Supervised learning magnetic skyrmion phases

I. A. Iakovlev, O. M. Sotnikov, V. V. Mazurenko
Theoretical Physics and Applied Mathematics Department,
Ural Federal University, Mira Street 19,
Ekaterinburg 620002, Russia

(Dated: March 20, 2018)

We propose a simple and transparent machine learning approach for recognition and classification of complex non-collinear magnetic structures in two-dimensional materials. It is based on the implementation of the single-hidden-layer neural network that only relies on the z projections of the spins. In this setup one needs a limited set of magnetic configurations to distiguish ferromagnetic, skyrmion and spin spiral phases, as well as their different combinations. The network trained on the configurations for square-lattice Heisenberg model with Dzyaloshinskii-Moriya interaction can classify the magnetic structures obtained from Monte Carlo calculations for triangular lattice and vice versa. Our approach is also easy to use for analysis of the numerous experimental data collected with spin-polarized scanning tunneling experiments.

Introduction.— A fascinating progress in development of neural-network-based approaches in condensed matter theory allows one to advance the methods for studying of physical properties of materials. For instance, a neural network representation of the quantum Hamiltonain wave function proposed by Carleo and Troyer [1] has revolutionized the field of simulation of complex many-body systems [2, 3]. Within such an approach it becomes possible to model frustrated systems for which existing Quantum Monte Carlo methods fail due to the sign problem. Another remarkable example of the innovations in artificial neural network learning is identification of the magnetic phases of the spin Hamiltonians widely used for description of the strongly correlated materials [4–9]. For instance, in the case of the two-dimensional Ising model the ferromagnetic and paramagnetic phases can be successfully recognized with a single hidden layer network [10]. Importantly, topological phases obtained with a more complex XY Hamiltonian [11] can be also classified with machine learning, however, in this case one needs to design a deep convolutional network and use the system of filters, which makes such an approach similar to the image recognition [12].

Therefore, an important question arises. Is it possible to use the machine learning approach in its simplest and transparent formulation with single hidden layer [10] to explore complex non-collinear magnetic phases of technological importance? In this respect topologically-protected magnetic skyrmions [13–17] and spin spiral states are the first candidates for such a consideration, since they can be used for creating novel magnetic memory devices [18]. Numerous experimental studies revealed skyrmion state in metallic ferromagnets with Dzyaloshinskii-Moriya interaction such as FeGe [19, 20], Fe monolayer on Ir(111) [21], MnGe [22], $Fe_xCo_{1-x}Si$ [23] in a narrow range of the external parameters, magnetic fields and temperatures. The experimental phase diagrams [23] contain significant transitional areas between different phases, which raises the problem of the precise definition of the skyrmion and spin spiral phase boundaries.

Here we show that a standard feed-forward network can be used efficiently for supervised learning on topologically-protected magnetic skyrmion states and spin spirals originated from the spin-orbit coupling. Fig.1 illustrates the idea of our approach. A non-collinear magnetic configuration obtained from the Monte Carlo simulations describing a two-dimensional ferromagnet with Dzyaloshinskii-Moriya interaction (Fig.1 a) is projected on the z axis (Fig.1 b). This z-projected magnetic structure is considered as input for the single-hidden-layer network (Fig.1 c). Having trained such a network on a limited set of the configurations belonging to pure ferromagnetic, skyrmion and spiral states on the square lattice we were able to recognize the states from completely different parts of the phase diagram. Moreover, we found that the trained network can classify the data collected for triangular lattice, which demonstrates universality of our approach.

Model and Method. — In our study to simulate the topological magnetic excitations we used the following spin Hamiltonian on the 48×48 square lattices:

$$H = -\sum_{i<j} J_{ij}\mathbf{S}_i\mathbf{S}_j - \sum_{i<j}\mathbf{D}_{ij}[\mathbf{S}_i \times \mathbf{S}_j] - \sum_i BS_i^z \quad (1)$$

where J_{ij} and $\mathbf{D}_{ij}$ are the isotropic exchange interaction and Dzyaloshinskii-Moriya vector, respectively. $\mathbf{S}_i$ is a unit vector along the direction of the ith spin and B denotes the z-oriented magnetic field. We take into account the interaction between nearest neighbours. The isotropic exchange interaction is positive in our simulations, which corresponds to the ferromagnetic case. The symmetry of the Dzyaloshinskii-Moriya vectors is of C_{4v} type, DMI has an in-plane orientation and perpendicular to the corresponding inter-site radius vector. The Hamiltonian was solved by using the classical Monte Carlo approach [24].

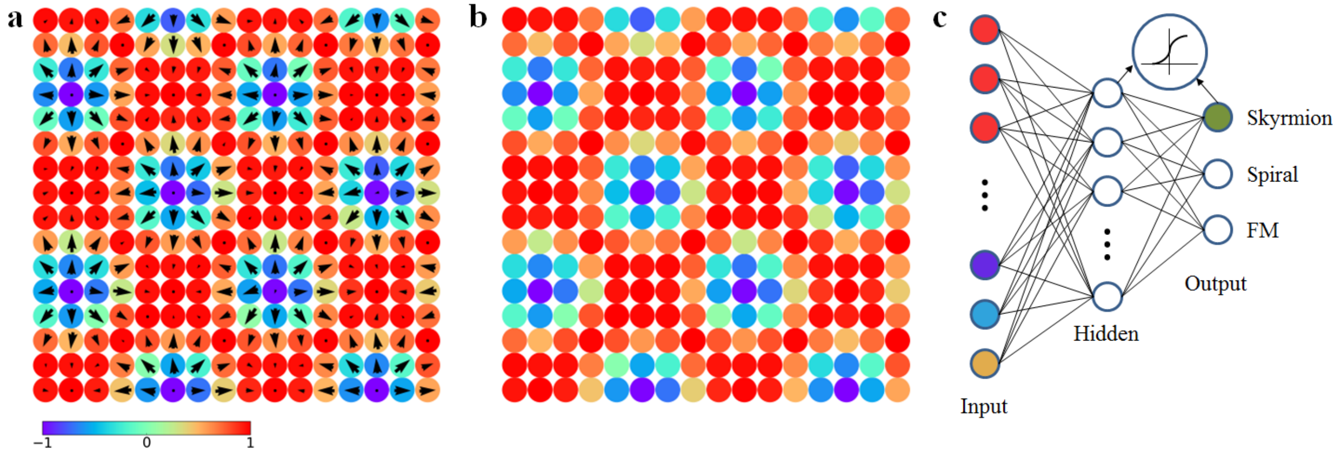

FIG. 1. Schematic representation of the machine learning process. (a) The skyrmion magnetic structure as obtained from the classical Monte Carlo simulations for a two-dimensional ferromagnet with Dzyaloshinskii-Moriya interaction at finite temperature and magnetic fields. Black arrows indicate the in-plane xy spin components. (b) The matrix contains the z projection of the spin structure to be classified. (c) Neural network with single hidden layer of sigmoid neurons. The values of the input neurons are equal to z components of the spins of the magnetic configuration.

To identify the different magnetic phases of the spin Hamiltonian, Eq.(1) we calculated spin-spin correlations functions [25], topological charges [26] and visualized a number of the spin configurations from each simulation. By using such information a neural network was trained as described below.

Machine learning.— In our study we employ a standard network architecture that is one-layer feed forward network presented in Fig.1 c. It consists of one hidden layer of sigmoid activation neurons and three output sigmoid neurons that activate depending on the particular magnetic phase. For the training set we generated 1000 configurations for each of ferromagnetic, skyrmion and spiral states corresponding to the areas marked in Fig.2. In these simulations we fixed $J = 1$ and used a uniform distribution for magnetic field and Dzyaloshinskii-Moriya interaction. The simulation temperatures were taken in the range $T \in [0.02, 0.1]$ in units of isotropic exchange interaction. Moreover, we generated 1000 configurations belonging to paramagnetic phase at high temperatures ($T \sim 10J$) and added them to the training set. For these paramagnetic configurations the signals of all the output neurons were set to zero.

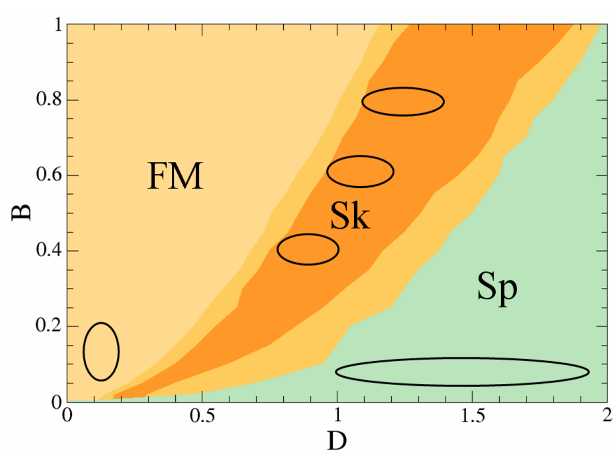

FIG. 2. Phase diagram in terms of Dzyaloshinskii-Moriya interaction and magnetic field. The abbrevation Sk, FM and Sp denote skyrmion lattice state, ferromagnetic and spin spiral state, respectively. The phase diagram was obtained at $T = 0.02$. All the parameters are given in units of J. Black ovals denote the phase areas used for supervised learning.

The main challenge in a machine learning for classification of magnetic phases is how to relate the states of the input neurons of the network to the particular magnetic configuration. As it was shown in Ref.10 in the case of the Ising model with $S^z = \pm 1$ there is one-to-one correspondence between the neuron values in the input layer and spins of the particular configurations. On the other hand, for the XY model solutions characterized by in-plane non-collinear magnetic states the authors of Ref.11 used the angle values determining the in-plane orientations of the spins.

In the case of the non-collinear magnetic configurations the situation is more complicated, since the orientation of a spin can not be described by single angle value. However, one can make use of that skyrmions are characterized by a typical profile, the core and background spins of a skyrmion align anti-parallel and parallel to the applied magnetic field (Fig.1 b), respectively. It means that the skyrmion excitation can be detected by analyzing the z components of the spins [27]. We use this fact to realize our neural network approach, the values of the input neurons are equal to the z components of the spins obtained from Monte Carlo simulations of Eq.(1). As we will show below, such an approach also works well in the case of the spin spiral and ferromagnetic phases.

The network was trained to minimize the error function. As the loss function we used a standard Mean Squared Error (MSE) function [28]. Weights of neurons were adjusted by means of back-propagation method. During the learning process, we calculated the mean value of the error function on each epoch in order to avoid overfeeding. The network was trained with different numbers of hidden neurons from 8 to 128. According to our simulations the network with 64 hidden neurons gives reliable results on the phases recognition. We found that further increase of hidden neurons number for considered case leads to decrease in recognition quality. Thus, the total number of adjustable parameters are $64L^2 + 192$ that is much smaller than in the previous work [11].

Phase diagram.—The developed neural network approach was used for construction of the phase diagram of the spin model, Eq.(1) on the square lattice [29]. From Fig.3 one can see that the trained network can successfully recognize all the phases of interest at low temperature, which follows from a comparison with the boundaries obtained by calculating the structure factor. It is worth mentioning that we obtained a large value of skyrmion number ($Q > 15$) for the parameters corresponding to the dark green area in Fig.3. Importantly, it is possible to perform a composition analysis of the transitional areas between different phases. For each point of the phase diagram one can define the values of the output neurons that indicate the contributions of the phases. It gives us opportunity to solve the complex problem of the definition of the phase boundaries and quantitatively characterize the transitional areas between different phases [30, 31].

Analysis of the classification process.— The results of the previous neural-network-based studies [1, 10, 11] stand new fundamental questions on how a network learns the different phases of matter. It was shown in Ref.10 that identification of the Ising model states is related to difference in total magnetization of the spin configurations belonging to different phases. In our case such an explanation can be also used, since the phases we simulated are characterized by different magnetizations. The magnetization per spin, $m(x) = \frac{1}{N}\sum_i^N S_i^z$ in the training set is in the range [0.91, 0.99], [0.38, 0.53] and [0, 0.03] for ferromagnetic, skyrmion and spin spiral phases, respectively. At the same time, the test sets include pure spin configurations that are characterized by wider ranges of the average magnetization: [0.84, 0.99] (ferromagnetic), [0.33, 0.69] (skyrmion) and [0, 0.07] (spin spiral). In agreement with Ref.10 we obtain that the components of the vector Wx (here W is the weights matrix between input and hidden layers) become linear functions of the magnetization $m(x)$. However, in our case the increase

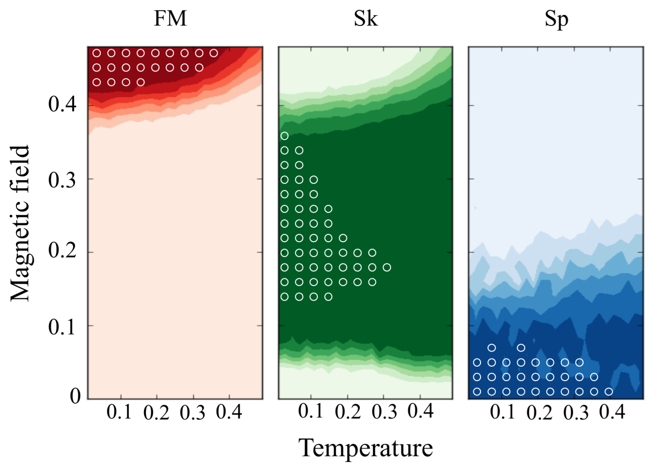

FIG. 3. Phase triptych obtained by using the neural network with 64 hidden neurons. Color intensities indicate the values of the output neurons for different phases, dark and light colors correspond to 1 and 0, respectively. The Dzyaloshinskii-Moriya interaction was choosen to be $D = 0.72$. All the parameters are given in units of J. White circles denote the phases boundaries defined with the spin structure factors.

of the number of the hidden neurons leads to a larger number of the neuron categories, which may mean that the magnetization is not the only parameter the network uses for recognition.

Another way to understand the neural network functioning is to visualize hidden layer neurons. By the example of the configuration with big skyrmions presented

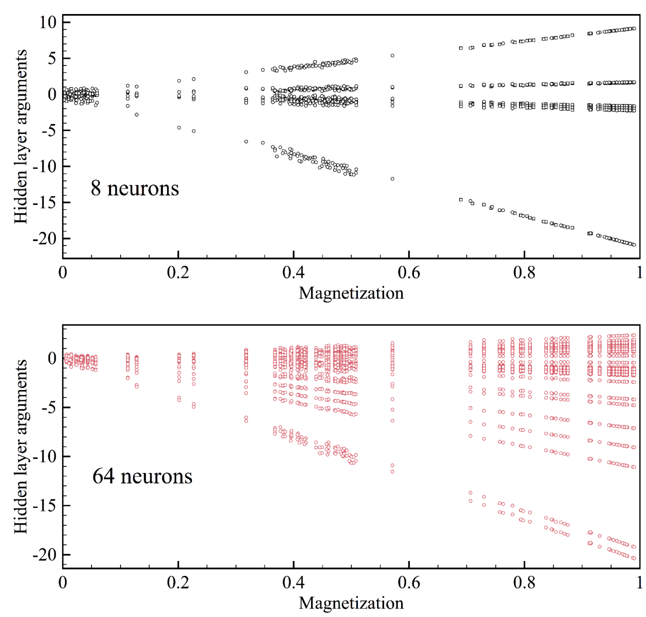

FIG. 4. Hidden layer arguments as a function of the z-oriented magnetization of the simulated spin configurations. (Top) 8-hidden-neuron network. (Bottom) 64-hidden-neuron network.

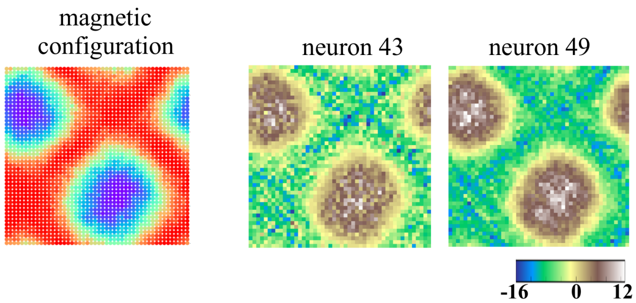

FIG. 5. (Left panel) z-projection of the skyrmion magnetic configuration obtained with the parameters $J = 1, D = 0.2, T = 0.02$ and $B = 0.01$. (Right panel) Visualization of the hidden layer neurons arguments.

in Fig.5 we performed such an analysis. Importantly, the size of the skyrmions in the training data set does not exceed $10a$, where a is the lattice constant, but we found, that the trained neural network correctly classifies the configurations with skyrmions of much larger diameter. Indeed, the diameter of the skyrmion in Fig.5 is about $35a$ and such a skyrmion state is uniquely recognized by the neural network even with 8 hidden neurons.

Fig.5 gives two-dimensional representation of two hidden neurons arguments that are the weights matrix mutliplied by spin z components corresponding to the magnetic configuration. The maximal and minimal intensities of the core and background areas of the skyrmions are different for these neurons. Nevertheless, one can easily recognize the original skyrmion structure. The visualization of the neural network weights by themselves does not give any useful information about network functioning.

As a hard test for our neural-network approach we generated 300 high temperature spiral configurations ($T \in [0.18; 0.26], D = 0.72, B = 0.03$). A typical example of such configurations is presented in Fig.6 (b). It is of labyrinth type and consists of the broken spin spirals that are distorted due to the temperature effects. Importantly, the training set contains only ideal spin spirals presented in Fig.6 (a). One can see that increase of the number of the hidden neurons leads to a decrease in the value of the output neuron corresponding to the skyrmion phase that provides a more accurate phase separation. Having analyzed this test set, we found that total number of clearly recognised configurations increased from 40% to 75% with using 8 and 64 hidden neurons respectively.

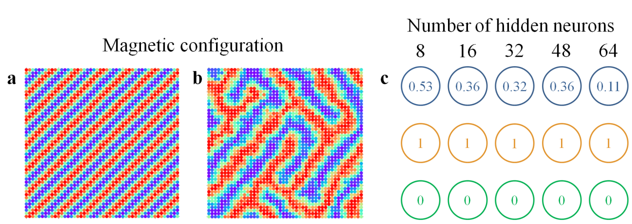

FIG. 6. (a) Example of spiral state (D=1.4, B=0.02, T=0.05, J=1) used for training the network. (b) Example of a complex spiral configuration from a test set obtained with D=0.72, B=0.03, T=0.22, J=1. (c) The output neurons values in the case of configuration (b) depending on the number of the hidden neurons. Numbers in blue, orange and green circles correspond to values of skyrmion, spiral and FM outputs respectively.

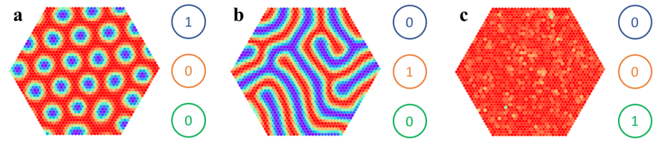

FIG. 7. Examples of skyrmion (a), spin spiral (b) and ferromagnetic (c) configurations stabilized on the triangular lattice and recognized with the neural network trained on square lattice data. Numbers in blue, orange and green circles correspond to values of skyrmion, spiral and FM outputs respectively.

Variation of the lattice structure.—The next step of our investigation is to examine the network trained on the square lattice magnetic configurations for recognizing the phases of the spin Hamiltonian on the triangular lattice [32]. For that we solved Eq.(1) with DMI of C_{3v} symmetry and generated magnetic configurations belonging to skyrmion, spin spiral and ferromagnetic phases as well as their mixtures. Fig.7 gives the corresponding examples. It was found that the trained network classifies of the skyrmion and ferromagnetic triangular-lattice configurations with high precision. In the case of the spin spiral states the classification accuracy is low, since such magnetic configurations (typical example is presented in (Fig.7 b) strongly differ from those we used in the training set (Fig.6 a).

Conclusions.— We have developed a neural-network-based approach for recognition magnetic phases of two-dimensional ferromagnets with Dzyaloshinskii-Moriya interaction in wide ranges of magnetic fields and temperatures. In contrast to previous works one needs to generate a limited set of magnetic configurations ($\sim$ 4000 in total) to train the network. It facilitates the construction of the phase diagram of the system in question during the Monte Carlo sampling. Complex and mixed ferromagnetic-skyrmion and skyrmion-spin-spiral configurations can be quantitatively described, which was not possible before. The calculations for spin Hamiltonians on the 128×128 square lattice also demonstrated high accuracy in classification magnetic phases. We have shown that the method does not sensitive to the particular lattice structure used for training. Our machine learning approach can be used for recognition of the skyrmion magnetic configurations observed in experiments.

Acknowledgements. We thank Andrey Bagrov, Frederic Mila and Jeanne Colbois for fruitful discussions.

[1] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[2] Y. Nomura, A. S. Darmawan, Y. Yamaji, M. Imada, Phys. Rev. B **96**, 205152 (2017).

[3] H. Saito, M. Kato, J. Phys. Soc. Jpn. **87**, 014001 (2018).

[4] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).

[5] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).

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

[7] Y. Zhang, R. G. Melko, and E.-A. Kim, Phys. Rev. B **96**, 245119 (2017).

[8] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).

[9] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).

[10] J. Carrasquilla, R. G. Melko, Nature Physics **13**, 431 (2017).

[11] M. J. S. Beach, A. Golubeva, R. G. Melko, arXiv:1710.09842

[12] Y. LeCun, Y. Bengio, G. Hinton, Nature **521**, 436 (2015).

[13] A. Bogdanov, A. Hubert, Phys. Stat. Sol. (b) **186**, 527 (1994).

[14] A. Bogdanov, A. Hubert, J. Magn. Magn. Mater. **138**, 255 (1994).

[15] S. Mühlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, P. Böni, Science **323**, 915 (2009).

[16] F. Jonietz, S. Mühlbauer, C. Pfleiderer, A. Neubauer, W. Münzer, A. Bauer, T. Adams, R. Georgii, P. Böni, R. A. Duine, K. Everschor, M. Garst, A. Rosch, Science **330**, 1648 (2010).

[17] A. Neubauer, C. Pfleiderer, B. Binz, A. Rosch, R. Ritz, P. G. Niklowitz, and P. Böni, Phys. Rev. Lett. **102**, 186602 (2009).

[18] F. Zheng, H. Li, S. Wang, D. Song, C. Jin, W. Wei, A. Kovács, J. Zang, M. Tian, Y. Zhang, H. Du, R. E. Dunin-Borkowski, Phys. Rev. Lett. **119**, 197205 (2017).

[19] X. Z. Yu, N. Kanazawa, Y. Onose, K. Kimoto, W. Z. Zhang, S. Ishiwata, Y. Matsui, Y. Tokura, Nature Materials **10**, 106 (2011).

[20] M. Nagao, Y.-G. So, H. Yoshida, K. Yamaura, T. Nagai, T. Hara, A. Yamazaki, and K. Kimoto, Phys. Rev. B **92**, 140415 (2015).

[21] S. Heinze, K. von Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, S Blügel, Nature Physics **7**, 713718 (2011).

[22] N. Kanazawa, J.-H. Kim, D. S. Inosov, J. S. White, N. Egetenmeyer, J. L. Gavilano, S. Ishiwata, Y. Onose, T. Arima, B. Keimer, and Y. Tokura, Phys. Rev. B **86**, 134425 (2012).

[23] X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, Y. Tokura, Nature **465**, 901904 (2010).

[24] The spin update scheme is based on the Metropolis algorithm. The systems in question are gradually (200 temperature steps) cooled down from high temperatures (T $\sim 3J$) to the required temperature. Each temperature step run consists of 1.5×10^6 Monte Carlo steps.

[25] D. I. Badrtdinov, S. A. Nikolaev, M. I. Katsnelson, and V. V. Mazurenko, Phys. Rev. B **94**, 224418 (2016).

[26] Berg, B., Lüscher, M. Nuclear Physics B **190**, 412 (1981).

[27] M. C. Ambrose and R. L. Stamps, New Journal of Physics **15**, 053003 (2013).

[28] The error function is given by $E(out^{ideal}, out^{actual}) = \frac{\sum_{i=1}^{n}(out_i^{ideal} - out_i^{actual})^2}{n}$, where n is the number of output neurons, out^{ideal} represents the training labels and out^{actual} the obtained values of the output neurons.

[29] To construct the phase diagram (Fig.3) we used a grid of 625 points on the temperature-magnetic field plane. For each point the values of the neural network output neurons were averaged over 10 Monte Carlo runs. Thus the total number of the Monte Carlo calculations was equal to 6250.

[30] S. El Hog, A. Bailly-Reyre, H. T. Diep, J. Magn. Magn. Mater. **452**, 32-38 (2018).

[31] I. A. Iakovlev, O. M. Sotnikov, V. V. Mazurenko, arXiv:1801.07896.

[32] We have solved spin Hamiltonain, Eq.(1) on the triangular lattice 48×48 with periodic boundary conditions. The supercell of the rhombic shape was replicated. A square area of 48×48 spins cropped from the replicated lattice was used to define the values of the neural network input neurons.