Machine-learning Skyrmions

Vinit Kumar Singh[1, *] and Jung Hoon Han[2, †]

[1]*Department of Physics, Indian Institute of Technology, Kharagpur 721302, India*
[2]*Department of Physics, Sungkyunkwan University, Suwon 16419, Korea*
(Dated: June 14, 2018)

Principles of machine learning (ML) are applied to models that support skyrmion phases in two dimensions. Most successful predictions were found when a convolutional neural network (CNN) layer was inserted as well as several layers of neural networks. A new training scheme based on features of the input configuration such as magnetization and spin chirality is introduced to make reliable predictions on the mixed phases, consisting of either a mixture of spiral and skyrmions or of skyrmions and ferromagnets. It proved possible to further train external parameters such as the external magnetic field and temperature and make reliable predictions on them. The predictive capacity of the ML continued to apply to configurations that are not generated by the original Hamiltonian used in the training stage, but a different Hamiltonian adiabatically connected to the original one.

The basic strategy behind teaching ML algorithm to recognize various phases of many-body systems[1–15], whether classical or quantum, is to train it on many examples of many-body configurations together with answers to the phases to which they belong. After the successful implementation of supervised learning as such, the ML algorithm can predict the phase of a new configuration, not drawn from the previous training set. If the input state is drawn from near the phase transition, the prediction is either one or the other side of the transition with certain probability. The gradual change in the probability with temperature or tuning parameter can be used to locate the critical temperature (classical) or the interaction strength (quantum) of the continuous phase transition point. Most studies in recent years have focused on the transition between ordered and disordered phases separated with a second-order critical point. Following the natural progression in the level of sophistication, models studied with the ML method have evolved from Ising[1–11] to planar (XY)[7,9,12–14], and most recently to Heisenberg[15] spins.

Here we address various aspects of the Heisenberg-Dzyaloshinskii-Moriya-Zeeman (HDMZ) spin Hamiltonian by the ML method:

$$H_{\text{HDMZ}} = -J \sum_{i \in L^2} \mathbf{n}_i \cdot (\mathbf{n}_{i+\hat{x}} + \mathbf{n}_{i+\hat{y}})$$
$$+ D \sum_i (\hat{y} \cdot \mathbf{n}_i \times \mathbf{n}_{i+\hat{x}} - \hat{x} \cdot \mathbf{n}_i \times \mathbf{n}_{i+\hat{y}}) - \mathbf{B} \cdot \sum_i \mathbf{n}_i. \quad (1)$$

This lattice model, usually solved in two-dimensional $L \times L$ square lattice, describes the magnetic interaction at the interface of a magnetic layer with a non-magnetic layer, or a magnetic layer exposed to vacuum. Its phase diagram, by now well-known, includes the skyrmion crystal over some intermediate field range, flanked by spiral phase at low field and ferromagnetic phase at high field[16–20].

Classifying mixed phases: As an initial application of ML ideas to the HDMZ model, we created a training set of configurations drawn from deep inside the spiral, skyrmionic, and ferromagnetic phases of the model (1). Our training set was generated by the Monte Carlo (MC) method with $D/J = \sqrt{6}$, corresponding to the spiral period $\lambda = 6$. (J will be set to unity from now on.) The inter-skyrmion distance in the skyrmion phase is also of the same order. The average spin chirality χ, defined in (2), over the (T, B) plane is presented in Fig. 1(a). Judging from the lack of significant spin chirality, the $B \in [0, 1.2]$ region over the temperature $T \in [0.03, 0.25]$ (box 1 in the chirality map, Fig. 1) can be said to belong to the spiral phase. Likewise, one can say with confidence that $B \in [3.4, 4.2]$ over $T \in [0.03, 0.25]$ belongs to the ferromagnetic phase (box 5 in Fig. 1). Here $T = 0.03$ is the lowest temperature reached in our MC calculation. The robust skyrmion phase can be found at $B \in [1.8, 2.6]$ (box 3). For each B, the temperature interval $T \in [0.03, 2.0]$ was divided into 40 steps using adaptive scheduling, *i.e.* exponentially decaying step size with a decay rate of 0.1. That gave us a total of 20 steps from the interval $T \in [0.03, 0.25]$, and with each step we drew 100 MC configurations. Finally, 17 different magnetic field values were selected, for a total of $20 \times 100 \times 17 = 34,000$ training configurations. (Training with the much larger training set of 330,000 configurations did not change the final results.)

The ML architecture[22] used for the training involved an initial CNN layer of 6×6 filter size (since the skyrmion diameter and period of spirals are both 6) and a second CNN layer with 3×3 filter size both accompanied with Max Pool filters, followed by flatten and two dense neural network layers containing 512, 1024 neurons respectively, which then led to the output layer. Batch normalization and dropout regularization are applied to outputs from each layer. The schematic diagram of the architecture can be found in the Supplementary Material (SM). The final outcome was then compared to one of three values, 0, 1, 2, corresponding to spiral, skyrmion, and ferromagnetic phases, respectively. The training data was initially prepared in terms of spin angles (θ_i, ϕ_i), which gave far poorer results than if the data was prepared in terms of
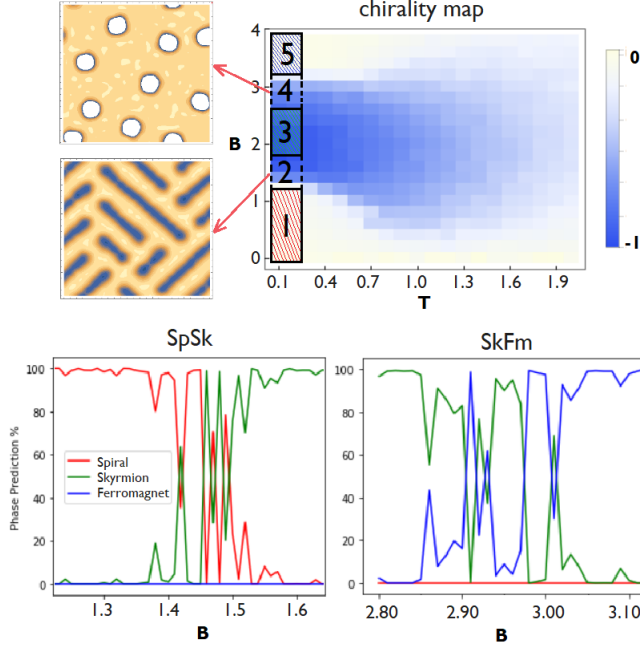
FIG. 1. (top) Spin chirality [χ in Eq. (2)] in the (T, B) plane obtained by MC calculation on the HDMZ Hamiltonian (1). Color scale represents the normalized value of the chirality. Boxes 1, 3, and 5 (2 and 4) represent regions where training (testing) data were taken for label predictions. Two configurations on the left show a typical SpSk and SkFm mixed state, respectively. The z-component of the local magnetization is used for the plots. (bottom) Probability of phase predictions in the SpSk and SkFm phases. The numbers represent averages over the testing set in the temperature interval $T \in [0.03, 0.25]$ at the same B value. The irregularities are not artifacts of the small data size.

the magnetization $\mathbf{n}_i = (\sin\theta_i \cos\phi_i, \sin\theta_i, \sin\phi_i, \cos\theta_i)$. All the ML analysis presented in this paper is thus based on the magnetization inputs. The architecture consisting purely of the deep neural network layer did not work as well as the one we used, involving the CNN filter layer as well. Further minute changes in the architecture had little impact on the overall quality of final results. Only the z-component was used for training in earlier work[15]. After training, the validation procedure gave nearly 100% correct values for the phase labels. This is not surprising given the fact that validation sets also came from deep inside one of the three phases.

It is well-known both experimentally and from simulations that a substantial mixed phase region exists in two-dimensional skyrmion matter[21]. They are mixed spiral and skyrmion (SpSk) regions at low fields, and mixed skyrmion and ferromagnetic (SkFm) regions at higher fields. Because of their presence, a sharp phase boundary separating one phase from another is difficult to define. The next logical step in our investigation is therefore to generate configurations that have such mixed characters, and ask the ML program to predict their phases. For

testing purpose we generate a new batch of MC configurations at $T \in [0.03, 0.25]$ and $B \in [1.2, 1.8]$ (box 2 in Fig. 1), which admittedly belongs to the SpSk mixed phase, and another batch at $T \in [0.03, 0.25]$ and $B \in [2.6, 3.4]$ (box 4 in Fig. 1) belonging to the SkFm mixed phase.

At each (T, B), 100 configurations were generated and fed to the machine for prediction. The answers given by the machine are then averaged over and shown as probabilities for spiral, skymion, and ferromagnetic phase, in Fig. 1 (b) and (c). Since the temperature range was quite small and the variations in the configurations were minor, the answers were averaged over all the 20 temperature steps in the interval $T \in [0.03, 0.25]$ as well. Despite the fact that each data point in Fig. 1(b) and (c) represents an average over $100 \times 20 = 2,000$ configurations, and that extremely fine steps in magnetic field $\Delta B = 0.01$ was used, the final results are far from being smooth. Our results are in stark contrast to earlier attempts on a vast array of models exhibiting second-order phase transition. There, the ML program trained exclusively on the configurations deep inside the ordered and disordered phases could successfully predict the phases near the critical point in a continuously varying manner[2,3,6–8,11–13]. The ML program trained on the three distinct phases of the skyrmion model, on the other hand, fails quite dramatically to make continuously varying predictions for its mixed phases. Rather than declaring it as a failure, we view it as the way the neural network can successfully perceive the first-order phase transition - with a substantial mixed-phase region - differently from the sharp second-order transitions. At the same time, it is a telling suggestion that one must seek other means of characterizing mixed phases. This problem is taken up next.

Feature predictions: The main characteristics of the skyrmion and the ferromagnetic phases are the average spin chirality and the magnetization, respectively, defined as (N=number of lattice sites)

$$\chi = (1/N) \sum_i (\mathbf{n}_i \cdot \mathbf{n}_{i+\hat{x}} \times \mathbf{n}_{i+\hat{y}}),$$
$$m = (1/N) \sum_i n_i^z. \tag{2}$$

The spiral phase is the one where none of these features takes on significant values. Instead of training the ML algorithm on the labels of configurations such as spiral, skyrmion, or ferromagnet, we train it on their features such as χ and m. Once the ML algorithm has been trained to predict those values correctly, the problem of labeling a given configuration is as good as solved. For instance an input configuration whose (χ) m is predicted to be near the the maximum allowed value can have no other label than skyrmion (ferromagnet). Intermediate values of both χ and m signify the SkFm mixed phase. Finally, a configuration with small but non-negligible χ and m are likely associated with the SpSk mixed phase.
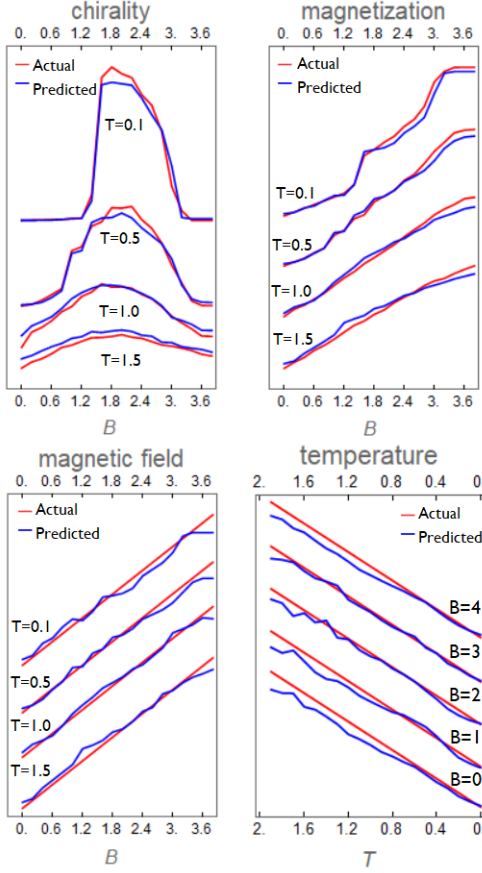
FIG. 2. Machine-predicted values of (χ, m, B, T) in blue curves, compared to their actual values in red. The training and the testing were done on MC configurations generated by the Hamiltonian (1). Different curves are offset for clarity.

The labeling problem is delegated to the human decision, while the machine is left to do its work at predicting quantitative features of the input.

We carried out supervised learning of the (χ, m) features on configurations drawn from a wide temperature range $0.1 \leq T \leq 2.0$ ($\Delta T = 0.1$) and magnetic field $0 \leq B \leq 4.0$ ($\Delta B = 0.2$) corresponding to the entire chirality map in Fig. 1. For each (T, B) we collect 200 MC-annealed configurations for training purpose. The temperature and magnetic field intervals were sufficiently fine, as can be seen from the high quality of the pre-smearing data of χ in Fig. 1. After training with a total of $200 \times 20 \times 21 = 82,000$ configuration, a fresh set of 40,000 configurations were generated to compare the machine-predicted (χ, m) against the actual values. As shown in Fig. 2, the agreement between predicted and actual values is very good across the entire phase diagram.

Encouraged by the success of feature prediction in terms of χ and m, we next ask if the machine can be trained to recognize the particular temperature and magnetic field from which the input configuration

originated. Figure 2 shows predicted values of (T, B) to be very close to the actual values[23]. In contrast to (χ, m) which are the *mechanical variables* of the configuration, (T, B) are *thermodynamic variables*. It turns out that ML works well in predicting both kinds of features.
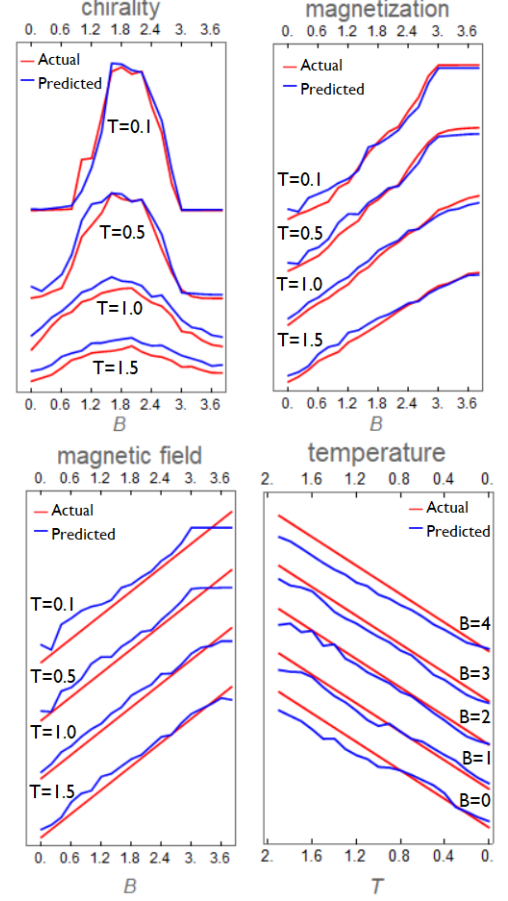

FIG. 3. Machine prediction of (χ, m, B, T) for MC configurations drawn from $H_{\mathrm{HDMZ}} + H_K$ with $K = 1$ and $p = 0.5$. The training itself was done on MC configurations generated by H_{HDMZ} alone.

Prediction for adiabatically connected phases: Various modifications can be added to the HDMZ Hamiltonian (1) to represent the realistic situation of the material. One that reflects the disorder in the material can be given, for instance, by[17–19]

$$H_K = -K \sum_{i \in \mathrm{random}} (S_i^z)^2. \tag{3}$$

The magnetic anisotropy term of strength K is added at the random sites occupying a fraction p of the whole lattice. The model $H(K, p) = H_{\mathrm{HDMZ}} + H_K$ represents an adiabatically connected family of Hamiltonians as long as K is sufficiently small compared to other energy scales. It is interesting to ask whether the ML algorithm, trained
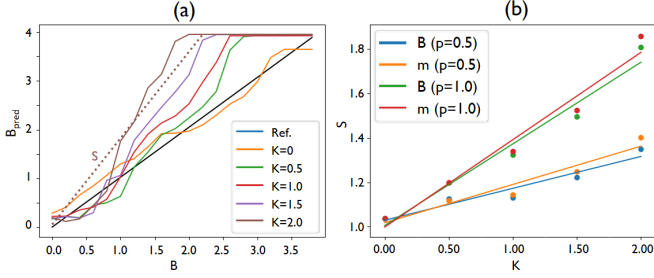
FIG. 4. (a) Predicted magnetic field B_{pred} at $T = 0.1$, for several values of K and $p = 1$. The reference line in black is the actual B. A linear fit (dashed line) to the transient part of the curve gives the slope s. (b) The slope s deduced from the B_{pred}. fit is shown as blue ($p = 0.5$) and green ($p = 1$) dots. They follow linear relationship as shown by lines of the same color. Twice the magnetic susceptibility α deduced from linear fits to m_{pred}, shown in yellow ($p = 0.5$) and magenta ($p = 1$), also gives the similar slope. Slopes for $p = 1$ is larger than the slopes for $p = 0.5$ in roughly obedience of the mean-field relationship $s = 2Kp\alpha$.

solely on configurations drawn from $H(0,0) = H_{\text{HDMZ}}$, can have predictive power over those generated from arbitrary $H(K,p)$. It is also a pragmatic question, when it comes to addressing the machine's predictive power over the experimental data, as real materials are never free of inhomogeneities and one does not have the *a priori* knowledge of the governing Hamiltonian.

A large number of configurations at $K = 1.0$ and $p = 0.5$ was generated by MC and tested by the ML algorithm, previously trained solely on the pristine Hamiltonian H_{HDMZ}. As shown in Fig. 3, very good fits of all features (χ, m, B, T) were obtained. Similar plots for several (K, p) values can be found in SM. The error in the prediction can be quantified by measuring $\Delta X \equiv \sum_{i=T,B} |X_{\text{pred.}} - X_{\text{act.}}|/400$, where 400 refers to the total number of (T, B) steps used in the generation of the test set, and $X = \chi, m, B, T$. Table I shows the mean errors in (χ, m, B, T) for several (K, p) values. Both $\Delta\chi$ and Δm remain less than 0.05 as K grows from 0 to 2 (recall $J = 1$ and $D = \sqrt{6}$). Note that χ and m have the maximum size of 1. On the other hand, there is a systematic growth in ΔB and ΔT as K becomes larger.

We plot predicted values $B_{\text{pred.}}$ against the actual B in Fig. 4(a) for several (K, p)'s. There is an approximate linear relationship in $B_{\text{pred.}}$ against B, at least until $B_{\text{pred.}}$ reaches saturation, with the slope that grows almost linearly with K, as shown in Fig. 4(b). The effect of the added anisotropy can be qualitatively understood within the mean-field picture by replacing $K \sum_{i \in \text{random}} (n_i^z)^2$ with $2Kpm \sum_{i \in L^2} n_i^z$, where p is the impurity fraction. Assuming the magnetization m depending linearly on B, $m = \alpha B$, the overall effect of

the random anisotropy term is to replace the external field B by the effective one, $B_{\text{eff}} = (1 + 2Kp\alpha)B$. The machine, having been trained solely on pristine H_{HDMZ}, knows nothing of the impurity effect *a priori* and "erroneously" predicts the renormalized B_{eff} for the input, thereby incidentally divulging the discrepancy between the training and testing Hamiltonian.

Such expectations are consistent with our numerical analysis of Fig. 4(b), showing almost linear increase in the slope of B_{pred} with K. To further prove this picture we obtain α independently from linear fits to predicted m values such as shown in Fig. 3. The two ways of extracting the susceptibility α agree very well. The ~ 2 times difference in the estimated slopes for $p = 0.5$ and $p = 1$ data are consistent with the mean-field picture of B_{eff}, as shown in Fig. 4(b). The under-estimation of the temperature by the machine, as shown in Fig. 3 and SM figure 2, can be also understood, qualitatively, as a result of K having the tendency to stiffen the spins and align them. At the same bare temperature T, configurations generated at finite K tend to have more alignment of spins, which is "erroneously" seen by the machine to be the consequence of lesser effective temperature $T_{\text{eff}} < T$. Our numerical experiment demonstrates how well the neural network can respond to perturbations in the model[25].

Concluding remarks: Ideas of machine learning have been applied to the skyrmion model with quite accurate predictions for both mechanical and thermodynamic features of the input. Especially the mechanical quantities can be predicted reliably for images produced by different Hamiltonians, indicating that the machine has learned the formulas for calculating them. The architecture design we found optimal for training of the skyrmion matter is a deep layer with CNN (SM figure 1).

(K,p)	$\Delta\chi$	Δm	ΔB	ΔT
$(0,0)$	0.026	0.027	0.032	0.028
$(1,0.5)$	0.030	0.029	0.060	0.037
$(1,1)$	0.038	0.024	0.087	0.066
$(2,0.5)$	0.042	0.025	0.083	0.064
$(2,1)$	0.042	0.024	0.152	0.100

TABLE I: Averaged variance between predicted and actual values of (χ, m, B, T).

[*] Electronic address: vinitsingh911@gmail.com

[†] Electronic address: hanjh@skku.edu

[1] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).

[2] L. Wang, Phys. Rev. B **94**, 195105 (2016).

[3] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).

[4] P. Ponte and R. G. Melko, Phys. Rev. B **96**, 205146 (2017).

[5] A. Morningstar and R. G. Melko, arXiv:1708.04622 (2017).

[6] A. Tanaka and A. Tomiya, J. Phys. Soc. Jpn. **86**, 063001 (2017).

[7] W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017).

[8] S. J. Wetzel and M. Scherzer, Phys. Rev. B **96**, 184410 (2017).

[9] S. J. Wetzel, Phys. Rev. E **96**, 022140 (2017).

[10] S. Iso, S. Shiba, and S. Yokoo, Phys. Rev. E **97**, 053304 (2018).

[11] D. Kim and D.-H. Kim, arXiv:1804.02171v1 (2018).

[12] C. Wang and H. Zhai, Phys. Rev. B **96**, 144432 (2017).

[13] M. J. S. Beach, A. Golubeva, and R. G. Melko, Phys. Rev. B **97**, 045207 (2018).

[14] C. Wang and H. Zhai, arXiv:1803.01205 (2018).

[15] I. A. Iakovlev, O. M. Sotnikov, and V. V. Mazurenko, arXiv:1883.06682v1 (2018).

[16] N. Nagaosa and Y. Tokura, Nature Nanotech. **8**, 899 (2013).

[17] J. P. Liu, Z. Zhang, and G. Zhao, *Skyrmions: topological structures, properties, and applications* (CRC Press, 2016)

[18] W. Jiang, G. Chen, K. Liu, J. Zang, S. G. E. Velthuis, and A. Hoffmann, *Phys. Rep.* **704**,1 (2017).

[19] A. Fert, N. Reyren, and V. Cros, Nature Reviews Materials **2**, 17031 (2017).

[20] J. H. Han, *Skyrmions in Condensed Matter* (Springer, 2017).

[21] X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, and Y. Tokura, Nature (London) **465**, 901 (2010).

[22] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, 2016).

[23] Averaging is done over the MC configurations generated at a fixed (T, B). Shown in the figures 2 and 3 are the averages of the machine-predicted values, and the averages of the actual values. There is a greater degree of fluctuation if the predictions of an individual configuration is compared with the actual value of that configuration.

[24] Supplementary Material

[25] Y. Wu, P. Zhang, H. Shen, and H. Zhai, arXiv:1802.03930 (2018).