


Machine-learning recognition of Dzyaloshinskii-Moriya interaction from magnetometryBradley J. Fugetta¹, Zhijie Chen¹, Dhritiman Bhattacharya¹, Kun Yue², Kai Liu¹, Amy Y. Liu¹, and Gen Yin^{1,*}¹*Department of Physics, Georgetown University, Washington, DC 20057, USA*²*Nvidia Corporation, Santa Clara, California 95051, USA* (Received 10 April 2023; revised 3 August 2023; accepted 30 August 2023; published 4 October 2023)

The Dzyaloshinskii-Moriya interaction (DMI), which is the antisymmetric part of the exchange interaction between neighboring local spins, winds the spin manifold and can stabilize nontrivial topological spin textures. Since topology is a robust information carrier, characterization techniques that can extract the DMI magnitude are important for the discovery and optimization of spintronic materials. Existing experimental techniques for quantitative determination of DMI, such as high-resolution magnetic imaging of spin textures and measurement of magnon or transport properties, are time-consuming and require specialized instrumentation. Here we show that a convolutional neural network can extract the DMI magnitude from minor hysteresis loops, or magnetic “fingerprints,” of a material. These hysteresis loops are readily available by conventional magnetometry measurements. This provides a convenient tool to investigate topological spin textures for next-generation information processing.

DOI: [10.1103/PhysRevResearch.5.043012](https://doi.org/10.1103/PhysRevResearch.5.043012)**I. INTRODUCTION**

The Dzyaloshinskii-Moriya interaction (DMI) is an antisymmetric exchange coupling between neighboring spins, and it is nonzero only in materials with broken central symmetry [1,2]. The symmetry can be broken intrinsically by the crystal structure, as in the case of B20 compounds [3–8], or it can be broken extrinsically by designing magnetic multilayer heterostructures [9–14]. Due to its antisymmetric nature, the DMI favors perpendicular configurations between neighboring spins, winding the spin manifold. As a result, it plays an essential role for topological spin textures, including magnetic skyrmions [11], vortices [15], bimerons [16], hedgehogs [17–20], chiral domain walls [21,22], and hopfions [23]. These textures are either directly stabilized by the DMI, or their behaviors are strongly impacted by the DMI magnitude. Such topological spin textures have promising potential as information carriers in next-generation spintronic devices for low-power and high-speed applications [11,24,25]. Even in systems with uniform spins, the DMI can induce phenomena such as nonreciprocal and topological magnon spectra [26–28], which are useful for radiofrequency devices. Quantitative understanding and control of the DMI magnitude in spintronic systems is therefore important for both fundamental and application purposes.

Significant efforts have been devoted to the quantitative determination of the DMI magnitude, especially for

thin-film multilayers, which are essential for device applications [29]. In fact, for a thin-film multilayer, the magnitude of the extrinsic DMI can be continuously modulated by carefully controlling the vertical profile of the heterostructure [30,31]. In addition, tunable control of interfacial DMI has recently been demonstrated via chemisorption or ionic gating, even after the materials systems have been synthesized [32–34]. Since details such as chemical bonding and atomic alignment at interfaces are important, theoretical prediction of extrinsic DMIs from first principles is very challenging. Experimentally, techniques such as domain-wall imaging [13,35–37], Brillouin light scattering [31,38–40], and loop-shift measurements using spin-orbit torque setups [37,41–44] can be used to obtain the DMI magnitude. However, these experiments are nontrivial and sometimes require trial-and-error iterations between experiments and modeling.

Recently, predictive machine-learning models have been demonstrated to successfully extract DMI magnitudes from ground-state spin textures with good accuracy [45–48]. These approaches typically use phenomenological Hamiltonians considering leading-order terms, such as the symmetric Heisenberg exchange, the uniaxial anisotropy, the long-range dipolar interaction, the Zeeman coupling, and the DMI. To generate the training data, either dynamical or Monte Carlo simulations are used to find the ground state of a Hamiltonian given a set of parameters. A neural network can then be trained to correlate the ground-state spin texture (inputs) and the Hamiltonian parameters (outputs) including the DMI. The success of this practice suggests that the ground-state spin texture indeed contains information about the intricate competition between the DMI and other terms in the Hamiltonian. This is intuitive since the DMI is the only term in the Hamiltonian that favors spin winding with a uniform chirality. However, experimental imaging of spin textures requires high spatial resolution as well as the ability to resolve the

*Gen.Yin@georgetown.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

magnetization vectors. High-quality, exposed pristine surfaces are therefore often necessary.

In contrast to magnetic imaging, magnetometry is easily accessible, e.g., via SQUID (superconducting quantum interference device), vibrating sample magnetometry (VSM), or the magneto-optical Kerr effect. However, as conventional magnetometry typically measures ensemble-averaged magnetic responses of a given system, it is not obvious that information about the DMI can be extracted directly from magnetometry alone. Nevertheless, magnetometry has been demonstrated to contain detailed information about the magnetization reversal process. Specifically, the first-order reversal curve (FORC) method, which utilizes many partial hysteresis curves, has been used to “fingerprint” the magnetization reversal process, providing information well beyond the ensemble average extracted from conventional major hysteresis loop measurements [49–51].

Here, we show that a convolutional neural network (CNN) can indeed extract the DMI magnitude from FORCs. With a data set of FORCs generated for thin films via micromagnetic simulations, we demonstrate that good accuracy can be achieved when the FORCs exhibit rich features. The network performs particularly well when the SO(3) symmetry is not significantly broken, i.e., when the uniaxial anisotropy is comparable to the shape anisotropy of a thin film. This neural network provides a convenient tool to extract DMI and investigate topological spin textures in magnetic multilayers using magnetometry.

II. FIRST-ORDER REVERSAL CURVES

The FORC method measures a collection of many partial hysteresis curves, which are known to contain important information about the distribution of magnetic properties as well as interactions between magnetic elements within a system. Typically, a sample is first saturated in a positive field, then brought to a particular reversal field H_R , and the magnetization M is measured with increasing applied field H back to positive saturation, thus tracing out a single FORC. The measurement is repeated at successively more negative H_R , leading to a family of FORCs. A second-order derivative is taken to extract a FORC distribution, $\rho(H, H_R) = -[\partial^2 M(H, H_R)/\partial H \partial H_R]/2$ [50,52,53]. The FORC distribution has been used to probe irreversible magnetization switching [50], capture magnetic distributions [54–56], distinguish different magnetic phases [57] and switching mechanisms [56], and study magnetic interactions [51,58,59].

While the FORC distribution mentioned above can be used to extract useful information, the vast parameter space of many competing terms in the Hamiltonian makes it difficult to interpret the DMI magnitude. Specifically, the analysis that can extract the DMI magnitude from FORCs is essentially unknown. However, this task can be easily converted to a pattern recognition problem for artificial intelligence (AI). With supervised learning, the first several convolutional layers in a CNN can usually identify complicated, nonlinear filters to extract needed features autonomously with relatively low computational and memory costs. To train such a CNN, we generated labeled data using micromagnetic simulations.

Consider a phenomenological Hamiltonian

$$H = -J_0 \sum_{\langle i,j \rangle} \vec{s}_i \cdot \vec{s}_j - \sum_{\langle i,j \rangle} D_0 (\hat{z} \times \hat{r}_{ij}) \cdot (\vec{s}_i \times \vec{s}_j) - K_0 \sum_i (\vec{s}_i \cdot \hat{n})^2 - \mu \sum_i (\mu_0 \vec{H}_{\text{ext}} + \mu_0 \vec{H}_{\text{dipole},i}) \cdot \vec{s}_i \quad (1)$$

where $\{\vec{s}_i\}$ denotes the normalized classical spins defined on a square lattice $\{i\}$ in the x - y plane, \vec{r}_{ij} is the position vector connecting sites i and j , \hat{n} is the direction of the uniaxial anisotropy, and $\langle i, j \rangle$ denotes pairs of neighboring sites. The evolution of the spin texture $\{\vec{s}_i\}$ governed by the Hamiltonian given by Eq. (1) can be simulated using micromagnetic simulations implemented by mumax³ [60], taking $A_{\text{ex}} = J_0/2a$ as the exchange stiffness, where J_0 is the symmetric Heisenberg exchange energy and a is the lattice constant. Similarly, $D = D_0/a^2$ is the DMI magnitude and $K_u = K_0/a^2 t$ is the uniaxial anisotropy energy density in a film of thickness t . Both the external magnetic field, \vec{H}_{ext} , and the dipolar fields $\{\vec{H}_{\text{dipole},i}\}$ are included, assuming periodic boundary conditions within the thin-film plane and a uniform magnetic moment of $\mu = a^2 t M_s$ within each site, where M_s is the saturation magnetization. FORCs are then simulated by the average spin of the ground state during the scan of \vec{H}_{ext} applied perpendicular to the thin film. The thermal fluctuation of finite temperature T is included by a stochastic effective field.

Each family of FORCs was simulated using a simulation box with an extent of 120 nm in the x and y directions split into a 16×16 grid with a periodic boundary condition of five repeats in the plane of the film. All samples had this geometry and used a damping factor of $\alpha = 0.5$ to enable fast access to the ground state. To simulate experimental uncertainty, both the magnitude of K_u and the direction of the easy axis (\hat{n}) were given an extra layer of randomness. The direction of the easy axis $\hat{n}(\theta, \phi)$ for the entire sample was chosen with the azimuthal angle ϕ being uniformly random within $[0, 2\pi]$ and the polar angle θ chosen from a normal distribution centered at 0 with a standard deviation of σ_θ , which was itself chosen from a uniform distribution within $[0^\circ, 10^\circ]$. The magnitude of K_u for the sample was chosen according to a normal distribution centered at K with a standard deviation of σ_K , which were themselves chosen from uniform distributions. We used Heun’s method implemented in mumax³ to solve the Landau-Lifshitz-Gilbert equation using adaptive time-stepping with $\text{MinDt} = 2.5 \times 10^{-13}$ s and $\text{MaxDt} = 5 \times 10^{-12}$ s. At every value of \vec{H}_{ext} , time first progressed for 0.1 ns and then evolved in increments of 0.125 ns until $\langle s_z \rangle$ changed by less than 5×10^{-3} between step increments.

III. MODEL AND TRAINING

To interface FORCs with a CNN, we convert each family of FORCs [one typical example shown in Fig. 1(a)] to an information-dense input image [Fig. 1(b)]. First, the full range of scanning magnetic field $[H_{\text{min}}, H_{\text{max}}]$ is discretized to 61 steps. A family of FORCs can then be rearranged as a 61×61 image, with each pixel denoting one discrete step of the field scan. During the scan, the normalized magnetization value given by $\langle s_z \rangle$ is discretized to integers within $[0, 255]$, which are denoted by the brightness of the pixels in Fig. 1(b). For each minor loop in Fig. 1(a), the applied

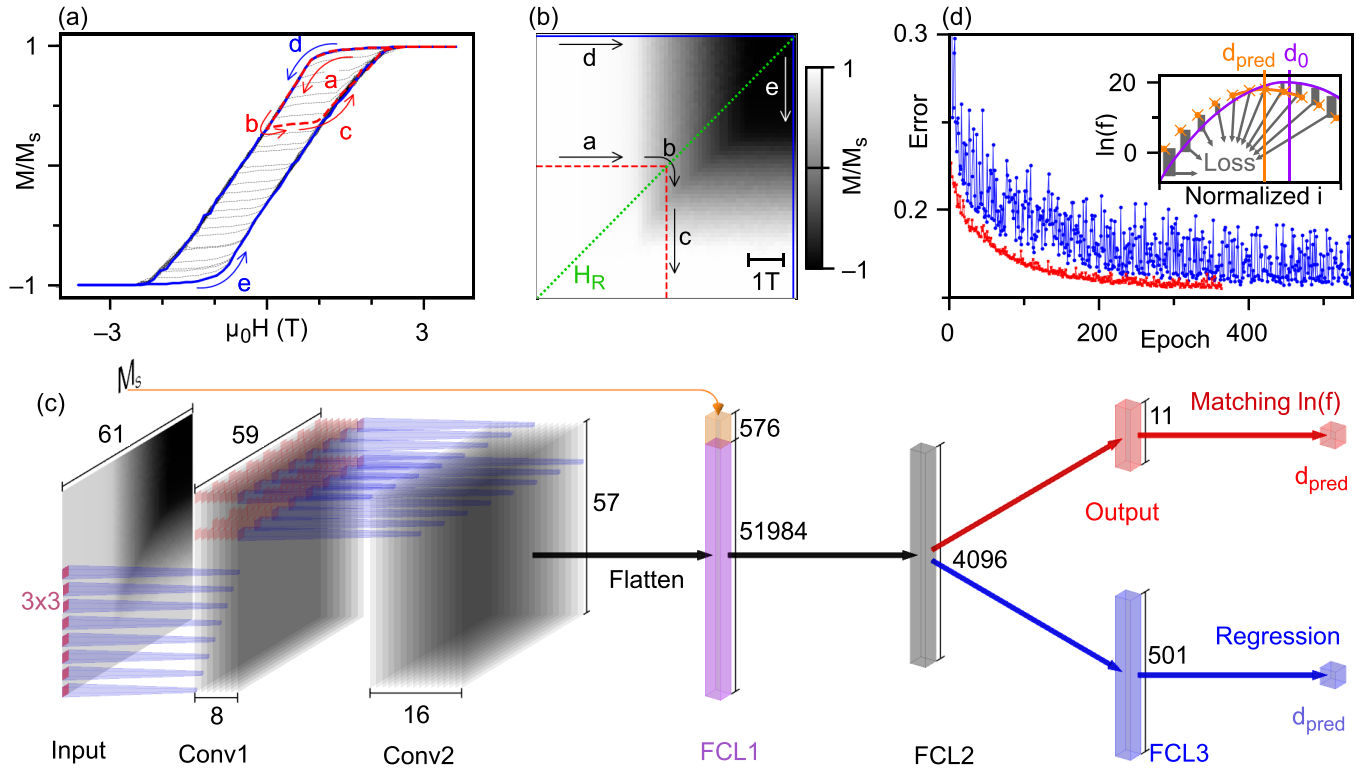


FIG. 1. Setup of the machine-learning problem. (a) A typical family of FORCs in our data set. One minor loop is shown by the red dashed line, whereas the full hysteresis loop is denoted by the solid blue line. Other minor loops are illustrated by the gray dotted curves. (b) The input image corresponding to the FORCs shown in (a), with each pixel denoting one discrete point during the field scan. The magnetization during the field scan is normalized to a 1-byte integer between 0 and 255, corresponding to the brightness of each pixel. (c) The structure of the CNN. The top path (red) denotes the distribution matching method, whereas the lower path (blue) represents a conventional regression. (d) The convergence of the training. The distribution-matching and conventional regression results are denoted by the red and blue curves, respectively. The inset illustrates the scheme of the loss function in the distribution-matching method.

field starts from positive saturation H_{\max} (the top right corner), then hits the reversal field H_R , and eventually goes back to positive saturation (red arrows $a \rightarrow b \rightarrow c$). Correspondingly, the pixel in Fig. 1(b) first scans horizontally from the left to the right (arrow a). When the scanning field in Fig. 1(a) reaches H_R , the scanning pixel hits the diagonal green dotted line in Fig. 1(b). This pixel then scans vertically downwards (arrows $b \rightarrow c$) as the field in Fig. 1(a) scans back to positive saturation. With this arrangement, each minor loop traced out by a FORC in Fig. 1(a) is mapped to a horizontal and a vertical segment of pixels in Fig. 1(b) connected at the dotted green line. When $H_R = H_{\max}$, the minor loop in Fig. 1(a) shrinks to zero, contributing only a bright pixel at the lower left corner in Fig. 1(b). As H_R becomes more negative, the minor loops in Fig. 1(a) enlarge, and eventually recover the full hysteresis loop when $H_R = H_{\min}$. Correspondingly, the path in Fig. 1(b) gradually becomes larger and eventually completes the image by filling the outermost row and column of pixels ($d \rightarrow e$). In our simulations, the range of the scan is $[-3.6 \text{ T}, +3.6 \text{ T}]$. Each pixel in the image thus corresponds to one step of 0.12 T . This choice of step size balances feature resolution and simulation time.

To ensure data diversity, each family of simulated FORCs is parametrized by $\{M_s, T, A_{\text{ex}}, K_u, D, t, \theta, \phi\}$, with each element being a double-precision floating-point number ran-

domly generated using the values and bounds shown in Table I. The normalized DMI magnitude $d = \frac{D-D_{\min}}{D_{\max}-D_{\min}}$ is used as the label. We determined the range of these material parameters based on reported experimental values. For M_s , we set the upper bound according to bulk cobalt, whereas the lower bound is set to mimic the reduced value in typical magnetic multilayers with nonmagnetic or antiferromagnetic components [57,61–64]. A_{ex} is set to vary between the cobalt value obtained in first-principles calculations [65] and those values previously obtained in soft magnetic multilayers or alloys [66,67]. The value of K is set between zero and a typical value for a Co/Pd multilayer with strong perpendicular magnetic anisotropy [57,61]. As the main target of this work, we set the DMI magnitude from zero to $5 \times 10^{-3} \text{ J/m}^2$, an upper bound exceeding most observed DMIs in multilayers [62,68,69]. We chose the bounds of temperature based on

TABLE I. The bounds of simulation parameters.

| | M_s (A/m) | T (K) | A_{ex} (J/m) | K (J/m ³) | D (J/m ²) | t (nm) | σ_θ (deg) | σ_K (%) |
|-----|------------------|------------|--------------------------|----------------------------|----------------------------|-------------|--------------------------|-------------------|
| Min | 2×10^5 | 135 | 1×10^{-12} | 0 | 0 | 10 | 0 | 0 |
| Max | 14×10^5 | 19 420 | 35×10^{-12} | 10^6 | 0.005 | 100 | 10 | 20 |

TABLE II. Parameters of our CNN.

| Layer | Input | Output | Kernel Size | Dropout |
|-------|----------------|----------------|--------------|---------|
| Conv1 | 61×61 | 59×59 | 3×3 | 0 |
| Conv2 | 59×59 | 57×57 | 3×3 | 0.5 |
| FCL1 | 52 560 | 4096 | N/A | 0.5 |
| FCL2 | 4096 | 11 | N/A | 0.5 |

a rough estimate of the Curie temperature: $T_C \approx \frac{A_{\text{ex}d}}{k_B}$, such that $T \in [\min(T_C) - \Delta T, \max(T_C) + \Delta T]$, leaving margins of $\Delta T = \frac{3}{4} \min(T_C)$ to sample the cases of $T < T_C$ and $T > T_C$. Note that these are the simulation temperatures. In micromagnetic simulations, the coarse-graining process leads to an effective physical temperature that is rescaled from the simulation temperature. In our case, the maximum effective temperature is estimated to be roughly 10–20 times smaller than the maximum simulation temperature [70,71].

We utilized Pytorch, a machine learning framework with a Python interface, to construct and train our CNN. The specific model of our CNN is illustrated in Fig. 1(c), where two convolutional layers (Conv1 and Conv2) and two fully connected layers (FCL1 and FCL2) are employed with ReLU (Rectified Linear Unit) activation function. The details of the model are listed in Table II. During training, a dropout rate of 0.5 was applied to Conv2 and all fully connected layers. The configuration of our CNN was informed by a series of trial-and-error tests in which we varied, among other things, the number of convolutional layers, the amount of feature maps in each layer, the size and number of fully connected layers, the activation functions between the layers, and the combination of dropout rates used for the layers. Similarly, we surveyed the parameters of the ADADELTA adaptive learning rate method, and eventually used a momentum of $\rho = 0.9$, a learning rate of $\gamma = 0.01$, and a weight decay of $\lambda = 0$ [72].

In practice, since M_s is readily available from typical magnetometry measurements, it is feasible to include it, along with the normalized FORCs, as input to the neural network. This is similar to the idea of a residual network [73] where additional information can skip some sections of the network to mimic long-term memory. In our CNN, M_s is first normalized to $[0,10]$ and then directly fed to FCL1 by appending 576 nodes of the same value, along with the flattened feature maps extracted by all the convolutional layers. The number of M_s nodes was large enough to avoid concerns of being dropped out while being small enough not to overwhelm the feature map information.

Different from a typical regression model, we implemented a distribution matching method as described by Eqs. (2)–(4),

$$\bar{y} = \exp[\vec{\mathcal{F}}_{\text{CNN}}(\bar{x})], \quad (2)$$

$$\mathcal{L} = \langle [\ln \bar{y} - \ln \vec{f}(d_0)]^2 \rangle, \quad (3)$$

$$d_{\text{pred}} = \arg \min'_d [\ln \bar{y} - \ln \vec{f}(d)]^2, \quad (4)$$

where $\vec{\mathcal{F}}_{\text{CNN}}(\bar{x})$ is the CNN output given input \bar{x} . Unlike a conventional regression method, the output involves more nodes than needed (in our case 11). We then map the output

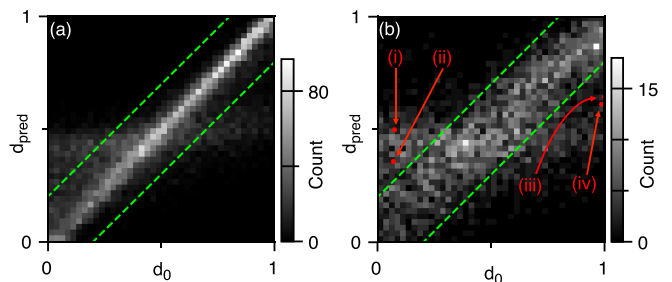


FIG. 2. Performance of the CNN. (a),(b) The statistics of the training outcome among the training set (a) and the testing set (b). The dashed green lines denote a prediction error of 20% marking our threshold for good guesses. The red arrows in (b) denote four representative cases [(i)–(iv)] among the faint band of wrong predictions.

to positive values $\bar{y} > 0$, and we train the CNN by matching \bar{y} to some smooth distribution function $\vec{f}(d) > 0$ centered at d by using the loss function \mathcal{L} defined in Eq. (3), where d_0 is the label. When making the prediction, the value of d_{pred} is determined by an argument minimization searching for the best fit between $\vec{f}(d)$ and \bar{y} . The prime in Eq. (4) denotes that the search for d_{pred} involves only certain (n) output nodes centered at $\arg \max(\bar{y})$. This distribution matching method allows not only for a bounded, continuous prediction from a discrete output, but it also enables the possibility to find the confidence of the prediction $\xi[\bar{y}, d_{\text{pred}}(\bar{y})]$ using the redundant output information, as will be discussed later.

In practice, we used Gaussian $f_i = e^C e^{-\frac{1}{2}(\frac{i-d}{\sigma})^2}$ for $\vec{f}(d)$, where i is the index of the output layer normalized within $[0,1]$, and ($C = 20, \sigma = 0.1, n = 5$) are chosen for the coefficients. This makes Eq. (4) equivalent to the search for a best-fit parabola based on five given points [Fig. 1(d), inset], for which the formalism is straightforward (see Appendix A). In principle, one could train a deep neural network to replace Eq. (4) and the model would become a standard regression [74]. To show this, we trained an alternative model by first feeding FCL2 to another fully connected layer, FCL3, with 501 nodes, from which the prediction is directly extracted, as illustrated by the lower path in Fig. 1(c). The convergence paths of the two models are compared in Fig. 1(d), where the standard regression model converges more slowly and exhibits more significant fluctuation of the error.

IV. RESULTS

To train the model, we generated 20 000 families of FORCs with recorded simulation parameters. Several examples of FORCs are shown in Appendix B. For the training, we used 80% of the data separated into 32 batches, which were shuffled after all batches are accessed during each training epoch. The rest of the data were held out as a test set, which was used to examine the performance after each epoch. The training continued until the best performance was not superseded for 200 epochs straight. After training, the average absolute error was ~ 0.096 and ~ 0.155 for the training and the testing sets, respectively. The statistics for these outcomes are shown in Figs. 2(a) and 2(b), where the bright major diagonal lines in both results correspond to accurate predictions. This indicates

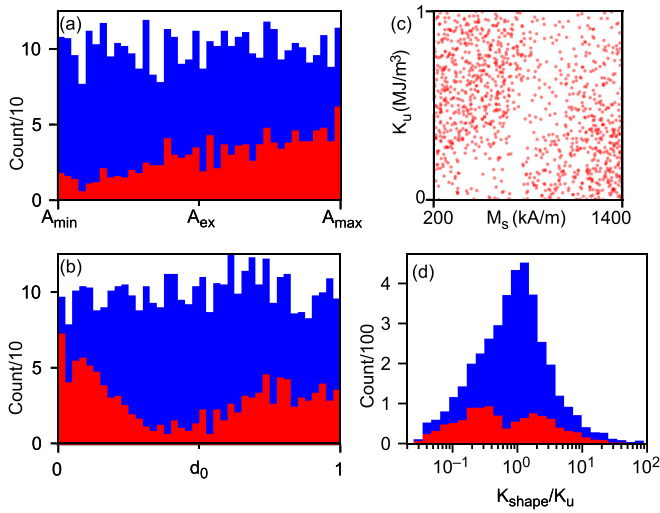


FIG. 3. The statistics of the predictions. The distributions of A_{ex} and d_0 among the testing data are shown in (a) and (b), respectively. The red histogram denotes the predictions with errors greater than 20%, whereas the blue ones illustrate the entire test set. (c) The distribution of M_s and K_u among the wrong predictions. (d) The distribution of $\frac{K_{\text{shape}}}{K_u}$ among the wrong predictions (red) and the entire test set (blue).

that the CNN can indeed extract the magnitude of DMI from FORCs well beyond random guessing. Further training can indeed improve the performance for the training set, but not for the test set, consistent with the trend expected for overfitting. In addition to the main diagonal lines, a faint, nearly flat background band also shows up in both results, suggesting that the CNN is unable to make the correct predictions in certain cases.

To understand the origin of the faint nearly flat band in Figs. 2(a) and 2(b), we compare the simulation parameters for the group of FORCs with “wrong” predictions to those for the entire data set. Here, we define “wrong” predictions to be those where $|d_{\text{pred}} - d_0| \geq 0.2$ [i.e., those points not between the green dashed lines in Fig. 2(b)]. Since the parameters are uniformly random, the test set is well balanced for all simulation parameters. This can be seen from the histograms (blue) of A_{ex} and d_0 in Figs. 3(a) and 3(b), respectively. However, among the wrong predictions (red), a skew towards larger values of A_{ex} can be clearly seen. This is intuitive considering that the Heisenberg exchange is the symmetric part of the exchange coupling that directly competes with the antisymmetric DMI. On the other hand, the histogram of d_0 has a dip near the center and rises near the edges, which is consistent with the faint flat band of wrong predictions in Figs. 2(a) and 2(b).

The main reason of the prediction uncertainty can be uncovered by comparing the statistics of M_s and K_u for the wrong predictions among the test set. As one can see in Fig. 3(c), the wrong predictions are heavily populated where either K_u is large and M_s is small or vice versa. This suggests that the competition between the easy-axis and the shape anisotropy plays an essential role. When magnetized uniformly, the demagnetization field contributed by the long-range dipolar interaction can be effectively seen as an easy-plane shape anisotropy with

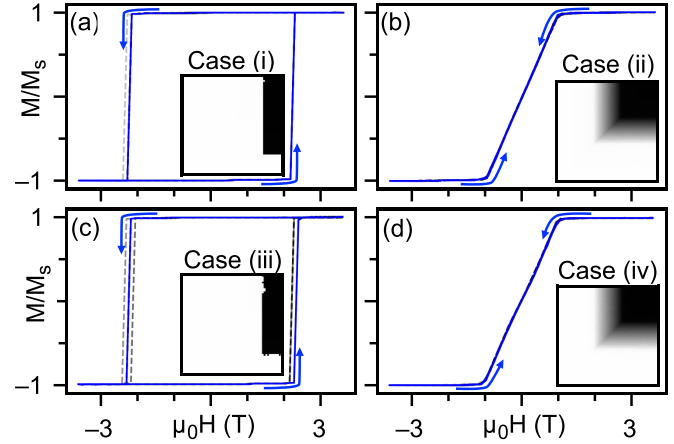


FIG. 4. Example inputs corresponding to uncertain predictions. (a)–(d) The hysteresis loops for Cases (i)–(iv) highlighted in Fig. 2(b). The inset in each case illustrates the corresponding input image fed to the CNN.

$K_{\text{shape}} = \frac{1}{2}\mu_0 M_s^2$. This results in a perpendicular hard axis that competes with the easy axis intrinsically hosted by the material. When these two terms almost cancel in the Hamiltonian, for each site the $\text{SO}(3)$ rotation symmetry is restored in the case of uniform spin. The average magnetization thus depends more on the exchange interactions between neighboring sites. This makes the competition between A_{ex} and D more pronounced, which makes it easier for the CNN to recognize the features of the DMI. Figure 3(d) shows histograms of K_{shape}/K_u for the full test set (blue) and the erroneous portion (red), which demonstrates that the CNN has more difficulty when either anisotropy dominates. This difficulty originates from the fact that the FORCs are featureless in either case. This can be seen in four representative FORCs [Figs. 4(a)–4(d)] corresponding to the wrong predictions highlighted in Fig. 2(b) as cases (i)–(iv). When K_u is very large [cases (i) and (iii)], the FORCs switch abruptly such that almost no minor loops show up within our 61-step resolution of H_R . On the other hand, when K_{shape} dominates [Cases (ii) and (iv)], the switching is gradual but the hysteresis vanishes, leaving a featureless curve that carries no information of the detailed spin texture. These FORCs remain featureless for both large and small values of D , suggesting that the DMI information is indeed overwhelmed.

The faint flat band formed by the wrong predictions in Figs. 2(a) and 2(b) can be understood by checking the learning strategy during training. At different stages of training, the statistics on the testing set are recorded, among which three representative snapshots are shown in Figs. 5(a)–5(c). It seems that the CNN first learns to guess around a single number because even this naive strategy reduces the average loss from random guesses by a factor of 5. After this, it starts to recognize the patterns associating the FORCs with the DMI and gradually forgets the naive strategy. As training continues, the CNN starts to make more accurate predictions for nearly all FORCs with better pronounced features, leaving the featureless FORCs in the faint band of wrong guesses due to the memory of the naive strategy. Note that further training can indeed remove this memory for the training data,

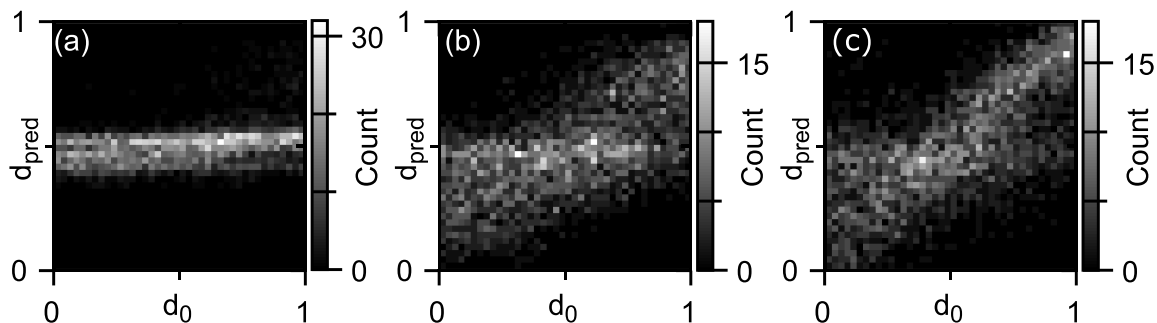


FIG. 5. Evolution of CNN performance. (a)–(c) The evolution of the CNN performance during training. The three panels (a)–(c) correspond to Epochs 3, 50, and 385, respectively.

whereas the performance on the testing data does not improve, suggesting overfitting. Videos showing the evolution of the CNN performance can be found in [75].

Our CNN is different from a standard regression model since it interprets the prediction using the distribution matching method mentioned above. The redundant output information can be used to identify a function $\xi[\bar{y}, d_{\text{pred}}(\bar{y})]$ that extracts the confidence of each prediction without knowing the correct answer d_0 . Although classification models naturally have this capability, the confidence interpretation for a typical regression model is nontrivial [76].

In our particular case, we empirically observed that the average log distance between the proper distribution and the output ($\Delta x = \langle \ln[\hat{f}(d_{\text{pred}})] - \ln \bar{y} \rangle$) is strongly correlated with d_{pred} . This can be seen in Fig. 6(a), where the variance of d_0 is illustrated for different combinations of Δx and d_{pred} . We

can then identify a linear function that separates the confident and uncertain predictions, $\Delta x = \frac{d_{\text{pred}}}{0.2} + 0.9$, as shown by the red line in Fig. 6(a). To further examine the validity of such observations, we illustrate the relation between Δx and the absolute error in Fig. 6(b), where the confident predictions are mainly contributed by those with errors less than 20%. Note that we do have correct predictions above the dark line in Fig. 6(b), which are essentially lucky guesses since the corresponding variance is large, as shown in Fig. 6(a). Finally, we define the overall metric of confidence within (0,1) by $\xi[\bar{y}, d_{\text{pred}}(\bar{y})] = \sigma[-\alpha(\Delta x - \frac{d_{\text{pred}}}{0.2} - 0.9)]$, which is a sigmoid function centered at the red line in Fig. 6(a), with the tolerance modulated by α . As shown in Figs. 6(c)–6(e), when modulating α one can eventually separate the confident predictions from the entire test set. This is not possible with a conventional regression model, where the only output information is the prediction. Note that $\xi[\bar{y}, d_{\text{pred}}(\bar{y})]$ is not unique, and the boundary of confident predictions is not necessarily a linear function. In principle, one can build another neural network to find a $\xi[\bar{y}, d_{\text{pred}}(\bar{y})]$ that performs better, which is beyond the scope of this work.

V. CONCLUSION

To conclude, we demonstrate that the magnitude of the DMI is indeed contained in the hysteretic magnetometry data. Without any information of spin-texture details, a CNN can recognize the DMI magnitude from feature-rich FORCs with an error of $\sim 15\%$. The prediction is particularly confident when the intrinsic perpendicular easy-axis anisotropy is comparable to the shape anisotropy, which can be fine-tuned in experiments. This brings about the possibility to evaluate the magnitude of DMI without any spin-texture characterization. Future directions to improve accuracy in applications to experimental data include the generation of higher-resolution simulation data within parameter ranges of particular experimental interest, as well as the injection of noise into the training data. Our results also suggest that there may exist a model that relates the DMI magnitude with the ensemble-averaged magnetization, which invites further theoretical investigations.

The database ‘‘HoyaFORCs’’ and the example code to train the CNN are available in Ref. [77].

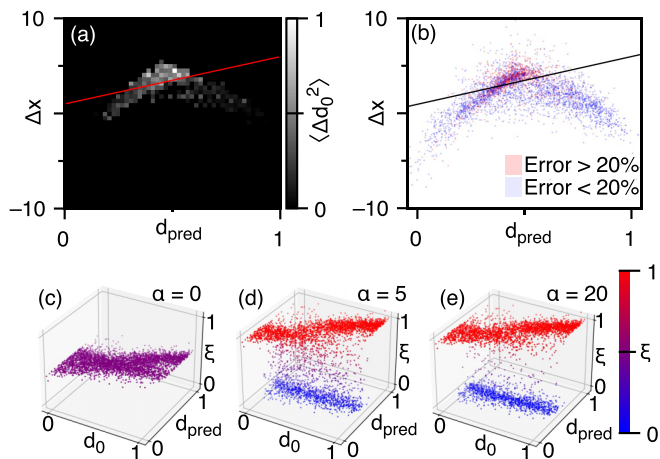


FIG. 6. Confidence metric established from the distribution-matching method. (a) The distribution of the variance of d_0 among the test set illustrated as a function of Δx and d_{pred} , where the red solid line denotes our selected boundary separating confident and uncertain predictions. The statistics within each pixel are only shown if the sample size is more than seven points. (b) The absolute error distribution of all predictions made among the test set. The dark line in (b) is the same as the red solid line in (a). (c)–(e) The separation between confident and uncertain predictions by modulating the threshold coefficient α .

ACKNOWLEDGMENTS

This work has been supported by the U.S. National Science Foundation Grants No. DMR-1950502 (B.F., A.Y.L.), No. DMR-2005108 (Z.C., D.B., K.L.), and No. ECCS-2151809 (G.Y., K.L.).

APPENDIX A: DISTRIBUTION-MATCHING METHOD

Our choice of target distribution function is $f_i = e^C e^{-\frac{1}{2}(\frac{t-d}{\sigma})^2}$, where i is the index of the output node normalized within $[0,1]$. As discussed in the main text, the prediction of the model is given by fitting a parabola to five points using the least-squares method. The choice of five points is according to our optimization balancing between estimation quality and computational efficiency. The chosen five points are centered around the maximum value of the output, as shown in Fig. 7(a). The loss function is defined as the fitting error averaged among all the discrete points, which is effectively the sum of the gray areas denoted in Fig. 7(a). Distinct from a typical classification problem, the distribution-matching method allows the prediction of continuous values within a specified range, using the output values among a discrete set of output nodes (in our case 11). The best-fit parabola $y = ax^2 + bx + c$ can be given by Eq. (A1), where the upper bar denotes the average among the discrete fitting points $\{x_i, y_i\}$. The corresponding prediction is given by Eq. (A2). When the maximum output value occurs near a boundary of the range, the five points closest to the boundary, within the range, are selected, as shown in Fig. 7(b),

$$\begin{bmatrix} \overline{x^2} & \overline{x} & 1 \\ \overline{x^3} & \overline{x^2} & \overline{x} \\ \overline{x^4} & \overline{x^3} & \overline{x^2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \overline{y} \\ \overline{xy} \\ \overline{xy^2} \end{bmatrix}, \quad (A1)$$

$$d_{\text{pred}} = -\frac{b}{2a}. \quad (A2)$$

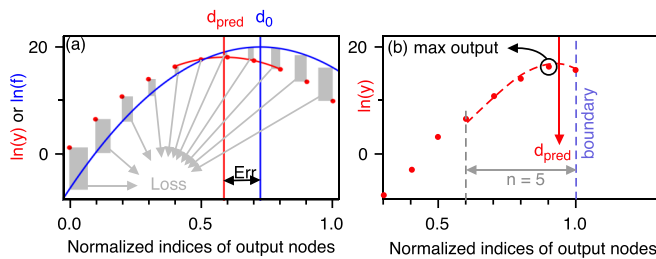


FIG. 7. Method to extract d_{pred} . (a) An example of the distribution-matching method. Only five output points near the maximum output node are selected to make the prediction. The loss function is defined by the average fitting error (gray area) between the target parabola (blue) and the output points (red). The prediction error is the distance between the blue and the red vertical lines. (b) Treatment for special cases when the maximum output node is near the boundary.

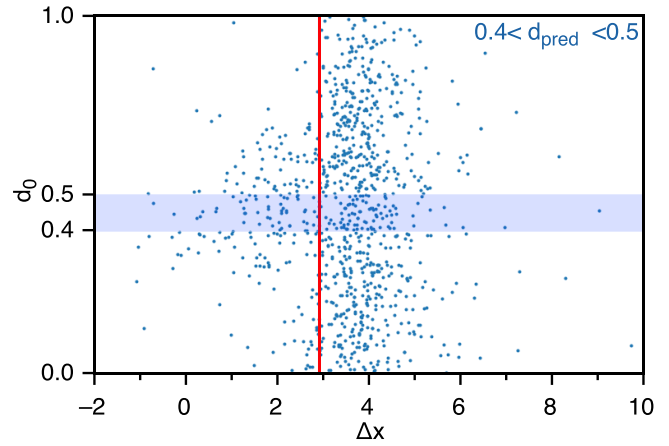


FIG. 8. Separation between confident and uncertain predictions. The values of labels d_0 corresponding to predictions within $0.4 < d_{\text{pred}} < 0.5$ are denoted by the light-blue points. The most accurate predictions are within the horizontal transparent bar denoting $0.4 < d_0 < 0.5$. The red line corresponds to our choice of separation, which leads to the choice of the solid red line in Fig. 5(a) (main text).

The extra information provided by the output layer is used to extract the confidence of the prediction without knowing the correct values of d_0 . As discussed in the main text, we observed that the quantity $\Delta x = \langle \ln [\hat{f}(d_{\text{pred}})] - \ln \bar{y} \rangle$ is highly associated with the variance of d_0 given a range of d_{pred} . As shown in Fig. 8, the distribution of d_0 within the range $0.4 < d_{\text{pred}} < 0.5$ is almost uniformly random for large values of Δx , whereas more accurate predictions can be found on the left. There seems to be a clear boundary between the two regions, which is denoted by the red solid line. The range of d_{pred} shown in Fig. 8 corresponds to the intersection between the main diagonal line and the faint band of wrong predictions shown in Figs. 2(a) and 2(b) in the main text. We carried out a similar analysis for all other ranges of d_{pred} , resulting in the choice of the overall separation as $\Delta x = \frac{d_{\text{pred}}}{0.2} + 0.9$. Finally, we define the overall metric of confidence within $(0,1)$ by $\xi[\bar{y}, d_{\text{pred}}(\bar{y})] = \sigma[-\alpha(\Delta x - \frac{d_{\text{pred}}}{0.2} - 0.9)]$, where σ is a sigmoid function and α is a parameter controlling the separation sharpness. As shown in Movie S3 [75], increasing the value of α gradually separates confident predictions away from the uncertain ones.

APPENDIX B: THE DATABASE HOYAFORCS

The data used to train and test the CNN contain 20 000 families of FORCs with different simulation parameters evenly distributed within the bounds listed in the main text. Six examples of the simulated magnetometry data and their corresponding input images are illustrated in Fig. 9. Note that we have identified 559 samples (out of 20 000) that do not reverse magnetization within the range of $[-3.6 \text{ T}, +3.6 \text{ T}]$.

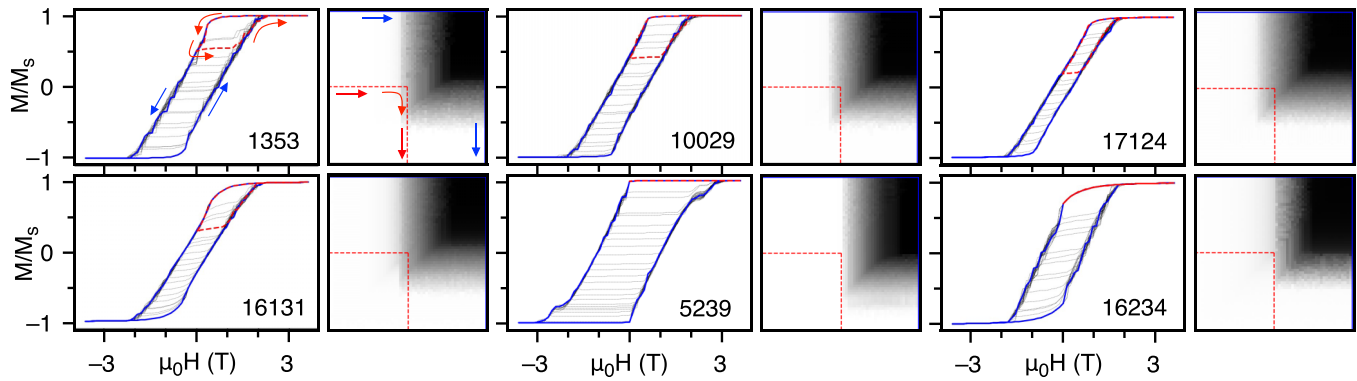


FIG. 9. Examples of items included in HoyaFORCs. Six examples of magnetometry data (left panels) and their corresponding input images (right panels) are demonstrated here. The blue curves are the full hysteresis loops, whereas the dashed gray lines show the minor loops. One example of the minor loops is highlighted by the red dashed line in each panel. The pixels in the input images corresponding to the full loops and the highlighted minor loops are also denoted by the solid blue and dashed red lines, respectively. The number in each panel is the ID of the item in the database.

In the spirit of taking a hands-off approach in the construction of the data set and avoiding post-simulation processing

to select data, we decided to keep these samples in the data set.

- [1] I. Dzyaloshinsky, A thermodynamic theory of “weak” ferromagnetism of antiferromagnetics, *J. Phys. Chem. Solids* **4**, 241 (1958).
- [2] T. Moriya, Anisotropic Superexchange Interaction and Weak Ferromagnetism, *Phys. Rev.* **120**, 91 (1960).
- [3] S. Mühlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, and P. Böni, Skyrmion lattice in a chiral magnet, *Science* **323**, 915 (2009).
- [4] 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).
- [5] X. Z. Yu, N. Kanazawa, Y. Onose, K. Kimoto, W. Z. Zhang, S. Ishiwata, Y. Matsui, and Y. Tokura, Near room-temperature formation of a skyrmion crystal in thin-films of the helimagnet FeGe, *Nat. Mater.* **10**, 106 (2011).
- [6] C. Pappas, E. Lelièvre-Berna, P. Falus, P. M. Bentley, E. Moskvina, S. Grigoriev, P. Fouquet, and B. Farago, Chiral paramagnetic Skyrmion-like phase in MnSi, *Phys. Rev. Lett.* **102**, 197202 (2009).
- [7] X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, and Y. Tokura, Real-space observation of a two-dimensional skyrmion crystal, *Nature (London)* **465**, 901 (2010).
- [8] S. X. Huang and C. L. Chien, Extended Skyrmion phase in epitaxial FeGe (111) thin films, *Phys. Rev. Lett.* **108**, 267201 (2012).
- [9] A. Crépieux and C. Lacroix, Dzyaloshinsky–Moriya interactions induced by symmetry breaking at a surface, *J. Magn. Magn. Mater.* **182**, 341 (1998).
- [10] N. Romming, C. Hanneken, M. Menzel, J. E. Bickel, B. Wolter, K. v. Bergmann, A. Kubetzka, and R. Wiesendanger, Writing and deleting single magnetic skyrmions, *Science* **341**, 636 (2013).
- [11] A. Fert, V. Cros, and J. Sampaio, Skyrmions on the track, *Nat. Nanotechnol.* **8**, 152 (2013).
- [12] S. Heinze, K. v. Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, and S. Blügel, Spontaneous atomic-scale magnetic Skyrmion lattice in two dimensions, *Nat. Phys.* **7**, 713 (2011).
- [13] P. Ferriani, K. von Bergmann, E. Y. Vedmedenko, S. Heinze, M. Bode, M. Heide, G. Bihlmayer, S. Blügel, and R. Wiesendanger, Atomic-scale spin spiral with a unique rotational sense: Mn monolayer on W(001), *Phys. Rev. Lett.* **101**, 027201 (2008).
- [14] W. Jiang, P. Upadhyaya, W. Zhang, G. Yu, M. B. Jungfleisch, F. Y. Fradin, J. E. Pearson, Y. Tserkovnyak, K. L. Wang, O. Heinonen, S. G. E. t. Velthuis, and A. Hoffmann, Blowing magnetic skyrmion bubbles, *Science* **349**, 283 (2015).
- [15] G. Siracusano, R. Tomasello, A. Giordano, V. Puliafito, B. Azzerboni, O. Ozatay, M. Carpentieri, and G. Finocchio, Magnetic radial vortex stabilization and efficient manipulation driven by the Dzyaloshinskii–Moriya interaction and spin-transfer torque, *Phys. Rev. Lett.* **117**, 087204 (2016).
- [16] X. Li, L. Shen, Y. Bai, J. Wang, X. Zhang, J. Xia, M. Ezawa, O. A. Tretiakov, X. Xu, M. Mruczkiewicz, M. Krawczyk, Y. Xu, R. F. L. Evans, R. W. Chantrell, and Y. Zhou, Bimeron clusters in chiral antiferromagnets, *npj Comput. Mater.* **6**, 169 (2020).
- [17] Y. Fujishiro, N. Kanazawa, T. Nakajima, X. Z. Yu, K. Ohishi, Y. Kawamura, K. Kakurai, T. Arima, H. Mitamura, A. Miyake, K. Akiba, M. Tokunaga, A. Matsuo, K. Kindo, T. Koretsune, R. Arita, and Y. Tokura, Topological transitions among skyrmion- and hedgehog-lattice states in cubic chiral magnets, *Nat. Commun.* **10**, 1059 (2019).
- [18] T. Tanigaki, K. Shibata, N. Kanazawa, X. Yu, Y. Onose, H. S. Park, D. Shindo, and Y. Tokura, Real-Space observation of short-period cubic lattice of skyrmions in MnGe, *Nano Lett.* **15**, 5438 (2015).
- [19] S. Okumura, S. Hayami, Y. Kato, and Y. Motome, Magnetic hedgehog lattices in noncentrosymmetric metals, *Phys. Rev. B* **101**, 144416 (2020).

- [20] N. Kent, N. Reynolds, D. Raftrey, I. T. G. Campbell, S. Virasawmy, S. Dhuey, R. V. Chopdekar, A. Hierro-Rodriguez, A. Sorrentino, E. Pereiro, S. Ferrer, F. Hellman, P. Sutcliffe, and P. Fischer, Creation and observation of Hopfions in magnetic multilayer systems, *Nat. Commun.* **12**, 1562 (2021).
- [21] G. Chen, J. Zhu, A. Quesada, J. Li, A. T. N'Diaye, Y. Huo, T. P. Ma, Y. Chen, H. Y. Kwon, C. Won, Z. Q. Qiu, A. K. Schmid, and Y. Z. Wu, Novel chiral magnetic domain wall structure in Fe/Ni/Cu (001) films, *Phys. Rev. Lett.* **110**, 177204 (2013).
- [22] K.-S. Ryu, L. Thomas, S.-H. Yang, and S. Parkin, Chiral spin torque at magnetic domain walls, *Nat. Nanotechnol.* **8**, 527 (2013).
- [23] J.-S. B. Tai and I. I. Smalyukh, Static Hopf solitons and knotted emergent fields in solid-state noncentrosymmetric magnetic nanostructures, *Phys. Rev. Lett.* **121**, 187201 (2018).
- [24] Y. Liu, W. Hou, X. Han, and J. Zang, Three-dimensional dynamics of a magnetic hopfion driven by spin transfer torque, *Phys. Rev. Lett.* **124**, 127204 (2020).
- [25] J. Zou, S. Zhang, and Y. Tserkovnyak, Topological transport of deconfined hedgehogs in magnets, *Phys. Rev. Lett.* **125**, 267201 (2020).
- [26] R. L. Melcher, Linear contribution to spatial dispersion in the spin-wave spectrum of ferromagnets, *Phys. Rev. Lett.* **30**, 125 (1973).
- [27] L. Udvardi and L. Szunyogh, Chiral Asymmetry of the spin-wave spectra in ultrathin magnetic films, *Phys. Rev. Lett.* **102**, 207204 (2009).
- [28] A. T. Costa, R. B. Muniz, S. Lounis, A. B. Klautau, and D. L. Mills, Spin-orbit coupling and spin waves in ultrathin ferromagnets: The spin-wave Rashba effect, *Phys. Rev. B* **82**, 014428 (2010).
- [29] M. Kuepferling, A. Casiraghi, G. Soares, G. Durin, F. Garcia-Sanchez, L. Chen, C. H. Back, C. H. Marrows, S. Tacchi, and G. Carlotti, Measuring interfacial Dzyaloshinskii-Moriya interaction in ultra-thin magnetic films, *Rev. Mod. Phys.* **95**, 015003 (2023).
- [30] G. Yu, P. Upadhyaya, Q. Shao, H. Wu, G. Yin, X. Li, C. He, W. Jiang, X. Han, P. K. Amiri, and K. L. Wang, Room-Temperature skyrmion shift device for memory application, *Nano Lett.* **17**, 261 (2017).
- [31] X. Ma, G. Yu, X. Li, T. Wang, D. Wu, K. S. Olsson, Z. Chu, K. An, J. Q. Xiao, K. L. Wang, and X. Li, Interfacial control of Dzyaloshinskii-Moriya interaction in heavy metal/ferromagnetic metal thin film heterostructures, *Phys. Rev. B* **94**, 180408(R) (2016).
- [32] G. Chen, A. Mascaraque, H. Y. Jia, B. Zimmermann, M. Robertson, R. L. Conte, M. Hoffmann, M. A. G. Barrio, H. F. Ding, R. Wiesendanger, E. Michel, S. Blügel, A. Schmid, and K. Liu, Large Dzyaloshinskii-Moriya interaction induced by chemisorbed oxygen on a ferromagnet surface, *Sci. Adv.* **6**, eaba4924 (2020).
- [33] G. Chen, M. C. Robertson, M. Hoffmann, C. Ophus, A. L. FernandesCauduro, R. LoConte, H. Ding, R. Wiesendanger, S. Blügel, A. e. Schmid, and K. Liu, Observation of hydrogen-induced Dzyaloshinskii-Moriya interaction and reversible switching of magnetic chirality, *Phys. Rev. X* **11**, 021015 (2021).
- [34] L. HerreraDiez, Y. T. Liu, D. A. Gilbert, M. Belmeguenai, J. Vogel, S. Pizzini, E. Martinez, A. Lamperti, J. B. Mohammedi, A. Laborieux, Y. Roussigné, A. J. Grutter, E. Arenholtz, P. Quarterman, B. Maranville, S. Ono, M. S. E. Hadri, R. Tolley, E. E. Fullerton, L. Sanchez-Tejerina *et al.*, Nonvolatile ionic modification of the Dzyaloshinskii-Moriya interaction, *Phys. Rev. Appl.* **12**, 034005 (2019).
- [35] M. Bode, M. Heide, K. v. Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel, and R. Wiesendanger, Chiral magnetic order at surfaces driven by inversion asymmetry, *Nature (London)* **447**, 190 (2007).
- [36] M. Heide, G. Bihlmayer, and S. Blügel, Dzyaloshinskii-Moriya interaction accounting for the orientation of magnetic domains in ultrathin films: Fe/W(110), *Phys. Rev. B* **78**, 140403(R) (2008).
- [37] J. Torrejon, J. Kim, J. Sinha, S. Mitani, M. Hayashi, M. Yamanouchi, and H. Ohno, Interface control of the magnetic chirality in CoFeB/MgO heterostructures with heavy-metal underlayers, *Nat. Commun.* **5**, 4655 (2014).
- [38] K. Zakeri, Y. Zhang, J. Prokop, T.-H. Chuang, N. Sakr, W. X. Tang, and J. Kirschner, Asymmetric spin-wave dispersion on Fe(110): Direct evidence of the Dzyaloshinskii-Moriya interaction, *Phys. Rev. Lett.* **104**, 137203 (2010).
- [39] X. Ma, G. Yu, S. A. Razavi, S. S. Sasaki, X. Li, K. Hao, S. H. Tolbert, K. L. Wang, and X. Li, Dzyaloshinskii-Moriya interaction across an antiferromagnet-ferromagnet interface, *Phys. Rev. Lett.* **119**, 027202 (2017).
- [40] R. Soucaille, M. Belmeguenai, J. Torrejon, J.-V. Kim, T. Devolder, Y. Roussigné, S.-M. Chérif, A. A. Stashkevich, M. Hayashi, and J.-P. Adam, Probing the Dzyaloshinskii-Moriya interaction in CoFeB ultrathin films using domain wall creep and Brillouin light spectroscopy, *Phys. Rev. B* **94**, 104431 (2016).
- [41] S. Fukami, C. Zhang, S. DuttaGupta, A. Kurenkov, and H. Ohno, Magnetization switching by spin-orbit torque in an antiferromagnet-ferromagnet bilayer system, *Nat. Mater.* **15**, 535 (2016).
- [42] C.-F. Pai, M. Mann, A. J. Tan, and G. S. D. Beach, Determination of spin torque efficiencies in heterostructures with perpendicular magnetic anisotropy, *Phys. Rev. B* **93**, 144409 (2016).
- [43] S. Ding, A. Ross, R. Lebrun, S. Becker, K. Lee, I. Boventer, S. Das, Y. Kurokawa, S. Gupta, J. Yang, G. Jakob, and M. Kläui, Interfacial Dzyaloshinskii-Moriya interaction and chiral magnetic textures in a ferrimagnetic insulator, *Phys. Rev. B* **100**, 100406(R) (2019).
- [44] J. Han, A. Richardella, S. A. Siddiqui, J. Finley, N. Samarth, and L. Liu, Room-temperature spin-orbit torque switching induced by a topological insulator, *Phys. Rev. Lett.* **119**, 077702 (2017).
- [45] M. Kawaguchi, K. Tanabe, K. Yamada, T. Sawa, S. Hasegawa, M. Hayashi, and Y. Nakatani, Determination of the Dzyaloshinskii-Moriya interaction using pattern recognition and machine learning, *npj Comput. Mater.* **7**, 20 (2021).
- [46] V. K. Singh and J. H. Han, Application of machine learning to two-dimensional Dzyaloshinskii-Moriya ferromagnets, *Phys. Rev. B* **99**, 174426 (2019).
- [47] H. Y. Kwon, H. G. Yoon, C. Lee, G. Chen, K. Liu, A. K. Schmid, Y. Z. Wu, J. W. Choi, and C. Won, Magnetic Hamiltonian parameter estimation using deep learning techniques, *Sci. Adv.* **6**, eabb0872 (2020).
- [48] D. Wang, S. Wei, A. Yuan, F. Tian, K. Cao, Q. Zhao, Y. Zhang, C. Zhou, X. Song, D. Xue, and S. Yang, Machine learning mag-

- netic parameters from spin configurations, *Adv. Sci.* **7**, 2000566 (2020).
- [49] C. R. Pike, A. P. Roberts, and K. L. Verosub, Characterizing interactions in fine magnetic particle systems using first order reversal curves, *J. Appl. Phys.* **85**, 6660 (1999).
- [50] J. E. Davies, O. Hellwig, E. E. Fullerton, G. Denbeaux, J. B. Kortright, and K. Liu, Magnetization reversal of Co/Pt: Microscopic origin of high-field magnetic irreversibility, *Phys. Rev. B* **70**, 224434 (2004).
- [51] D. A. Gilbert, G. T. Zimanyi, R. K. Dumas, M. Winklhofer, A. Gomez, N. Eibagi, J. L. Vicent, and K. Liu, Quantitative decoding of interactions in tunable nanomagnet arrays using first order reversal curves, *Sci. Rep.* **4**, 4204 (2014).
- [52] A. P. Roberts, C. R. Pike, and K. L. Verosub, First-order reversal curve diagrams: A new tool for characterizing the magnetic properties of natural samples, *J. Geophys. Res.: Solid Earth* **105**, 28461 (2000).
- [53] C.-I. Dobrotă and A. Stancu, What does a first-order reversal curve diagram really mean? A study case: Array of ferromagnetic nanowires, *J. Appl. Phys.* **113**, 043928 (2013).
- [54] M. T. Rahman, R. K. Dumas, N. Eibagi, N. N. Shams, Y.-C. Wu, K. Liu, and C.-H. Lai, Controlling magnetization reversal in Co/Pt nanostructures with perpendicular anisotropy, *Appl. Phys. Lett.* **94**, 042507 (2009).
- [55] J. A. D. Toro, M. Vasilakaki, S. S. Lee, M. S. Andersson, P. S. Normile, N. Yaacoub, P. Murray, E. H. Sánchez, P. Muñoz, D. Peddis, R. Mathieu, K. Liu, J. Geshev, K. N. Trohidou, and J. Nogués, Remanence plots as a probe of spin disorder in magnetic nanoparticles, *Chem. Mater.* **29**, 8258 (2017).
- [56] R. K. Dumas, C.-P. Li, I. V. Roshchin, I. K. Schuller, and K. Liu, Magnetic fingerprints of sub-100 nm Fe dots, *Phys. Rev. B* **75**, 134405 (2007).
- [57] D. A. Gilbert, B. B. Maranville, A. L. Balk, B. J. Kirby, P. Fischer, D. T. Pierce, J. Unguris, J. A. Borchers, and K. Liu, Realization of ground-state artificial skyrmion lattices at room temperature, *Nat. Commun.* **6**, 8462 (2015).
- [58] E. C. Burks, D. A. Gilbert, P. D. Murray, C. Flores, T. E. Felter, S. Charnvanichborikarn, S. O. Kucheyev, J. D. Colvin, G. Yin, and K. Liu, 3D nanomagnetism in low density interconnected nanowire networks, *Nano Lett.* **21**, 716 (2021).
- [59] E. D. Biasi, J. Curiale, and R. D. Zysler, Quantitative study of FORC diagrams in thermally corrected Stoner–Wohlfarth nanoparticles systems, *J. Magn. Magn. Mater.* **419**, 580 (2016).
- [60] A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. V. Waeyenberge, The design and verification of MuMax3, *AIP Adv.* **4**, 107133 (2014).
- [61] J. M. Shaw, S. E. Russek, T. Thomson, M. J. Donahue, B. D. Terris, O. Hellwig, E. Dobisz, and M. L. Schneider, Reversal mechanisms in perpendicularly magnetized nanostructures, *Phys. Rev. B* **78**, 024414 (2008).
- [62] O. Boulle, J. Vogel, H. Yang, S. Pizzini, D. d. S. Chaves, A. Locatelli, T. O. Menteş, A. Sala, L. D. Buda-Prejbeanu, O. Klein, M. Belmeguenai, Y. Roussigné, A. Stashkevich, S. M. Chérif, L. Aballe, M. Foerster, M. Chshiev, S. Auffret, I. M. Miron, and G. Gaudin, Room-temperature chiral magnetic skyrmions in ultrathin magnetic nanostructures, *Nat. Nanotechnol.* **11**, 449 (2016).
- [63] J.-H. Shim, A. A. Syed, Y. Shin, J.-W. Kim, H.-G. Piao, S.-H. Lee, K. M. Lee, J.-R. Jeong, D.-H. Kim, and D. E. Kim, Ultrafast dynamics of exchange stiffness in Co/Pt multilayer, *Commun. Phys.* **3**, 74 (2020).
- [64] M. Heigl, S. Koraltan, M. Vaňatka, R. Kraft, C. Abert, C. Vogler, A. Semisalova, P. Che, A. Ullrich, T. Schmidt, J. Hintermayr, D. Grundler, M. Farle, M. Urbánek, D. Suess, and M. Albrecht, Dipolar-stabilized first and second-order antiskyrmions in ferrimagnetic multilayers, *Nat. Commun.* **12**, 2611 (2021).
- [65] R. Moreno, R. F. L. Evans, S. Khmelevskiy, M. C. Muñoz, R. W. Chantrell, and O. Chubykalo-Fesenko, Temperature-dependent exchange stiffness and domain wall width in Co, *Phys. Rev. B* **94**, 104433 (2016).
- [66] C. Eyrich, W. Huttema, M. Arora, E. Montoya, F. Rashidi, C. Burrowes, B. Kardasz, E. Girt, B. Heinrich, O. N. Mryasov, M. From, and O. Karis, Exchange stiffness in thin film Co alloys, *J. Appl. Phys.* **111**, 07C919 (2012).
- [67] S. D. Pollard, J. A. Garlow, J. Yu, Z. Wang, Y. Zhu, and H. Yang, Observation of stable Néel skyrmions in cobalt/palladium multilayers with Lorentz transmission electron microscopy, *Nat. Commun.* **8**, 14761 (2017).
- [68] C. Moreau-Luchaire, C. Moutafis, N. Reyren, J. Sampaio, C. A. F. Vaz, N. V. Horne, K. Bouzehouane, K. Garcia, C. Deranlot, P. Warnicke, P. Wohlhüter, J.-M. George, M. Weigand, J. Raabe, V. Cros, and A. Fert, Additive interfacial chiral interaction in multilayers for stabilization of small individual skyrmions at room temperature, *Nat. Nanotechnol.* **11**, 444 (2016).
- [69] A. Soumyanarayanan, M. Raju, A. L. G. Oyarce, A. K. C. Tan, M.-Y. Im, A. P. Petrović, P. Ho, K. H. Khoo, M. Tran, C. K. Gan, F. Ernult, and C. Panagopoulos, Tunable room-temperature magnetic skyrmions in Ir/Fe/Co/Pt multilayers, *Nat. Mater.* **16**, 898 (2017).
- [70] G. Grinstein and R. H. Koch, Coarse graining in micromagnetics, *Phys. Rev. Lett.* **90**, 207201 (2003).
- [71] M. B. Hahn, Temperature in micromagnetism: Cell size and scaling effects of the stochastic Landau–Lifshitz equation, *J. Phys. Commun.* **3**, 075009 (2019).
- [72] M. D. Zeiler, ADADELTA: An Adaptive Learning Rate Method, [arXiv:1212.5701](https://arxiv.org/abs/1212.5701) [cs].
- [73] K. He, X. Zhang, S. Ren, and J. Sun, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Piscataway, NJ, 2016), pp. 770–778.
- [74] J. Y. Khoo and M. Heyl, Quantum entanglement recognition, *Phys. Rev. Res.* **3**, 033135 (2021).
- [75] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.5.043012> for the videos and their captions.
- [76] R. F. Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, Predictive inference with the jackknife+, *Ann. Stat.* **49**, 486 (2021).
- [77] <https://github.com/bfugetta/HoyaFORCs>