Active Learning for Discovering Complex Phase Diagrams with Gaussian Processes

Max Zhu[*]
Statistical Laboratory
University of Cambridge
mz406@cam.ac.uk

Jian Yao[*]
Department of Physics
Southern University of Science and Technology
yaoj2020@mail.sustech.edu.cn

Marcus Mynatt
Department of Physics
University of Florida
marcus.mynatt@ufl.edu

Hubert Pugzlys
Department of Physics
University of Florida
phubert.pugzlys@ufl.edu

Shuyi Li
Department of Physics
University of Florida
lis3@ufl.edu

Chunjing Jia
Department of Physics
University of Florida
chunjing@phys.ufl.edu

Qingyuan Zhao
Statistical Laboratory
University of Cambridge
qyzhao@statslab.cam.ac.uk

Abstract

We introduce a Bayesian active learning algorithm that efficiently elucidates phase diagrams. Using a novel acquisition function that assesses both the impact and likelihood of the next observation, the algorithm iteratively determines the most informative next experiment to conduct and rapidly discerns the phase diagrams with multiple phases. Comparative studies against existing methods highlight the superior efficiency of our approach. We demonstrate the algorithm's practical application through the successful identification of a skyrmion phase diagram. Finally, we provide code at https://github.com/Maccy-Z/Active_Learning_Phase_Diagram, as well as a graphical user interface application.

1 Introduction and Background

The phase diagram represents an indispensable map for understanding the different states of matter and the transitions between them under varying conditions of temperature and pressure. This conceptual tool is not only foundational in thermodynamics but also a cornerstone in material science, where it guides the synthesis and characterization of materials like alloys Okamoto et al. [2016]. In the realm of superconducting physics, phase diagrams serve a critical role in identifying the conditions under which materials exhibit superconductivity. The ability to pinpoint the exact pressures and temperatures that facilitate superconductivity enables researchers to design new materials and improve upon existing technologies Court and Cole [2020]. As such, the phase diagram is an invaluable reference for scientists and engineers seeking to manipulate the physical states of matter to innovative ends. It is crucial to determine the phase diagrams of different systems. The traditional method of determining the phase diagram, whether experimentally or theoretically, entails a grid search across the entire phase space, which could become exceedingly time-consuming as the dimension of the parameter space and the number of phases increase, or if experiments are slow.

[*]Equal Contribution

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

Active learning Settles [2009] is an iterative process where an algorithm repeatedly determines how to sample new observations based on existing observations. Originally proposed for complex optimizations Kushner [1964], it found applications in a broad range of problems including optimizing material design Lookman et al. [2019]. Dai and Glotzer [2020] applied active learning to uncover phase diagrams using a Gaussian Process (GP) model Rasmussen and Williams [2006]. Using GPs to quantify uncertainty, their method prioritizes selecting points near the predicted phase boundary and of high model uncertainty. However, we believe their acquisition function can be improved and their method is limited to modeling two phases. Tian et al. [2020] investigated phase diagrams with more than two phases, a classifier predicts phases and generates phase diagrams through prediction boundaries. However, the uncertainty of the phase boundary is not quantified, and direct observation of the phase boundary is required to perform phase boundary learning.

We propose a Bayesian active learning framework for efficiently discovering phase diagrams. We introduce a specially designed acquisition function with a GP-based model. Compared to existing approaches, our method has the following features: firstly, the phase diagram can contain any number of phases. Secondly, instead of direct phase transition measurements, only point measurements of the phase are necessary. Thirdly, our method is able to adapt to noisy measurements of phase. As a result, our algorithm converges to the true phase diagram considerably faster than existing baselines.

We apply our algorithm to investigate the phase diagram of a spin system on monolayer honeycomb lattice under the influence of magnetic field. Our algorithm efficiently identifies the skyrmion—a topological spin configuration—phase diagram with just a small sample. As an exotic state of matter, skyrmion has great potential in next-generation technological applications in microwave detectors Finocchio et al. [2015], nano oscillators Liu et al. [2015], Zhou et al. [2015], Carpentieri et al. [2015], logic gates Zhang et al. [2015], and quantum computation Psaroudaki and Panagopoulos [2021] etc.

2 Methodology

Our objective is to model a phase diagram with n phases (e.g. state of matter) determined by k parameters $\mathbf{x} \in \mathbb{R}^k$ (e.g. temperature or interaction strength), where experiments (or simulations) can determine the phase $y(\mathbf{x}) \in \{1, ..., n\}$ at point $\mathbf{x}$. Using the active learning framework, at step t, the model takes existing observations $\mathcal{T} = \{\mathbf{x}_i, y_i\}_{i=1}^t$ and determines the next point to sample, $\mathbf{x}_{t+1}$. An experiment is performed to give $y_{t+1}(\mathbf{x}_{t+1})$ and this process is repeated. A good algorithm generates a model that accurately predicts $y(\mathbf{x})$ using the fewest observations possible.

Our Model: The phase diagram is modeled as a collection of n GPs, $\{f_1(a_1|\mathbf{x})...f_n(a_n|\mathbf{x})\}$, one for each phase, where $a_i, i \in \{1...n\}$, gives the magnitude of each phase and is used for fitting and constructing the phase diagram as follows. At step t of our algorithm, the GPs are fitted to the current observations $\mathcal{T}$, $f_i(a_i|\mathbf{x}) = f_i(a_i|\mathbf{x}, \mathcal{T})$. To sample the phase at point $\mathbf{x}$, we sample from each of the GPs at $\mathbf{x}$ and take the phase as the maximum over all the phases, giving a phase diagram:

$$\mathcal{P}_{\mathcal{T}}(\mathbf{x}) = \underset{i}{\operatorname{argmax}}\, a_i, a_i \sim f_i(a_i|\mathbf{x}, \mathcal{T}), \mathcal{P}(\mathbf{x}) \in \{1, ..., n\}. \tag{1}$$

To simplify computation, we also define the most probable phase diagram as $P_{\mathcal{T}}(\mathbf{x}) = \operatorname{argmax}_i \mu_i(\mathbf{x}|\mathcal{T})$, where μ_i is the posterior mean of f_i.

Fitting to observations: As there are more than two phases, we extend standard Gaussian process classifiers (GPC) Krasser [2020], Rasmussen and Williams [2006] for multiclass observations. Each phases's GP f_i is individually fitted as a binary GPC in a one-vs-rest setup using $p(y(\mathbf{x}) = i|a_i, \mathcal{T}) = \sigma(a_i), a_i \sim f_i(a_i|\mathbf{x}, \mathcal{T})$. The posterior $f_i(a_i|\mathbf{x}, \mathcal{T})$ is fitted with the Laplace approximation, with additional details in Appendix 7.2. Each phase is fitted independently. By using GPCs instead of directly fitting GPs to binary observations, as in Dai and Glotzer [2020], our proposed method enhances training stability and naturally incorporates measurement errors. There are other approaches to fitting multi-class Gaussian process classifiers, such as Galy-Fajou et al. [2019]; however, they are more complex or slower to fit.

Acquisition function: Following the active learning algorithm, the next point to sample $\mathbf{x}_{t+1}$ is determined using the acquisition function $A(\mathbf{x})$ that measures how good each point is to sample and greedily selected with $\mathbf{x}_{t+1} = \operatorname{argmax}_{\mathbf{x}} A(\mathbf{x})$ by sampling $A(\mathbf{x})$ over a grid. We propose utilizing the expected change in the phase diagram after performing an experiment at that point as the

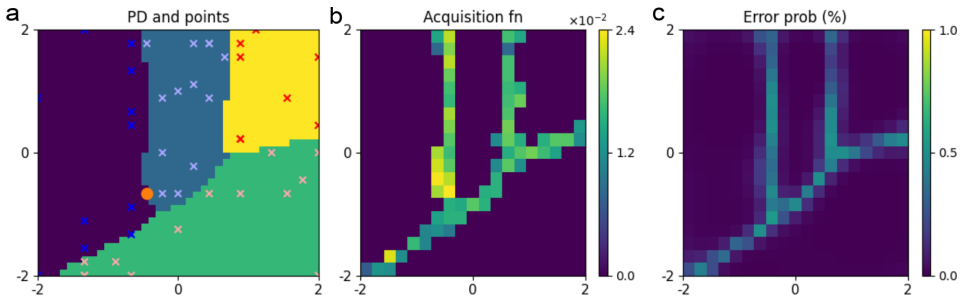

Figure 1: Visualization of our model internal state on a phase diagram with four phases. (**a**) Predicted most likely phase (shaded) with previous observations (crosses), colored by the phase observed. The recommended point to sample is shown as an orange dot. (**b**) Acquisition function rating points to sample. (**c**) Predicted error (%) of observing different phases from predicted.

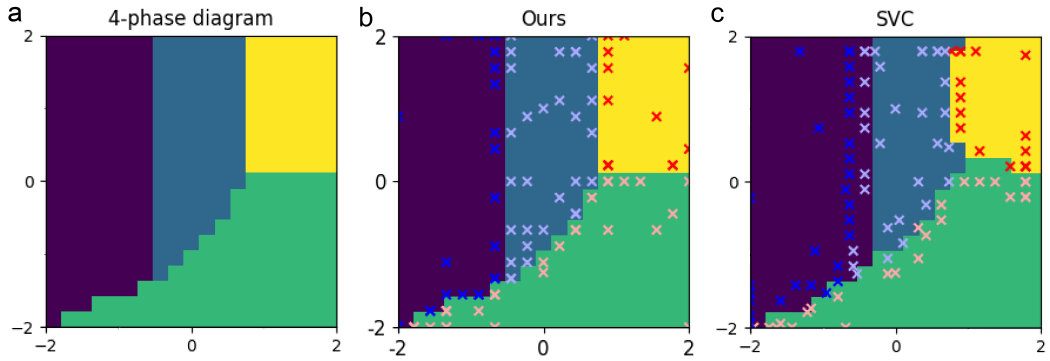

Figure 2: (**a**) A phase diagram of four phases used for testing, with phases shown with shaded color. Predicted phase diagram after 100 observations from our model (**b**) and SVC (**c**), with predicted phase shaded and observations as crosses colored by observed phase.

acquisition function $A(\mathbf{x})$. A distance function between two phase diagrams is defined as

$$D(P_1, P_2) = \frac{1}{V}\int \mathbb{1}(P_1(\mathbf{x}) \neq P_2(\mathbf{x}))d\mathbf{x} \approx \frac{1}{|\mathcal{G}|}\sum_{\mathbf{x}\in\mathcal{G}} \mathbb{1}(P_1(\mathbf{x}) \neq P_2(\mathbf{x})), \tag{2}$$

the fractional area where the two phase diagrams differ. This is numerically approximated by sampling over a grid of pre-chosen regular points, $\mathcal{G}$ in parameters space. We choose $A(\mathbf{x}) = \mathbb{E}[D(P_\mathcal{T}, P_{\mathcal{T}\cup\{\mathbf{x},\hat{y}\}})], \hat{y} \sim \mathcal{P}_\mathcal{T}(\mathbf{x})$ to be the expected distance between the current most probable phase diagram $P_\mathcal{T}$ and the most probable phase diagram after sampling at a new point $\mathbf{x}$, assuming observation $\hat{y}$ is distributed according to the uncertainty of our model. The expectation averages over the possible observations $\hat{y}$. This acquisition function combines the impact of an observation on the entire phase diagram and the probability of that observation and balances the exploration of unknown regions with the refinement of known boundaries. This process is visualized in Figure 1. Appendix 7.3 describes implementation details of f_i and several numerical optimizations, with these optimizations, each step requires ~25 seconds on an 8-core processor.

3 Synthetic Experiments

Our model is evaluated against Dai and Glotzer [2020] (GPR) and a modified version of Approach 1 in Tian et al. [2020] (SVC) as baselines. Details of these baselines are given in Appendix 7.1. Our experiments are restricted to $k = 2$ but our methods can be adapted to higher k.

Two phase problem: Firstly, we evaluated the models on a noisy toy phase diagram of two phases, $P(x_1, x_2) = \mathbb{1}(x_2 > \sin(0.5\pi x_1))$ where $x_1, x_2 \in [-2, 2]$ represent experimental parameters. Uniform noise, $\epsilon \sim [-0.2, 0.2]$ was added to x_1, x_2 every observation to simulate experimental errors. Models were initialized with 2 observations. Figure 3 and Table 1 show the percent difference

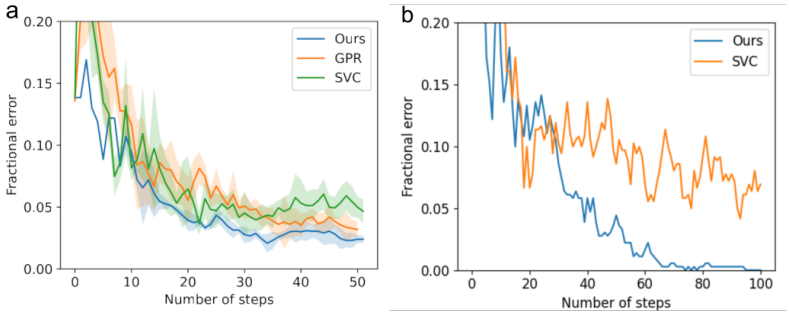

Figure 3: Fractional error between predicted and true phase diagrams versus number of samples taken for phase diagrams of (**a**) two phases and (**b**) four phases. Shaded area in (**a**) shows standard deviation over 6 runs with different random noise, while (**b**) is noiseless.

between the predicted and true phase diagram (fractional error) vs. the number of observations, averaged over 6 runs. Our method surpasses the baselines in accuracy after 12 observations.

Four phase problem: Our model and Tian (SVC) were evaluated on a more complex phase diagram of four spaces inspired by the pressure-temperature phase diagram of materials (Figure 2 (**a**)). Figure 3 (**b**) shows the error vs. number of observations and final predictions are shown in Figure 2 (**b**) and (**c**). Our model outperforms the SVC after 30 steps, and nearly perfectly predicts the phase diagram after 100 steps. SVC predictions are unaligned with observations, which was also observed in Ref Tian et al. [2020]. Figure 1 visualizes our model's predicted phase diagram, acquisition function, and error probability. Figure A1 in Appendix shows the evolution of our models predictions.

4 Application on Heisenberg Model with Antisymmetric Interaction

A motivating application is the exploration of exotic phases in frustrated magnetic systems. One particular phase that has recently garnered significant attention is magnetic skyrmions.Evans et al. [2014a], Finocchio et al. [2016], Wang et al. [2022] In the two-dimensional scenario, magnetic skyrmions manifest as real-space topological structures, formed by localized spin vectors that smoothly transition from their cores to their peripheries while maintaining fixed helicities. The real-space topology of a skyrmion is characterized by its skyrmion number $Q = \frac{1}{4\pi}\int \boldsymbol{m}\cdot(\frac{\partial \boldsymbol{m}}{\partial x}\times\frac{\partial \boldsymbol{m}}{\partial y})dxdy$, where $\boldsymbol{m}$ is the unit vector of local magnetization and the integral is taken over a two-dimensional space. To investigate the skyrmion phase, we focus on a Heisenberg model with antisymmetric interaction on the honeycomb lattice:

$$H = -J\sum_{\langle i,j\rangle}\boldsymbol{S_i}\cdot\boldsymbol{S_j} - \sum_{\langle i,j\rangle}\boldsymbol{d_{i,j}}\cdot(\boldsymbol{S_i}\times\boldsymbol{S_j}) - D\sum_i(S_i^z)^2 - h\sum_i S_i^z, \tag{3}$$

where $\boldsymbol{S_i}$ represents the classical spin vector at site i with $|\boldsymbol{S_i}| = 1$ and S_i^z is z-component. $J > 0$ is the nearest neighbor (NN) ferromagnetic coupling, and $\boldsymbol{d_{i,j}} = d\ (\boldsymbol{z}\times\boldsymbol{u_{i,j}})$ is the NN Dzyaloshinskii–Moriya (DM) interaction Dzyaloshinsky [1958], Moriya [1960], where $\boldsymbol{z}$ and $\boldsymbol{u_{i,j}}$ are unit vectors, respectively perpendicular to the magnetic layer plane and pointing from site i to site j, as shown in Fig. 4 (**a**). $D > 0$ represents easy-axis single-ion anisotropy, and $h > 0$ is the magnetic field. Besides the skyrmion phase, we have identified another two distinct magnetic phases in this system: the spiral phase and polarized paramagnet, in which the spin configuration has magnetic wave vector $\boldsymbol{q}$ and $\boldsymbol{0}$. An example of the spiral phase and skyrmion phase is illustrated in Fig. 4 (**b**) (**c**).

In order to determine the magnetic state configuration, we perform atomistic spin dynamics simulation by VAMPIRE Evans et al. [2014b] under fixed $J = 1$ meV, $d/J = 0.5$ and temperature lowered to $T/J = 10^{-4}$. The system size is chosen to be $L_x = L_y = 50a$ with periodic boundary condition, where a is the lattice constant. In constructing the phase diagram for $D/J \in (0, 0.5)$ and $h/J \in (0, 0.25)$ using our algorithm, we start with 6 random initial starting points, then increase the number of observations to $t = 16, 26, 36, 46, 56, 66, 76$. The resulting phase diagrams are illustrated in Fig. 4 (**d**). As a comparison, the black lines in each plot represent the true phase boundaries

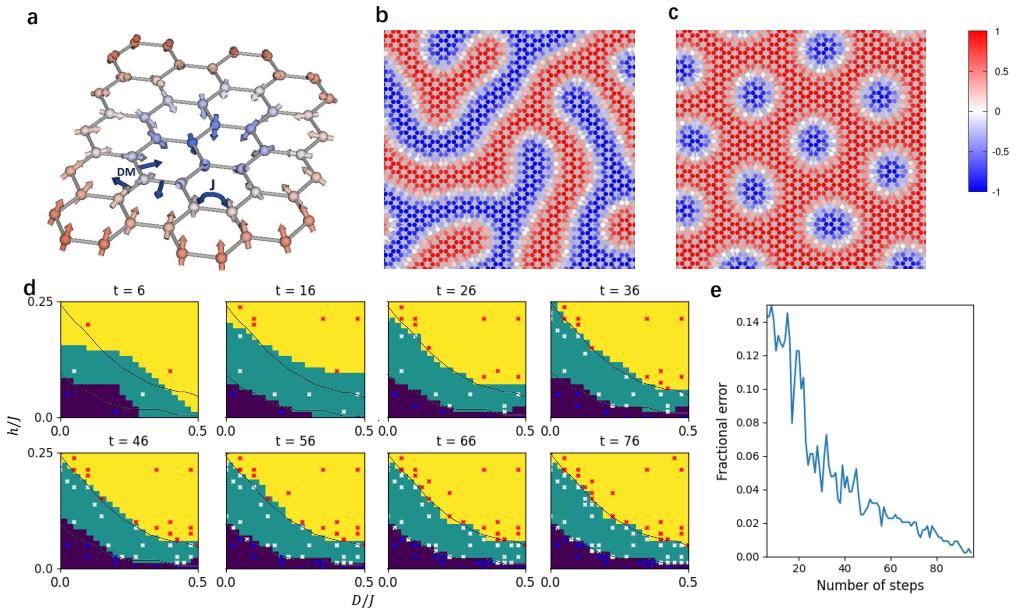

Figure 4: Phase diagram of Heisenberg model with antisymmetric interaction using our approach. (**a**) A schematic of the model Hamiltonian and the microscopic spin configuration for skyrmion. J and DM represent the spin exchange and DM vector for Dzyaloshinskii–Moriya interaction respectively. Single-ion anisotropy D and magnetic field h are not shown. (**b**) and (**c**) are real-space spin configurations of (**b**) spiral and (**c**) skyrmion phases. Color denotes S_z component (colorbar shown) and arrows denote the spin direction in (**a**) and (S_x, S_y) components in (**b**) and (**c**). Here (**b**) and (**c**) are obtained at $(D, h)/J = (0.056, 0)$ and $(0.028, 0.7)$, respectively. (**d**) Phase diagrams as functions of D/J and h/J, for $d = 0.5J$ and temperature $T = 10^{-4}J$, obtained after t observations. Crosses with colors represent points sampled: blue for spiral phase, white for skyrmion phase, and red for polarized paramagnet. We use purple, green, and yellow to fill out inferred regions of the three different phases respectively. As a comparison, black lines show the true phase boundaries obtained by grid sampling. (**e**) Fractional error as a function of the number of observations.

obtained by grid sampling. The fractional error decreases as the number of observations grows, becoming smaller than 0.25% after $t = 90$, as illustrated in Fig. 4 (**e**).

5 Conclusion and future work

Our proposed method is able to efficiently determine phase diagrams by identifying where to sample new observations, using a Bayesian approach to quantify the uncertainty. Demonstrated to be effective for synthetic experiments and for skyrmion phase diagram problem, this approach holds promise to significantly accelerate the identification of phase diagrams, especially when the dimension of the parameters space expands. For future work, we would like to extend our approach by improving how to combine the individual GPC phase diagrams using a single multiclass Gaussian process. Additionally, we would like to explicitly incorporate known experimental errors into our model.

6 Acknowledgments

We acknowledge support for this work from GSK and the Cambridge Centre for AI and Medicine. C. J. Jia acknowledges the support from Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, under Contract No. DE-AC02- 76SF0051.

References

H. Okamoto, M.E. Schlesinger, and E.M. Mueller. *Alloy Phase Diagrams*. ASM International, 04 2016. ISBN 978-1-62708-163-4. doi: 10.31399/asm.hb.v03.9781627081634. URL https://doi.org/10.31399/asm.hb.v03.9781627081634.

Callum J. Court and Jacqueline M. Cole. Magnetic and superconducting phase diagrams and transition temperatures predicted using text mining and machine learning. *npj Computational Materials*, 6(1):18, Mar 2020. ISSN 2057-3960. doi: 10.1038/s41524-020-0287-8. URL https://doi.org/10.1038/s41524-020-0287-8.

Burr Settles. Active learning literature survey. *University of Wisconsin-Madison Department of Computer Sciences*, 2009.

H. J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.*, 1964.

Turab Lookman, Prasanna V. Balachandran, Dezhen Xue, and Ruihao Yuan. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials*, 2019.

Chengyu Dai and Sharon C. Glotzer. Efficient phase diagram sampling by active learning. *The Journal of Physical Chemistry B*, 124(7):1275–1284, 2020. doi: 10.1021/acs.jpcb.9b09202. URL https://doi.org/10.1021/acs.jpcb.9b09202. PMID: 31964140.

Carl Edward Rasmussen and Christopher K. I. Williams. Gaussian processes for machine learning. *The MIT Press, 2006. ISBN 0-262-18253-X*, 2006.

Yuan Tian, Ruihao Yuan, Dezhen Xue, Yumei Zhou, Yunfan Wang, Xiangdong Ding, Liu Juan, and Turab Lookman. Determining multi-component phase diagrams with desired characteristics using active learning. *Advanced Science*, 8:2003165, 11 2020. doi: 10.1002/advs.202003165.

G. Finocchio, M. Ricci, R. Tomasello, A. Giordano, M. Lanuzza, V. Puliafito, P. Burrascano, B. Azzerboni, and M. Carpentieri. Skyrmion based microwave detectors and harvesting. *Applied Physics Letters*, 107(26):262401, 12 2015. ISSN 0003-6951. doi: 10.1063/1.4938539. URL https://doi.org/10.1063/1.4938539.

R. H. Liu, W. L. Lim, and S. Urazhdin. Dynamical skyrmion state in a spin current nano-oscillator with perpendicular magnetic anisotropy. *Phys. Rev. Lett.*, 114:137201, Mar 2015. doi: 10.1103/PhysRevLett.114.137201. URL https://link.aps.org/doi/10.1103/PhysRevLett.114.137201.

Y. Zhou, E. Iacocca, A. A. Awad, R. K. Dumas, F. C. Zhang, H. B. Braun, and J. Åkerman. Dynamically stabilized magnetic skyrmions. *Nature Communications*, 6(1):8193, Sep 2015. ISSN 2041-1723. doi: 10.1038/ncomms9193. URL https://doi.org/10.1038/ncomms9193.

Mario Carpentieri, Riccardo Tomasello, Roberto Zivieri, and Giovanni Finocchio. Topological, non-topological and instanton droplets driven by spin-transfer torque in materials with perpendicular magnetic anisotropy and dzyaloshinskii–moriya interaction. *Scientific Reports*, 5(1):16184, Nov 2015. ISSN 2045-2322. doi: 10.1038/srep16184. URL https://doi.org/10.1038/srep16184.

Xichao Zhang, Motohiko Ezawa, and Yan Zhou. Magnetic skyrmion logic gates: conversion, duplication and merging of skyrmions. *Scientific Reports*, 5(1):9400, Mar 2015. ISSN 2045-2322. doi: 10.1038/srep09400. URL https://doi.org/10.1038/srep09400.

Christina Psaroudaki and Christos Panagopoulos. Skyrmion qubits: A new class of quantum logic elements based on nanoscale magnetization. *Phys. Rev. Lett.*, 127:067201, Aug 2021. doi: 10.1103/PhysRevLett.127.067201. URL https://link.aps.org/doi/10.1103/PhysRevLett.127.067201.

Martin Krasser. Gaussian processes for classification. *http://krasserm.github.io/2020/11/04/gaussian-processes-classification/*, 2020.

Theo Galy-Fajou, Florian Wenzel, Christian Donner, and Manfred Opper. Multi-class gaussian process classification made conjugate: Efficient inference via data augmentation. *Uncertainty in artificial intelligence*, 2019.

R F L Evans, W J Fan, P Chureemart, T A Ostler, M O A Ellis, and R W Chantrell. Atomistic spin model simulations of magnetic nanomaterials. *Journal of Physics: Condensed Matter*, 26(10): 103202, feb 2014a. doi: 10.1088/0953-8984/26/10/103202. URL https://dx.doi.org/10.1088/0953-8984/26/10/103202.

Giovanni Finocchio, Felix Büttner, Riccardo Tomasello, Mario Carpentieri, and Mathias Kläui. Magnetic skyrmions: from fundamental to applications. *Journal of Physics D: Applied Physics*, 49 (42):423001, sep 2016. doi: 10.1088/0022-3727/49/42/423001. URL https://dx.doi.org/10.1088/0022-3727/49/42/423001.

Kang Wang, Vineetha Bheemarasetty, Junhang Duan, Shiyu Zhou, and Gang Xiao. Fundamental physics and applications of skyrmions: A review. *Journal of Magnetism and Magnetic Materials*, 563:169905, 2022. ISSN 0304-8853. doi: https://doi.org/10.1016/j.jmmm.2022.169905. URL https://www.sciencedirect.com/science/article/pii/S0304885322007909.

I. Dzyaloshinsky. A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics. *Journal of Physics and Chemistry of Solids*, 4(4):241–255, 1958. ISSN 0022-3697. doi: https://doi.org/10.1016/0022-3697(58)90076-3. URL https://www.sciencedirect.com/science/article/pii/0022369758900763.

Tôru Moriya. Anisotropic superexchange interaction and weak ferromagnetism. *Phys. Rev.*, 120: 91–98, Oct 1960. doi: 10.1103/PhysRev.120.91. URL https://link.aps.org/doi/10.1103/PhysRev.120.91.

R F L Evans, W J Fan, P Chureemart, T A Ostler, M O A Ellis, and R W Chantrell. Atomistic spin model simulations of magnetic nanomaterials. *Journal of Physics: Condensed Matter*, 26(10): 103202, feb 2014b. doi: 10.1088/0953-8984/26/10/103202. URL https://dx.doi.org/10.1088/0953-8984/26/10/103202.

SheffieldML. Gpy. In *https://github.com/SheffieldML/GPy*, 1.10.

B. Berg and M. Lüscher. Definition and statistical distributions of a topological number in the lattice o(3) σ-model. *Nuclear Physics B*, 190(2):412–424, 1981. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(81)90568-X. URL https://www.sciencedirect.com/science/article/pii/055032138190568X.

Gen Yin, Yufan Li, Lingyao Kong, Roger K. Lake, C. L. Chien, and Jiadong Zang. Topological charge analysis of ultrafast single skyrmion creation. *Phys. Rev. B*, 93:174403, May 2016. doi: 10.1103/PhysRevB.93.174403. URL https://link.aps.org/doi/10.1103/PhysRevB.93.174403.

M. Lakshmanan. The fascinating world of the landau–lifshitz–gilbert equation: an overview. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1939):1280–1300, 2011. doi: 10.1098/rsta.2010.0319. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2010.0319.

7 APPENDIX

Accuracy(%)	Number of observations 10	20	30	40	50
GPC (Ours)	0.094	**0.039**	**0.028**	**0.030**	**0.024**
GPR	0.139	0.055	0.051	0.052	0.033
SVC	**0.071**	0.055	0.043	0.047	0.052

Table 1: Accuracy of predicted phase diagrams for binary classification after each algorithm after different numbers of observations. Results are averaged over 6 runs. The lowest error for each step is highlighted in **bold**.

7.1 Baseline Details

We describe the two baselines here. Dai and Glotzer [2020] fit a Gaussian process (GP) is fit to binary observations $y = \{-1, +1\}$ and a heuristic acquisition function rates points, $\mathbf{x}$, to select, $A(\mathbf{x}) = \sqrt{var(y)}/(|E(y)| + \epsilon)$ where $var(y), E(y)$ and ϵ are the GP's variance, mean and a numerical stability factor. This acquisition function balances selecting points close to the currently predicted phase boundary where $E(y) \approx 0$ and regions of high uncertainty where $var(y)$ is large. The numerical stability factor prevents dividing by 0. Similar to our method, at every step, $A(\mathbf{x})$ is evaluated and the point with the highest value is selected as the next observation. Tian et al. [2020] presents a method for learning phase diagrams of 2+ phases in Method 2. However, this is not applicable to our setup since we cannot measure the location of phase transitions directly. Instead, we use a modified version of Method 1. A support vector classifier is fit to the observations of position and phase. Phase boundaries are predicted from classification boundaries. For active learning, we sample points along the phase boundary that have the highest weighted average distance from existing samples, which acts as a proxy for regions of the highest uncertainty. This is because support vector classifiers cannot directly quantify prediction uncertainty.

7.2 Fitting the Gaussian process classifier

GPCs are used to model a probabilistic event over space $\mathbf{x}$. At a point x, the probability of observing a positive outcome is $p(y = 1|\mathbf{x}) = \sigma(a(\mathbf{x}))$, where $a(\mathbf{x})$ is an unknown function, and σ maps $a(\mathbf{x})$ to a probability. Since $a(\mathbf{x})$ is unknown, so we model $a(\mathbf{x})$ with a Gaussian Process, f. Hence, $p(y = 1|\mathbf{x}) = \int f(a)\sigma(a)da$. To find the posterior distribution of $f(a|\mathbf{x}, \mathcal{T})$, we can show using Bayes rule that $f(a|\mathbf{x}, \mathcal{T}) = \frac{\sigma(a)f(a|\mathcal{T})}{\int P(y|a)f(a|\mathcal{T})da}$. By estimating $f(a|\mathcal{T})$ with the Laplace approximation and forcing the posterior to be Gaussian, it can be shown Rasmussen and Williams [2006] that the posterior can be written as $f(a|\mathbf{x}, \mathcal{T}) = \mathcal{N}(a|\mu_a, \sigma_a^2)$, where $\mu_a = K(\mathbf{X}, \mathbf{x})^T(\mathbf{Y} - \boldsymbol{\sigma}(\mathbf{a}))$ and $\sigma_a^2 = K(\mathbf{x}, \mathbf{x}) - K(\mathbf{X}, \mathbf{x})^T(\mathbf{W}^{-1} + K(\mathbf{X}, \mathbf{X}))^{-1}K(\mathbf{X}, \mathbf{x})$. Here $\mathbf{W} = diag(\boldsymbol{\sigma}(\mathbf{a})(1 - \boldsymbol{\sigma}(\mathbf{a}))$, the GP kernel is $\mathbf{K}()$, previous observations are $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{a}$ can be found numerically.

For a phase diagram of n phases, n GPCs are fitted, one to each phase. Given a dataset of observations $\{y_1...y_t\}$, the i'th GPC is fitted by mapping the observations to 1 if $y = i$ and 0 otherwise.

7.3 Implementation Details

All of our Gaussian processes, $f_i(x)$, use the stationary Matern 5/2 kernel, $K(\mathbf{x}, \mathbf{x'}) = K(r) = \sigma^2(1 + \sqrt{5}r/\rho + 5r^2/3\rho^2)exp(-\sqrt{5}r/\rho)$, where r is the distance between $\mathbf{x}$ and $\mathbf{x'}$. The kernel parameters are σ, the magnitude of correlation between observations, and ρ, setting the lengthscale two observations are correlated. These are fitted using maximum likelihood estimation. $p(a|\mathbf{y'})$ is fitted using the Laplace approximation using GPy SheffieldML [1.10]. For further details about the GPC and how we fit the GP to observations, we recommend Rasmussen and Williams [2006] and Krasser [2020].

To find $\text{argmax}_x A(x)$, we sample points on a grid (19×19 in our experiments). Optimizations can be made to speed up the search. Firstly, calculating the model's probability of observing a phase $p(P(\mathbf{x} = j))$ amounts to finding the probability a Gaussian is greater than multiple different

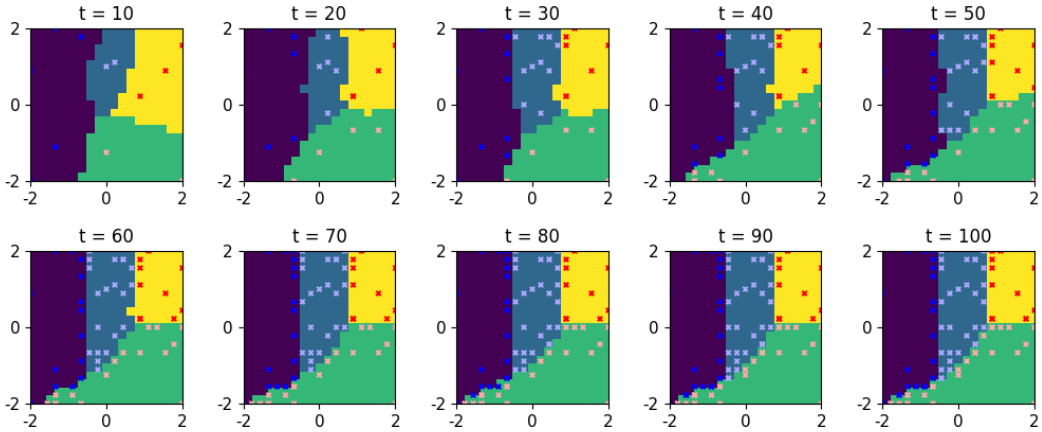

Figure A1: Observations and predictions of our model on the synthetic phase diagram of four phases after different numbers of observations. Observations are shown by crosses. The color indicates the observed phase.

Phase	Q	M
Spiral	low or 0	low
Skyrmion	high	high
Paramagnet	0	1

Table 2: Criterion used to assist the determination of three phases

Gaussians, $p(P(\mathbf{x}) = j) = p(f_j(x) > \{f_k(x)\}_{k \neq j})$. This can be done efficiently using the Gauss-Hermite quadrature instead of random sampling. Secondly, once we have probabilities of observing each phase, we only need to sample $A(x)$ at points where there is significant uncertainty, we set $A(\mathbf{x}) = 0$ if the probability of observing any single phase is greater than 90%. Thirdly, since we are using a local kernel for our Gaussian Process, the effect of new observations is local so $D(P_T, P_{T+1}(\mathbf{x}))$ can be calculated around $\mathbf{x}$ instead of the full area.

7.4 The criterion to determine phases of the honeycomb lattice

The phases under different parameter sets can be easily identified by observing real-space spin textures (fig). We have also calculated some physical quantities to precisely identify the different phases. We employ the discrete version of topological charge Q that identifies the existence of skyrmion phase. This version was first introduced in Berg and Lüscher [1981] and used in Yin et al. [2016]. The calculation of Q starts by triangulating the entire lattice and then counting the solid angels Ω_Δ for each triangle $\Delta(\boldsymbol{S_1}, \boldsymbol{S_2}, \boldsymbol{S_3})$ determined by

$$exp(i\frac{\Omega_\Delta}{2}) = \rho^{-1}[1 + \boldsymbol{S_1} \cdot \boldsymbol{S_2} + \boldsymbol{S_2} \cdot \boldsymbol{S_3} + \boldsymbol{S_3} \cdot \boldsymbol{S_1} + i\boldsymbol{S_1} \cdot (\boldsymbol{S_2} \times \boldsymbol{S_3})], \tag{4}$$

where $\rho = [2(1 + \boldsymbol{S_1} \cdot \boldsymbol{S_2})(1 + \boldsymbol{S_2} \cdot \boldsymbol{S_3})(1 + \boldsymbol{S_3} \cdot \boldsymbol{S_1})]^{1/2}$, then Q is given by summing over all of the triangles:

$$Q = \frac{1}{4\pi} \sum_\Delta \Omega_\Delta. \tag{5}$$

Here the solid angle Ω_Δ ranges from -2π to 2π and the branch cut is the negative real axis when calculating Ω_Δ. We also calculate standard magnetization:

$$M = \frac{1}{N} \sum_i S_i^z, \tag{6}$$

where N is the number of sites in the entire lattice.

In Table. 2 , we show the criterion we use to assist the determination of phases.

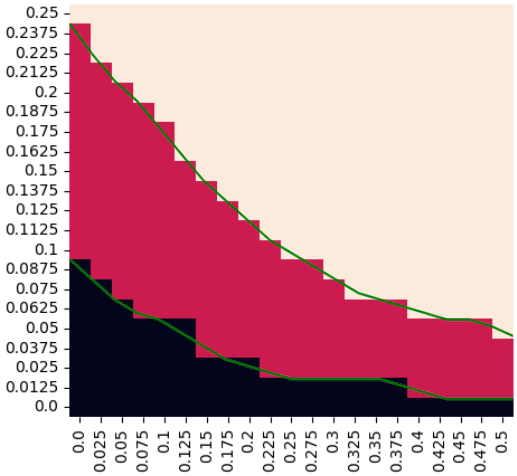

Figure A2: Phase diagrams of the honeycomb lattice as a functon of single-ion anisotropy D/J (x-axis) and the magnetic field h/J (y-axis) for fixed Dzyaloshinskii–Moriya interaction $d/J = 1/2$ and temperature $T = 10^{-4}J$ obtained by grid sampling. Black: spiral phase; red: skyrmion phase; pink: polarized paramagnet. Green lines represent the smoothed phase boundaries.

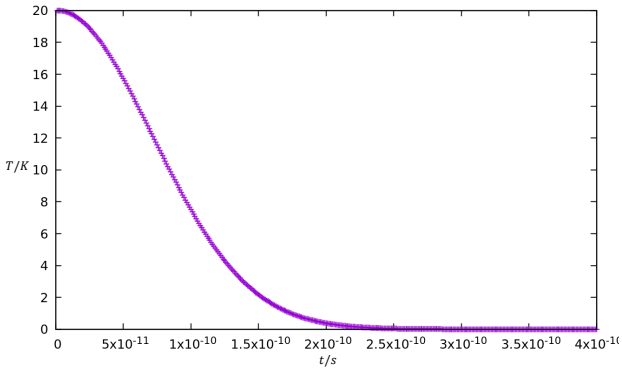

Figure A3: Temperature as a function of simulation time in our atomistic spin dynamics simulation.

7.5 Phase diagram obtained by grid sampling

In Fig. A2, we show phase diagram of the honeycomb lattice obtained by 21×21 grid sampling and the phase boundaries (smoothed). To simulate the spin configuration of the honeycomb lattice, we simulate the atomistic spin dynamics using in VampireEvans et al. [2014b]. Starting from a relatively high temperature, we simulate the evolution of spins using the Landau–Lifshitz–Gilbert (LLG) equation Lakshmanan [2011], cooling to set temperature we set with custom timesteps. In this work, we cool the system from $T = 20K$ to $T = 10^{-4}J(0.00116K)$ (Fig. A3), at each temperature during cooling process, the system evolves 10^4 steps with each step $10^{-16}s$, total 4×10^6 steps in the whole process of a simulation.