagram of the double-exchange model. *Phys. Rev. B* **64**, 054408 (2001).
3. Furukawa, N. & Motome, Y. Order N Monte Carlo algorithm for fermion systems coupled with fluctuating adiabatical fields. *J. Phys. Soc. Jpn.* **73**, 1482–1489 (2004).
4. Alvarez, G., Nukala, P. K. V. V. & D'Azevedo, E. Fast diagonalization of evolving matrices: application to spin-fermion models. *J. Stat. Mech.* **2007**, P08007 (2007).
5. Barros, K. & Kato, Y. Efficient Langevin simulation of coupled classical fields and fermions. *Phys. Rev. B* **88**, 235101 (2013).
6. Ruderman, M. A. & Kittel, C. Indirect exchange coupling of nuclear magnetic moments by conduction electrons. *Phys. Rev.* **96**, 99–102 (1954).

7. Kasuya, T. A theory of metallic ferro- and antiferromagnetism on Zener's model. *Prog. Theor. Phys.* **16**, 45–57 (1956).

8. Yosida, K. Magnetic properties of Cu-Mn alloys. *Phys. Rev.* **106**, 893–898 (1957).

9. Akagi, Y., Udagawa, M. & Motome, Y. Hidden multiple-spin interactions as an origin of spin scalar chiral order in frustrated kondo lattice models. *Phys. Rev. Lett.* **108**, 096401 (2012).

10. Ozawa, R. et al. Vortex crystals with chiral stripes in itinerant magnets. *J. Phys. Soc. Jpn.* **85**, 103703 (2016).

11. Batista, C. D., Lin, S.-Z., Hayami, S. & Kamiya, Y. Frustration and chiral orderings in correlated electron systems. *Rep. Prog. Phys.* **79**, 084504 (2016).

12. Hayami, S., Ozawa, R. & Motome, Y. Effective bilinear-biquadratic model for noncoplanar ordering in itinerant magnets. *Phys. Rev. B* **95**, 224424 (2017).

13. Kurz, P., Bihlmayer, G., Hirai, K. & Blügel, S. Three-dimensional spin structure on a two-dimensional lattice: Mn/Cu(111). *Phys. Rev. Lett.* **86**, 1106–1109 (2001).

14. Martin, I. & Batista, C. D. Itinerant electron-driven chiral magnetic ordering and spontaneous quantum hall effect in triangular lattice models. *Phys. Rev. Lett.* **101**, 156402 (2008).

15. Paul, S., Haldar, S., von Malottki, S. & Heinze, S. Role of higher-order exchange interactions for skyrmion stability. *Nat. Commun.* **11**, 4756 (2020).

16. Hayami, S. & Motome, Y. Topological spin crystals by itinerant frustration. *J. Phys. Condens. Matter* **33**, 443001 (2021).

17. Ye, J. et al. Berry phase theory of the anomalous hall effect: application to colossal magnetoresistance manganites. *Phys. Rev. Lett.* **83**, 3737–3740 (1999).

18. Onoda, M., Tatara, G. & Nagaosa, N. Anomalous hall effect and skyrmion number in real and momentum spaces. *J. Phys. Soc. Jpn.* **73**, 2624–2627 (2004).

19. Yi, S. D., Onoda, S., Nagaosa, N. & Han, J. H. Skyrmions and anomalous Hall effect in a Dzyaloshinskii-Moriya spiral magnet. *Phys. Rev. B* **80**, 054416 (2009).

20. Kato, Y., Martin, I. & Batista, C. D. Stability of the spontaneous quantum hall state in the triangular kondo-lattice model. *Phys. Rev. Lett.* **105**, 266405 (2010).

21. Hamamoto, K., Ezawa, M. & Nagaosa, N. Quantized topological Hall effect in skyrmion crystal. *Phys. Rev. B* **92**, 115417 (2015).

22. Göbel, B., Mook, A., Henk, J. & Mertig, I. Unconventional topological Hall effect in skyrmion crystals caused by the topology of the lattice. *Phys. Rev. B* **95**, 094413 (2017).

23. Shao, Q. et al. Topological Hall effect at above room temperature in heterostructures composed of a magnetic insulator and a heavy metal. *Nat. Electron.* **2**, 182–186 (2019).

24. Mallik, R., Sampathkumaran, E. V., Paulose, P. L., Sugawara, H. & Sato, H. Magnetic anomalies in Gd_2PdSi_3. *Pramana - J. Phys.* **51**, 505–509 (1998).

25. Saha, S. R. et al. Magnetic anisotropy, first-order-like metamagnetic transitions, and large negative magnetoresistance in single-crystal Gd_2PdSi_3. *Phys. Rev. B* **60**, 12162–12165 (1999).

26. Kurumaji, T. et al. Skyrmion lattice with a giant topological Hall effect in a frustrated triangular-lattice magnet. *Science* **365**, 914–918 (2019).

27. Chandragiri, V., Iyer, K. K. & Sampathkumaran, E. V. Magnetic behavior of $Gd_3Ru_4Al_{12}$, a layered compound with distorted kagomé net. *J. Phys. Condens. Matter* **28**, 286002 (2016).

28. Hirschberger, M. et al. Skyrmion phase and competing magnetic orders on a breathing kagomé lattice. *Nat. Commun.* **10**, 5831 (2019).

29. Wang, Z., Su, Y., Lin, S.-Z. & Batista, C. D. Skyrmion crystal from RKKY interaction mediated by 2D electron gas. *Phys. Rev. Lett.* **124**, 207201 (2020).

30. Takagi, R. et al. Multiple-q noncollinear magnetism in an itinerant hexagonal magnet. *Sci. Adv.* **4**, eaau3402 (2018).

31. Fujita, H., Nakagawa, Y. O., Sugiura, S. & Oshikawa, M. Construction of Hamiltonians by supervised learning of energy and entanglement spectra. *Phys. Rev. B* **97**, 075114 (2018).

32. Mühlbauer, S. et al. Skyrmion lattice in a chiral magnet. *Science* **323**, 915–919 (2009).

33. Yu, X. Z. et al. Real-space observation of a two-dimensional skyrmion crystal. *Nature* **465**, 901–904 (2010).

34. Yu, X. Z. et al. Near room-temperature formation of a skyrmion crystal in thin-films of the helimagnet FeGe. *Nat. Mater.* **10**, 106–109 (2011).

35. Seki, S., Yu, X. Z., Ishiwata, S. & Tokura, Y. Observation of skyrmions in a multi-ferroic material. *Science* **336**, 198–201 (2012).

36. Adams, T. et al. Long-wavelength helimagnetic order and skyrmion lattice phase in Cu_2OSeO_3. *Phys. Rev. Lett.* **108**, 237204 (2012).

37. Yu, X. et al. Magnetic stripes and skyrmions with helicity reversals. *Proc. Natl. Acad. Sci. USA* **109**, 8856–8860 (2012).

38. Yu, X. Z. et al. Biskyrmion states and their current-driven motion in a layered manganite. *Nat. Commun.* **5**, 3198 (2014).

39. Dzyaloshinsky, I. A thermodynamic theory of "weak" ferromagnetism of anti-ferromagnetics. *J. Phys. Chem. Solids* **4**, 241–255 (1958).

40. Moriya, T. AnisoTropic Superexchange Interaction And Weak Ferromagnetism. *Phys. Rev.* **120**, 91–98 (1960).

41. Okubo, T., Chung, S. & Kawamura, H. Multiple-q states and the skyrmion lattice of the triangular-lattice heisenberg antiferromagnet under magnetic fields. *Phys. Rev. Lett.* **108**, 017206 (2012).

42. Leonov, A. O. & Mostovoy, M. Multiply periodic states and isolated skyrmions in an anisotropic frustrated magnet. *Nat. Commun.* **6**, 8275 (2015).

43. Lin, S.-Z. & Hayami, S. Ginzburg-Landau theory for skyrmions in inversion-symmetric magnets with competing interactions. *Phys. Rev. B* **93**, 064430 (2016).

44. Hayami, S., Lin, S.-Z. & Batista, C. D. Bubble and skyrmion crystals in frustrated magnets with easy-axis anisotropy. *Phys. Rev. B* **93**, 184413 (2016).

45. Jonietz, F. et al. Spin transfer torques in MnSi at ultralow current densities. *Science* **330**, 1648–1651 (2010).

46. Yu, X. Z. et al. Skyrmion flow near room temperature in an ultralow current density. *Nat. Commun.* **3**, 988 (2012).

47. Schulz, T. et al. Emergent electrodynamics of skyrmions in a chiral magnet. *Nat. Phys.* **8**, 301–304 (2012).

48. Nagaosa, N. & Tokura, Y. Topological properties and dynamics of magnetic skyrmions. *Nat. Nanotechnol.* **8**, 899–911 (2013).

49. Zhang, S.-S., Ishizuka, H., Zhang, H., Halász, G. B. & Batista, C. D. Real-space Berry curvature of itinerant electron systems with spin-orbit interaction. *Phys. Rev. B* **101**, 024420 (2020).

50. Simeth, W. et al. A microscopic Kondo lattice model for the heavy fermion antiferromagnet $CeIn_3$. https://arxiv.org/abs/2208.02211 (2022).

51. Weiße, A., Wellein, G., Alvermann, A. & Fehske, H. The kernel polynomial method. *Rev. Mod. Phys.* **78**, 275–306 (2006).

52. Wang, Z., Chern, G.-W., Batista, C. D. & Barros, K. Gradient-based stochastic estimation of the density matrix. *J. Chem. Phys.* **148**, 094107 (2018).

53. Wang, Z., Barros, K., Chern, G.-W., Maslov, D. L. & Batista, C. D. Resistivity minimum in highly frustrated itinerant magnets. *Phys. Rev. Lett.* **117**, 206601 (2016).

54. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).

55. Garel, T. & Doniach, S. Phase transitions with spontaneous modulation-the dipolar Ising ferromagnet. *Phys. Rev. B* **26**, 325–329 (1982).

56. Kurumaji, T. et al. Néel-type skyrmion lattice in the tetragonal polar magnet $VOSe_2O_5$. *Phys. Rev. Lett.* **119**, 237201 (2017).

57. Wang, Z., Su, Y., Lin, S.-Z. & Batista, C. D. Meron, skyrmion, and vortex crystals in centrosymmetric tetragonal magnets. *Phys. Rev. B* **103**, 104408 (2021).

58. Zhang, S., Changlani, H. J., Plumb, K. W., Tchernyshyov, O. & Moessner, R. Dynamical structure factor of the three-dimensional quantum spin liquid candidate $NaCaNi_2F_7$. *Phys. Rev. Lett.* **122**, 167203 (2019).

59. Chern, G.-W., Barros, K., Wang, Z., Suwa, H. & Batista, C. D. Semiclassical dynamics of spin density waves. *Phys. Rev. B* **97**, 035120 (2018).

60. Zhang, P. & Chern, G.-W. Arrested phase separation in double-exchange models: large-scale simulation enabled by machine learning. *Phys. Rev. Lett.* **127**, 146401 (2021).

61. Novikov, I., Grabowski, B., Körmann, F. & Shapeev, A. Magnetic moment tensor potentials for collinear spin-polarized materials reproduce different magnetic states of bcc Fe. *npj Comput. Mater.* **8**, 1–6 (2022).

62. Arale Brännvall, M., Gambino, D., Armiento, R. & Alling, B. Machine learning approach for longitudinal spin fluctuation effects in bcc Fe at T_c and under earth-core conditions. *Phys. Rev. B* **105**, 144417 (2022).

AUTHOR CONTRIBUTIONS

The project was initiated by V.S. and C.D.B. and was coordinated by them. V.S. developed the codes for the machine learning protocol and the optimization with low-energy Hamiltonian. Z.W. developed the optimization and LL dynamics code for high-energy Hamiltonian. V.S. conducted the numerical simulations and machine learning analysis with guidance from C.D.B. The manuscript was written by V.S. and C.D.B. with input and review from Z.W. All the authors participated in discussing the data and its interpretation.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-023-01137-x.

Correspondence and requests for materials should be addressed to Vikram Sharma.

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

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