UNIVERSITY OF SOUTHAMPTON

DOCTORAL THESIS

Statistical modelling of skyrmion magnetic materials with defects

Author:
Vanessa NEHRUJI
(ORCID: 0000-0002-7606-7442)

Supervisor:
Dr. Ondrej HOVORKA
Co-supervisor:
Prof. Hans FANGOHR

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

Computational Engineering & Design Group
Engineering and Physical Science

February 1, 2025

Declaration of Authorship

I, Vanessa NEHRUJI, declare that this thesis titled, "Statistical modelling of skyrmion magnetic materials with defects" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

UNIVERSITY OF SOUTHAMPTON

Abstract

Computational Engineering & Design Group

Engineering and Physical Science

Doctor of Philosophy

Statistical modelling of skyrmion magnetic materials with defects

by Vanessa NEHRUJI

Skyrmions, a whirling magnetic texture, stabilised by the Dzyaloshinskii-Moriya interaction, emerge as promising candidates for next-generation data particles, owing to their compact size and topological stability. However, defects introduced during the manufacturing process can unpredictably affect the stability and dynamics of skyrmions, compromising data reliability. To address this, we conduct a comprehensive study using statistical physics-based modeling and machine learning data analysis to assess the impact of defects on the equilibrium properties of skyrmions. In our investigation, we uncover a novel disorder-driven continuous phase transition from a hexagonally-ordered arrangement of skyrmions (OSkL), in defect-free systems, to a disordered array of skyrmions (DSkL) at high defect levels. We pinpoint the OSkL-DSkL transition, representing the minimum defect level required to disrupt the OSkL, via the spin-spin correlation analysis. Due to the lack of appropriate order parameter formalism, we employ deep learning dimensionality reduction methods, which yield alternative transformations of the spin variables, to differentiate other phases present in such materials. In addition to phase classification, we use deep learning methods, based on the U-Net network, to estimate the Hamiltonian parameters and precisely identify defect locations. This more integrated approach, combining machine learning with numerical methods, provides insights into the complex interplay between defects and skyrmions, offering potential pathways for new experimentation and a wide range of technological applications.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Ondrej Hovorka, and Dr. Srinandan Dasmahapatra for their invaluable contributions and fruitful discussions on skyrmions and machine learning throughout my PhD journey, and also for their encouragement to persevere until the end. Their support has been instrumental in both shaping this thesis and helping me navigate through all the challenges of this PhD journey. I am particularly indebted to Ondrej for granting access to his mean-field numerical code, which I have used extensively in my research.

I gratefully acknowledge the financial support provided by the UK Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training under grant number EP/L006766/1. This funding has been crucial in allowing me to pursue this research with dedication and focus.

My heartfelt appreciation goes to my friends and family for their unconditional support over all these years, for giving me a space away from work and research, and for always lending a listening ear. I am especially grateful to my fiancé, Jacques, for his endless love and patience, which have been a constant source of strength for me throughout this journey. This thesis would not be possible without you.

Finally to my beloved mother, Shiyamala: though you are no longer here to witness this achievement, I will never forget your unwavering belief in me and I will carry that with me wherever I go. Your tireless dedication to my education, especially in my formative years, laid the foundation for this PhD, and I only wish you were here to see it. Your strength continues to inspire me, and I strive to honor your memory through my work. I hope one day to be as brilliant as you were, Amma.

Contents

List of Figures

List of Tables

In loving memory of my mum, Shiyamala

Chapter 1

Introduction

1.1 Motivation

Modern technological advancements increasingly rely on vast datasets, propelling the demand for robust hardware with enhanced data density and reduced energy consumption. Skyrmions, novel topological quasi-particles, offer a promising alternative to electric charge as primary data carriers in memory devices, owing to the reduced energy needed for writing, deleting and transfer of data via spin-polarised currents[1–3]. Their small sizes (mere nanometres[4, 5]) allow for high data density, and their topological protection ensures the non-volatility of data[6, 7].

The viability of a skyrmion-based device depends on the stability of skyrmions across the wide range of temperatures and fields required by analogous electronic devices[8–10]. Skyrmions primarily exist in a limited number of bulk low-symmetry crystals (B20 alloys[11–14]), where they are stable within a narrow range of fields and at very low temperatures. Optimising skyrmion stability over a wider range is feasible through various methods, including sample thickness reduction[12, 14, 15], pressure application[16], epitaxial strain introduction, and doping[17]—all of which introduce defects into the system.

Structural defects are particularly prevalent in sputtered thin films[18–20], which constitute the majority of skyrmion-hosting materials. These thin films are favoured for their ease of integration into multilayer heterostructures[21], allowing for fine-tuning of skyrmion properties through interactions with other layers[18, 20, 22–34].

However, the random and uncontrolled nature of these defects often leads to undesirable consequences. These include the suppression of the skyrmion phase[23], disruption of lattice tight-packing (resulting in diminished data density)[20, 30, 32–34], pinning of skyrmions to the crystal lattice impeding their mobility within the material (a critical aspect for skyrmion-based devices)[20, 22–24, 32], and even the annihilation of skyrmions, jeopardising data preservation. Defects, when controlled, can be exploited to create novel device ideas[35–38] or stabilise metastable skyrmions[13, 16, 17, 39, 40], but when unwanted, they can have unpredictable behaviour on the spin textures. Through statistical modelling of the defect distribution, our research aims to expand our understanding on the impact of defects on the skyrmion lattice phase.

1.2 Our Research Contribution

The majority of prior studies have primarily focused on investigating the mechanisms behind the pinning of skyrmions with defects stemming from atomic origins, such as atomic vacancies, or local variations in exchange, Dzyaloshinskii-Moriya interaction (DMI), and anisotropy strength. The effects of defects have been also noted in several experimental studies[18, 41]. However, the research is still in its infancy, largely due to difficulty in simulating and analysing the spin texture of all potential defect distributions. Therefore, instead of examining the effects of individual defects, this thesis explores the role of entire statistical distributions of defects on the formation of skyrmion phases. To facilitate effective computational modelling, we first developed a lattice-resolved mean-field model of the classical Heisenberg Hamiltonian with Dzyaloshinskii–Moriya interaction and random pinning field, which we validated by comparing it to experimental measurements of Fe_3GeTe_2 two-dimensional van der Waals magnet. We subsequently used this model in conjunction with pattern-searching deep learning models to analyse the hysteretic behaviour of skyrmion materials across different levels of defects.

The defects in our material are modelled as a spatially varying random pinning field (H_σ), following a Gaussian distribution[42]. External magnetic field sweeps,

within a 'clean' system, with very low pinning field distribution, reveal an instantaneous switch from the uniform state to an ordered skyrmion lattice (OSkL). Conversely 'dirty' systems, with significant pinning field distribution, display a gradual proliferation of skyrmions, forming instead a disordered array of skyrmions (DSkL). We argue that this shift from an abrupt emergence of skyrmion lattice to a smooth appearance of isolated skyrmion clusters as the magnetic field is swept along a hysteresis loop highlights the existence of a novel disorder-driven phase transition from OSkL to DSkL, with a critical disorder of $H_{\sigma,crit}$. This phase transition is signified by the divergence of the spin-spin correlation mapping between configurations along a field sweep at the skyrmion nucleation boundary. Further analysis shows that larger skyrmions are more sensitive to the defect field, suggesting that systems with high DMI (resulting in smaller skyrmions) are likely to achieve higher data density due to the tighter packing within an ordered lattice.

While spin-spin correlation mapping provides a rough estimate of the OSkL-DSkL transition boundary, its applicability is limited to the skyrmion nucleation boundary, being overly sensitive to minor configuration changes within a phase. Moreover, no other known property or order parameter can discern all phases within the entire magnetic-field parameter space. Hence, we used dimensionality reduction techniques, such as **variational autoencoders** (VAE)[43], to efficiently derive a set of macroscopic properties from spin variables that maximises phase information. VAE effectively clusters most phases in materials with defects, providing a more comprehensive method to identify phases and accurately pinpoint the transition boundaries.

Despite the random nature of defects induced by sputtering, supervised training of a neural network enables the development of a model capable of identifying defect locations in new materials. If the interactions between a skyrmion and defect are fully understood, then this approach can offer valuable insights into skyrmion stability and dynamics, without the need for extensive experimental testing. Through adaptation of the U-Net model[44], we can effectively extract both the underlying defect field and all additional Hamiltonian parameters, whether they vary locally at each lattice site or are applied globally, even in polycrystalline thin film materials.

1.3 Thesis Outline

In this thesis, the first three chapters serve as an introduction to the field, offering a summary of magnetic modelling, including an outline of the specific computational methods used for both simulation and data analysis, and a broad background on the current status of skyrmion research. The four concluding chapters present the main outcomes of our research. The initial chapter validates the mean-field model against experimental work carried out on 2D FGT flakes. The next chapter focuses on characterising the impact of defects on magnetic configurations, including the identification of a potential OSkL-DSkL phase transition. The final two chapters explore the application of deep learning techniques for phase classification and the estimation of Hamiltonian parameters from materials with defects.

Chapter 2 covers the theoretical background to understand the equilibrium behaviour of magnetic systems. It explores fundamental models such as the Ising and Heisenberg models, detailing the intrinsic interactions among magnetic moments and the incorporation of defects into these models. Finally, it discusses the two numerical methods used in our research for finding ground state spin configurations: the Monte Carlo method and mean field approximation.

Chapter 3 presents an overview of machine learning (ML), providing an in-depth explanation of neural networks, the training process, and introduces various layers that potentially reduce network complexity. Furthermore, it discusses the relevant architectures and models used in our study, including the U-Net for defect detection (see Chapter 8), and dimensionality reduction techniques for phase classification (see Chapter 7).

Chapter 4 introduces the concept of a magnetic skyrmion, and provides an overview of the key developments in skyrmion research. It underscores the challenges posed by material defects, particularly in thin film skyrmionic materials, and outlines our specific contributions to understanding the impact of these defects.

In **Chapter 5**, we validate the mean field model by reproducing similar spin configurations and field-sweep behaviour of the 2d magnet $Fe_{3-x}GeTe_2$. Our simulations correctly generates Néel skyrmions that are primarily stabilised by dipolar interactions, and accurately recovers the transitions between the monodomain,

skyrmion and stripe domain phases at various levels of Fe composition.

The impact of defects on equilibrium states is explored in **Chapter 6**, where we note a blurring of first-order transitions at increased defect levels, characterised by a gradual proliferation of skyrmions in the system. Using the spin-spin correlation mapping between configurations along the field sweep, we identify the critical transition boundary between OSkL and DSkL and analyse its behaviour as a function of skyrmion size.

Rather than manually computing macroscopic parameters such as skyrmion numbers, or counting the number of reciprocal space peaks to identify the various phases, **Chapter 7** uses unsupervised deep learning methods to find latent variables that contain the maximum information about a phase. This quasi-order parameter could differentiate between all phases present in defect materials, other than between the OSkL and DSkL due to a limitation of convolutional layers.

Chapter 8 uses the U-Net architecture to predict the random pinning field for an unknown spin texture to a high degree of accuracy. The architecture was then modified to predict any combination of local and global Hamiltonian parameters, even in polycrystalline materials featuring regions of varying thickness. The key components of U-Net that render such predictions feasible include convolutional layers, max-pool layers, the symmetric structure of the network, and most importantly, the incorporation of skip connections.

Finally, **Chapter 9** summarises the key outcomes of our research and discusses future improvements and applicability of the work.

Chapter 2

Background in Magnetism

The following chapter introduces the theoretical concepts required to understand phases and the transitions between them, starting from a macroscopic standpoint using thermodynamics, before delving into the microscopic models used in statistical mechanics. It covers key topics including the Ising model (order-disorder phase transition), Heisenberg model, defect modelling and, numerical methods (specifically equilibrium states) concluding with discussions on techniques for identifying skyrmions and skyrmion lattices within ground state spin configurations.

2.1 Magnetic Phases

Though predating the discovery of the atom, thermodynamics has emerged as a foundational framework capable of accurately characterising the behaviour of many-body systems in thermal equilibrium. Despite the myriad interactions among its innumerable particles, their collective dynamics converge to only a limited number of thermodynamic states, commonly referred to as macrostates. This macroscopic state can be effectively captured using only a minimal set of state quantities, for instance the internal energy of a system can be derived solely from its entropy and net magnetization, $U = U(S, \boldsymbol{M})$[1]. Thermodynamics is thus able to describe the macroscopic behaviour of complex systems, even in the absence of detailed microscopic information.

The fundamental equation of thermodynamics encapsulates the energy flow within a system, attributing changes in internal energy to the transfer of heat or work.

[1] Thermal equilibrium occurs when the internal energy is minimised at a fixed S and $\boldsymbol{M}$, or equivalently, from maximising the entropy of an isolated system with a fixed internal energy and net magnetisation, $S = S(U, \boldsymbol{M})$

Specifically, in the case of work induced by a magnetic field, $\boldsymbol{H}$, the first law is given as follow,

$$dU = \delta Q + \delta W = TdS + \mu_0 \boldsymbol{H} \cdot d\boldsymbol{M} \tag{2.1}$$

where dU, δQ, and δW are the changes in internal energy, heat and applied work on the system. T is the temperature, dS is the change in entropy, μ_0 is the magnetic permeability, $\boldsymbol{H}$ is the external field and $d\boldsymbol{M}$ is the change in net magnetisation of the sample. However, the practical application of this equation is limited, notably due to the inherent difficulty in manipulating the quantities $(S, \boldsymbol{M})$ within an experimental setting. Consequently, they are transformed into $(T, \boldsymbol{H})$ using Legendre transforms, resulting in the Helmholtz free energy[2],

$$dF(T, \boldsymbol{H}) = -SdT - \mu_0 \boldsymbol{M} \cdot d\boldsymbol{H}. \tag{2.3}$$

From this transformation, the system's magnetisation can be computed at fixed temperature as follows,

$$\boldsymbol{M} = -\left(\frac{1}{\mu_0}\frac{\partial F}{\partial \boldsymbol{H}}\right)_T. \tag{2.4}$$

The state space encompasses all possible values of the minimal set of state quantities permissible for the system under consideration, with each point in this space representing a unique macrostate. A phase is classified as a region over the state space where the system's macroscopic properties exhibit minute variation. Conversely, transitions between phases result in abrupt and discontinuous changes. One of the most extensively studied phase transitions in magnetic systems is from a paramagnetic state (where spins are randomly oriented) at high temperatures to a ferromagnetic state (with all spins aligned in the same direction) at low temperatures. The length of the reduced magnetisation, m, jumps from 0 to 1 between the disordered and ordered phase as temperature decreases. As it distinguishes the ordered state from the disordered state, m is termed the order parameter.

[2]This is the equivalent of minimising the internal energy under the constraints of constant temperature and field,

$$F(T, \boldsymbol{H}) = U(S, \boldsymbol{M}) - TS - \mu_0 \boldsymbol{H} \cdot \boldsymbol{M} \tag{2.2}$$

Landau derives the order parameter through a series expansion of the Helmholtz free energy in terms of m, yielding the following relationship,

$$F(T, m) = F_0 + a(T - T_c)m^2 + bm^4 + \ldots, \tag{2.5}$$

where a and b are the coefficients of the Taylor expansion and $a, b > 0$. Only even powers of m are considered due to the up-down symmetry of the magnetised state. Equilibrium occurs when the potential is at a minimum, therefore,

$$\left(\frac{\partial F}{\partial m}\right)_T = 2a(T - T_c)m + 4bm^3 = 0. \tag{2.6}$$

The solutions to the above equation are given below,

$$m = 0, \qquad m^2 = -\frac{a(T - T_c)}{2b}. \tag{2.7}$$

The first condition holds true when $T \geq T_c$, corresponding to the paramagnetic phase. For the second condition, the ferromagnetic state has $m = 1$ at $T = 0$, resulting in a transition temperature of $T_c = \sqrt{2b/a}$. T_c is the Curie temperature, denoting the boundary between the ordered and disordered phases.

Although thermodynamics offers a macroscopic description of magnetic materials under the effect of external conditions, including at transition boundaries, it lacks a microscopic reasoning for such behaviours. This can only be found in statistical mechanics, which applies probability theory to a representation of the microscopic model and can determine macroscopic properties via statistical averages. The simplest of such models is the Ising model.

2.1.1 Ising Model

The Ising model is the earliest microscopic model for ferromagnetic materials, and was proposed by Ising in 1925 for the purpose of understanding the paramagnetic-ferromagnetic phase transition[45]. The magnetic material is represented by a set of equally-spaced points on a hypercubic lattice of dimension d[3]. Each lattice point is assigned a spin and the collection of all spins in the lattice is denoted by set Λ. The

[3] In other words, a 1D system is a chain of lattice points, 2D system is a plane of points, etc.

spins themselves are constrained to take discrete values, $s_i \in \{-1, 1\}$ for all $s_i \in \Lambda$. A spin configuration is a lattice with all spins assigned to appropriate values, $\Sigma = (s_i)_{i\in\Lambda}$. The interactions between N spins is encoded in following the Hamiltonian,

$$\mathcal{H}(s) = -\frac{1}{2}J\sum_{i}^{N}\sum_{j\in i_{nn}} s_i s_j - \mu_0 H\sum_{i}^{N} s_i. \tag{2.8}$$

The first term is the pairwise exchange interaction between neighbouring spins, s_i and s_j, with a strength of J. The factor of $\frac{1}{2}$ is introduced to prevent double-counting of spin interactions. When $J > 0$, the ordered state is ferromagnetic, with adjacent spins aligned parallel to one another, while $J < 0$ favors an antiferromagnetic ordering with spins aligning anti-parallel. The sum is taken only over the nearest neighbours of a spin, $j \in i_{nn}$, as this is a short-ranged effect. The second term is a uniform external field H acting on every spin, s_i, on the lattice.

All potential spin configurations form a set, Ω. In the canonical ensemble, where the system is in contact with a heat bath[4], the probability measure over Ω follows the Boltzmann distribution. The probability of a specific microstate $\mu \in \Omega$, at a macrostate energy of E_μ, can be determined from this distribution as follows,

$$p(\mu) = \frac{e^{-\beta E_\mu}}{Z} = \frac{e^{-\beta \mathcal{H}(\mu)}}{Z}, \tag{2.9}$$

where $\beta = \frac{1}{k_B T}$, and Z is the partition function[5], calculated as a sum of the Boltzmann factors over all accessible states in Ω:

$$Z = \sum_{\mu\in\Omega} e^{-\beta \mathcal{H}_\mu}. \tag{2.10}$$

Macroscopic properties can be computed using p_μ as ensemble averages. For example, net magnetisation, m, is given as,

$$\langle m\rangle = \sum_{\mu} m_\mu p_\mu = \sum_{\mu}\sum_{i}^{N} s_{i,\mu}\frac{1}{Z}e^{-\beta\mathcal{H}_\mu} \tag{2.11}$$

[4]The system is at fixed temperature.
[5]This is a normalising factor such that $p(\Omega) = 1$.

where $\langle m \rangle$ is the ensemble average of m. It further provides a microscopic description for the Helmholtz free energy,

$$F = -\frac{1}{\beta} \ln Z. \tag{2.12}$$

The paramagnetic-uniform phase transition is a second-order (continuous) transition which should show discontinuities in $\frac{\partial^2 F}{\partial H^2}$. However, in a 1D chain of spins, the exact computation of the free energy reveals no non-analyticities, indicating an absence of a finite-temperature phase transition. Consequently, Ising deemed the model inadequate for representing real magnetic materials and abandoned it. Onsager later refuted this claim, demonstrating the existence of this phase transition in a 2D system using transfer matrices, and analytically determining a critical temperature of $T_c = 2/\ln(1+\sqrt{2}) = 2.269$ for $J = 1$[46]. Numerical methods, such as the Monte Carlo discussed later in Section 2.4.1, further validate the existence of a phase transition, showing a discontinuity in the thermally averaged susceptibility, $\chi = \frac{\partial^2 F}{\partial H^2}$. Example configurations at various temperatures are also illustrated in Figure 2.1A.

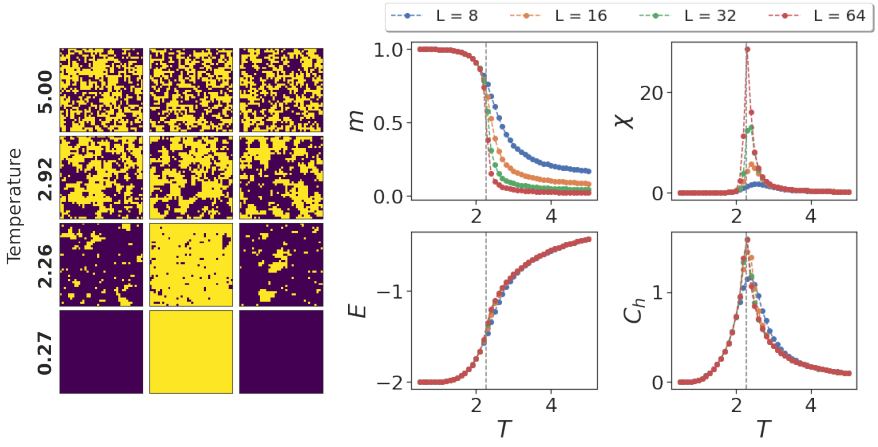

FIGURE 2.1: a) Example Ising configurations at a range of temperatures $T \in [0.25, 4.0]$ when $J = 1$, generated with Monte Carlo simulations, at lattice size, $L = 40$. b) Thermal statistics of the average magnetisation, m, the magnetic susceptibility, χ, energy, E, and specific heat, C_h, at different lattice sizes, averaged over 1000 Monte Carlo steps and 4000 realizations. The grey dashed line shows the critical temperature of $T_c = 2.2691$.

2.2 Heisenberg Model

The Heisenberg model expands the Ising model by allowing more degrees of freedom in the spin definition. The Heisenberg spin, $\boldsymbol{s}_i$, is represented as a unit vector, $|\boldsymbol{s}_i| = 1$, allowing for rotations in SO(3). Consequently, a wider range of interactions can be incorporated into the Hamiltonian, with the main ones highlighted below,

$$\mathcal{H} = \mathcal{H}_{\text{exch}} + \mathcal{H}_{\text{dm}} + \mathcal{H}_{\text{ani}} + \mathcal{H}_{\text{ext}} + \mathcal{H}_{\text{demag}} \tag{2.13}$$

The first term is the isotropic exchange, the second term is an antisymmetric exchange term called the Dzyaloshinskii-Moriya (DM) interaction, the third term is the anisotropy of the material, the fourth term is the external field, and the last term is the effect of the demagnetising field caused by the magnetism of the moments. The remainder of this section discusses these terms in more detail.

2.2.1 Isotropic Exchange

The Heisenberg system, like the Ising model, has a spin-spin pairwise interaction, resulting in ferromagnetic behaviour in the material. The exchange is a quantum mechanical effect that arises due to the indistinguishable nature of the system's unpaired electrons, forcing the total wavefunction to be antisymmetric on the exchange of any two electrons. Due to this symmetry and the Pauli Exclusion principle, stating that no two particles can occupy the same quantum state within an atom, two electrons can either couple in a spin-singlet ($S = 0$) or a spin-triplet ($S = 1$) state. The energy due to Coulomb repulsion between the electrons can be determined solely from the spin component of the wavefunction, leading to the following interaction[47],

$$\mathcal{H}_{\text{exch}} = -\frac{1}{2} J_{ex} \sum_{i}^{N} \sum_{j \in i_{nn}} (\boldsymbol{s}_i \cdot \boldsymbol{s}_j). \tag{2.14}$$

In ferromagnetism, the triplet spin state is preferred and $J_{ex} > 0$, leading to a parallel alignment of spins. Conversly, antiferromagnetism prefers the singlet state, leading to $J_{ex} < 0$ and an anti-parallel alignment. The preferred spin configuration of these different states are shown in Figure 2.2. Here, we model isotropic exchange as a short

range interaction, applicable only when the electron orbitals overlap, and therefore the sum is performed only over the nearest neighbours of $\boldsymbol{s}_i$.

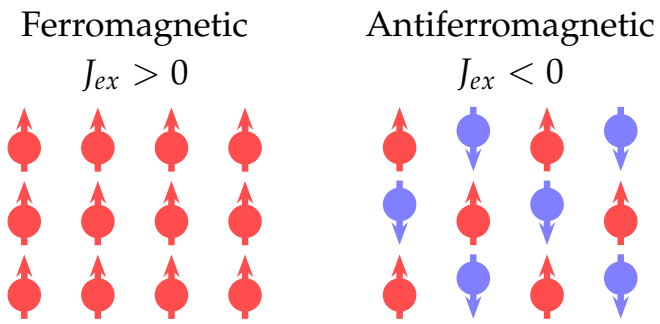

FIGURE 2.2: Example spin configurations showing the ferromagnetic state with $J_{ex} > 0$ (left), and the antiferromagnetic state with $J_{ex} < 0$ (right).

2.2.2 Dzyaloshinskii-Moriya Interaction

Historically, typically antiferromagnetic crystals, such as α-Fe_2O_3[48], $MnCO_3$ and $CoCO_3$, exhibit "weak" ferromagnetism, unexplained by the Hamiltonian interactions of the time. In 1958, Dzyaloshinskii used Landau's theory of symmetry arguments to introduce a new term in the free energy expansion, which favours canted spin configurations[49]. This phenomenological treatment could, unfortunately, only specify that $\boldsymbol{D}$ was parallel to the trigonal axis of α-Fe_2O_3, and was unable to explain its origins, nor how to calculate the strength of $\boldsymbol{D}$.

Moriya later provided a microscopic justification, expanding on Anderson's theory of superexchange, with spin-orbit coupling, in which the largest contribution in anisotropic superexchange has the same antisymmetric form derived by Dzyaloshinskii[50]. In the atomistic Hamiltonian, this appears as,

$$\mathcal{H}_{\text{dm}} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\in i_{nn}} \boldsymbol{D}_{ij}\cdot\left(\boldsymbol{s}_i \times \boldsymbol{s}_j\right), \tag{2.15}$$

where $\boldsymbol{D}_{ij} = J_{dm}\boldsymbol{D}_e$ is the DM vector whose strength (J_{dm}) and orientation ($\boldsymbol{D}_e$) are determined by the neighbouring vectors. If the lattice structure is known, then the $\boldsymbol{D}_{ij}$ vector between atoms i and j can be derived using Moriya's rules[50]. Figure 2.3 shows certain orientations of the DM vector and the skyrmion structures they produce.

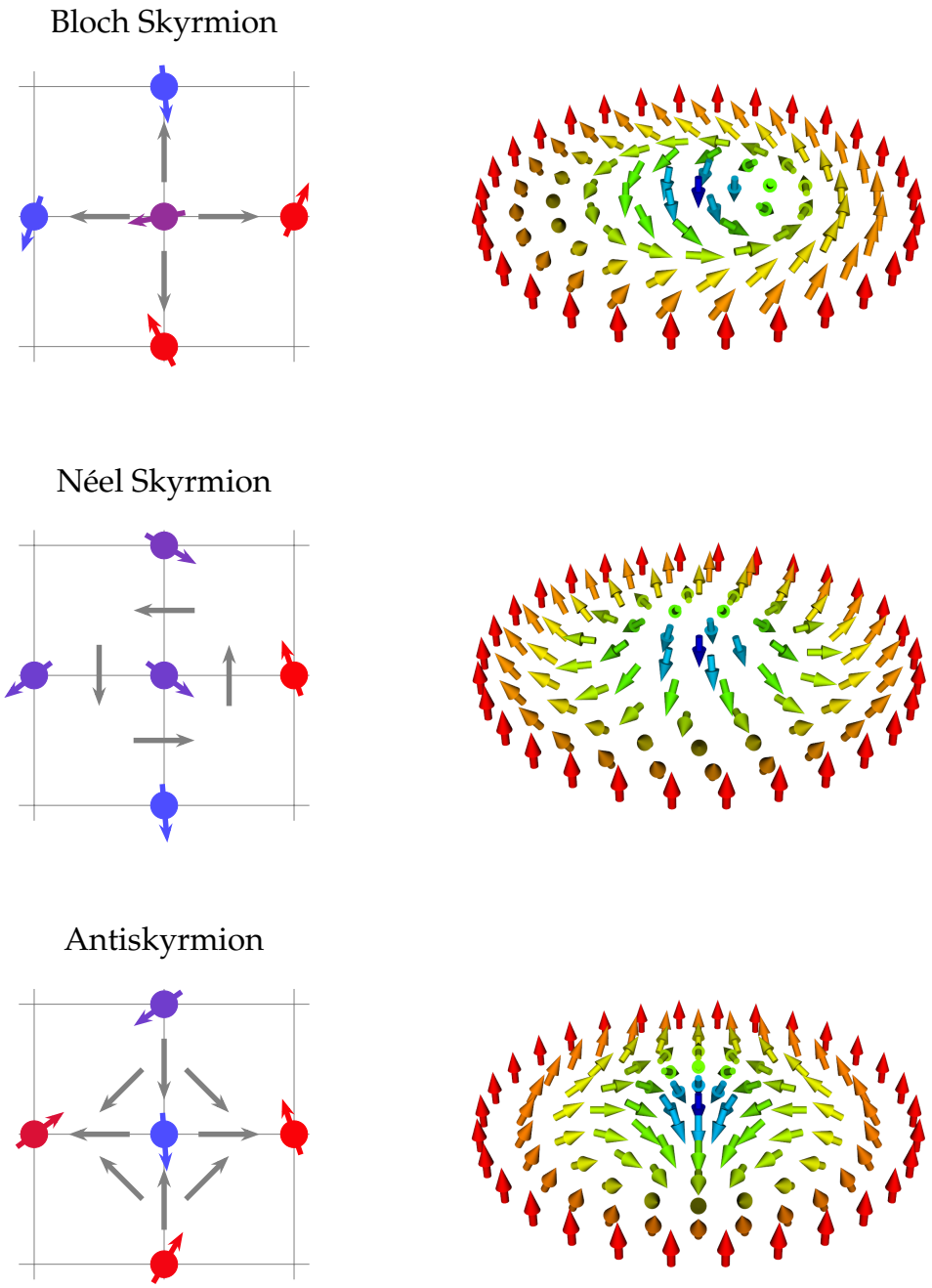

FIGURE 2.3: Dzyaloshinkskii-Moriya vectors (grey) are given for spins on a square lattice, and the corresponding skyrmion structure is plotted on the right. a) Bloch skyrmions are created with radial DM vectors, b) Néel with azimuthal vectors, and c) Antiskyrmions require a mixture of radial and azimuthal DM vectors.

The DM interaction must be included when modelling non-centrosymmetric materials[6]. Although it is weak in comparison to isotropic exchange, particularly in bulk materials[51], it is responsible for the creation of canted magnetic structures

[6]Materials with no inversion symmetry.

such as the helical, conical and skyrmion structures[52]. For this reason, many studies have focused on techniques that increase the strength of this interaction, including by artificially inducing inversion symmetry breaking at the interface between a ferromagnetic layer over a non-magnetic heavy metal layer[23, 25] (see Figure 2.4).

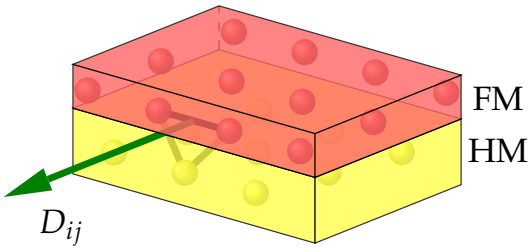

FIGURE 2.4: Interfacial Dzyaloshinskii-Moriya interaction due to symmetry breaking between a ferromagnetic layer and a heavy metal substrate. Reproduced from Figure 4.5.

In our implementation, $\boldsymbol{D}_e$ lies radially along the vector connecting two spins, r_{ij}, but this can be rotated to allow for other skyrmion types.

2.2.3 Magnetocrystalline Anisotropy

Magnetocrystalline anisotropy is an intrinsic property that stems from the shape and crystalline structure of the material, causing spins to prefer alignment to a particular direction (easy axis), or in a particular plane (easy plane). Anisotropy is the energy required to shift the alignment of a spin vector from the easy axis to a hard axis; from first-principles, it is the energy needed to overcome the barrier caused by spin-orbit coupling[53].

Spin-orbit coupling is a relativistic effect, whereby the motion of an electron about the nucleus creates a magnetic field that acts on the electron, but only in the rest frame of the electron. Consequently, a magnetic moment in that field has an energy of $\Delta E = -\boldsymbol{\mu} \cdot \boldsymbol{B} \propto \boldsymbol{L} \cdot \boldsymbol{S}$, where the field is proportional to the orbital angular momentum, $\boldsymbol{B} \propto \boldsymbol{L}$ and the moment to the spin angular momentum, $\boldsymbol{\mu} \propto \boldsymbol{S}$. When spin direction changes, the tightly coupled orbital axis attempts to follow. The electronic orbitals, however, are also coupled to the underlying lattice via the crystal field, leading to a competition between the orbital-lattice and spin-orbital coupling[54].

Anisotropy, therefore, arises in two ways: if the orbital-lattice coupling dominates the spin-orbital coupling, then $\boldsymbol{L}$ is tightly coupled to the lattice and anisotropy stems from the spin-orbital energy, $\propto \boldsymbol{L} \cdot \boldsymbol{S}$; in contrast, if spin-orbital coupling dominates, then $\boldsymbol{L}$ couples to $\boldsymbol{S}$, and anisotropy emerges from the orbital-lattice energy where the orbital momentum magnitude varies with the magnetisation direction[55, 56].

In a hexagonal or tetragonal system, there is a single axis of high symmetry which can be atomistically modelled as,

$$\mathcal{H}_{\text{ani}} = -J_k \sum_{i}^{N} (\boldsymbol{s}_i \cdot \boldsymbol{e}_i)^2, \tag{2.16}$$

where J_k is the strength of the anisotropy, and $\boldsymbol{e}_i$ is the high symmetry axis. This is also known as uniaxial anisotropy. It is particularly crucial to model the effects of anisotropy in finite-sized systems or systems with reduced dimensions such as thin films. In these cases both the strength of the spin-orbit coupling and the magnitude of the orbital momentum is increased, resulting in an stronger anisotropy effects[57].

2.2.4 External Field

Applying an external Zeeman field to a magentic system causes spins to align preferentially along the axis of the field. This applies uniformly on every spin, $\boldsymbol{s}_i$, as follows,

$$\mathcal{H}_{\text{ext}} = -\mu \sum_{i}^{N} \boldsymbol{H} \cdot \boldsymbol{s}_i, \tag{2.17}$$

where μ is the magnetic moment.

2.2.5 Demagnetising Field

A spin is a magnetic dipole which generates its own magnetic field. The interaction between a spin and the magnetic field of its neighbouring spins is the dipolar interaction, with the sum of these interactions being the demagnetising field. In the atomistic model, the dipolar interaction is calculated as follows,

$$\mathcal{H}_{\text{demag}} = -\frac{J_{\text{demag}}}{2} \sum_{ij} -\frac{\boldsymbol{s}_i \cdot \boldsymbol{s}_j}{r_{ij}^3} + 3\frac{(\boldsymbol{s}_i \cdot \hat{\boldsymbol{r}}_{ij})(\boldsymbol{s}_j \cdot \hat{\boldsymbol{r}}_{ij})}{r_{ij}^3}. \tag{2.18}$$

Here, J_{demag} is the strength of the dipolar interaction, r_{ij} is the distance between spins i and j, and the sum is taken over all other spins in the material. Due to the high computational expense[7] and to avoid artificial anisotropies, the calculation is simplified such that only neighbours ascribed within a circle of radius r_c of spin i are considered (see Figure 2.5).

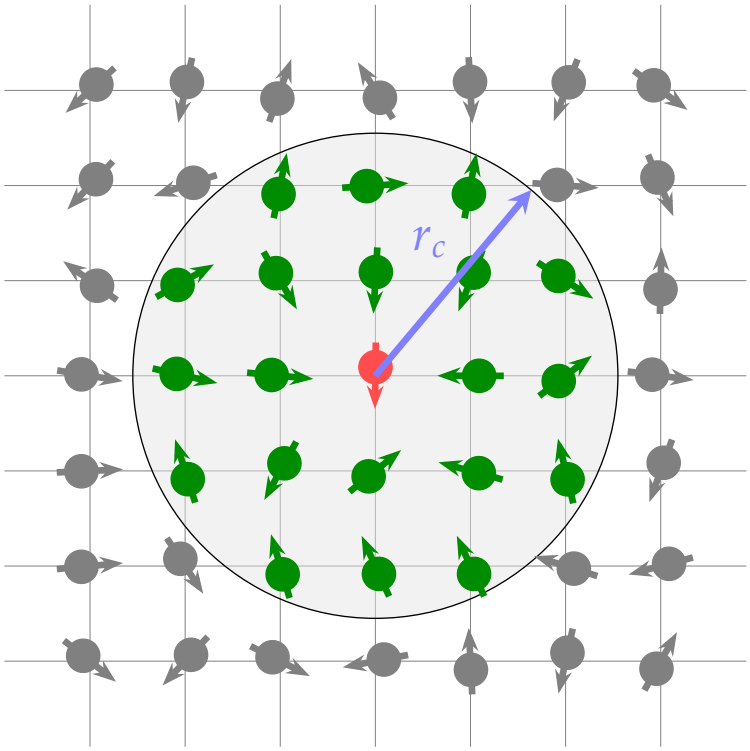

FIGURE 2.5: The large grey circle is the cutoff radius, r_c, when calculating the dipolar interactions for the red spin. The demagnetising field is the sum of all pairwise interactions between the red spin and each of the green spins.

The demagnetising field prefers an alignment of spins that minimises the total self-magnetisation of the material and plays a significant role in thin film materials. It competes with the effects of isotropic exchange and uniaxial anisotropy[8], favouring in-plane spins due to the factor $(\boldsymbol{s}_i \cdot \hat{\boldsymbol{r}}_{ij})$. The sum over the various dipolar neighbours divides the magnetic material into domains aligned along the longest axis to minimise the energy, see Figure 2.6.

[7] $O(n^2)$ complexity
[8] If $\boldsymbol{e} = (0, 0, 1)$, this would prefer spin aligned out of plane.

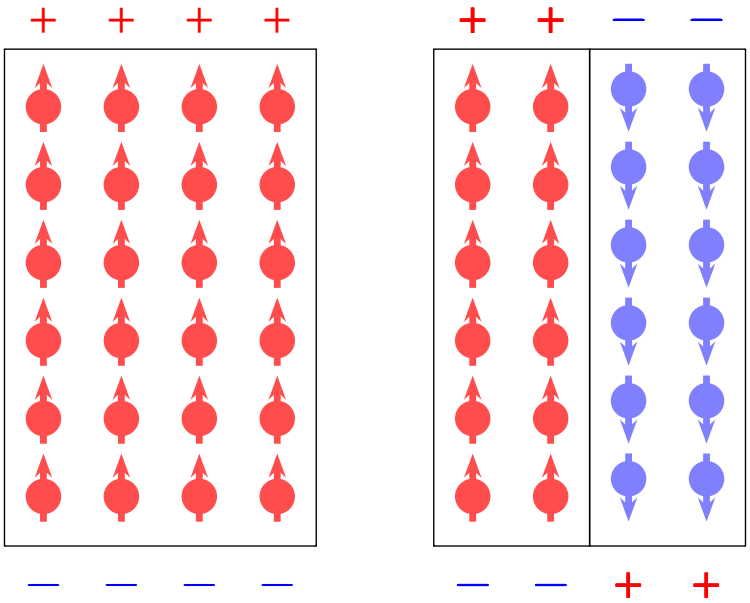

FIGURE 2.6: Given a magnetic material that is uniformly magnetised, the dipolar interaction seeks to minimise the self-magnetisation of the body by separating into domains along the longest axis.

2.3 Modelling Defects

Defects play a crucial role in shaping the properties of materials, significantly influencing the interaction strength between dipoles and, consequently, the resulting ground-state configurations. These imperfections in crystal structures can be classified according to their dimensionality, ranging from zero-dimensional point defects to three-dimensional precipitates. Point defects, which are the focus of this study, include vacancies (missing atoms), interstitials (atoms occupying non-lattice positions), and substitutions (foreign atoms at lattice or interstitial sites)[58, 59]. While extended defects such as edge/screw dislocations or planar boundaries exist[58, 59], their modeling is more complex[59], hence this research primarily concentrates on the distribution of point defects. These defects can occur accidentally during material growth or be intentionally introduced through methods like irradiation. The presence of defects can significantly alter material properties, and this study aims to investigate how they specifically affect the ground state spin configurations in magnetic materials.

Defects are prevalent in thin film heterostructures as a result of the the fabrication process[13, 15, 17, 18, 20, 22–34, 39]. There are two primary methods used

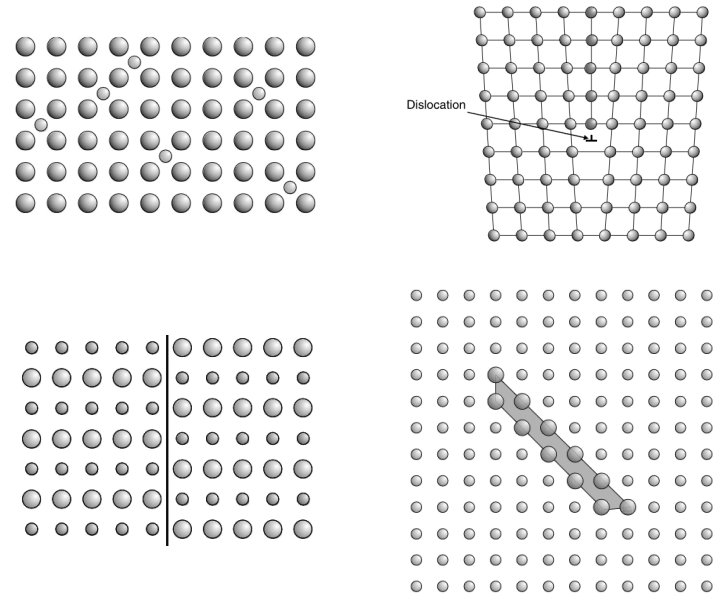

FIGURE 2.7: Examples of defects in crystalline solids. a) Point defects (interstitials), b) line defects (edge dislocation), c) planar defects (antiphase boundary), and d) volume defects (precipitates).

for growing thin layers: molecular beam epitaxy (MBE), which preserves crystalline structure but requires longer deposition time, and sputtering, which deposits layers quickly but with less regard for structure. Sputtering is the most common growth process, but resultant thin films are typically amorphous or at best polycrystalline, lacking the single-crystalline structure[9](see Figure 2.8). Despite this, skyrmions can still emerge in all these diverse thin film types[61–67].

In atomistic simulations, defects can be modelled by either modifying the lattice structure itself (e.g. removing certain sites) or by altering the spin interactions. The earliest example of defect modelling is Random Field Ising Model (RFIM), where a site-dependent pinning field (H_i) is added to each spin[68, 69],

$$\mathcal{H}_{\text{RFIM}} = -\frac{J}{2}\sum_{i}^{N}\sum_{j\in i_{nn}} s_i s_j - \mu\sum_{i} H_i s_i. \tag{2.19}$$

Here, s_i are Ising spins, and H_i is a Gaussian random variable drawn from $\mathcal{N}(0, H_\sigma)$, with H_σ characterising the level of defects in the system. Even minor additions of this defect type can significantly influence the phases observed in clean systems,

[9]Polycrystalline films contain small grains (also called crystallites) of differing sizes and random orientations, rather than the perfectly-aligned structure of single crystalline material[60]. Amorphous films have a higher level of disorder, with no long-range geometric order in the structure[60].

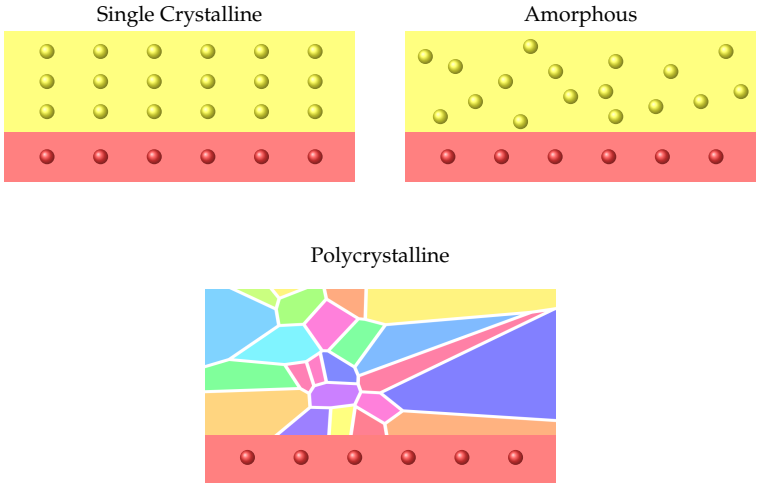

FIGURE 2.8: Types of thin film structures: a) single crystalline, where the thin film is stacked regularly (yellow) over the substrate (red); b) amorphous, where atoms are distributed in completely random fashion; c) polycrystalline, where atoms form grains with different sizes and orientations (generated via voronoi tessellation).

hence referred to as strong disorder[10][68, 70]. For instance, the paramagnetic-ferromagnetic phase transition, analytically proven to occur in 2D systems by Onsager, is suppressed at even low levels of defects ($\sigma \ll J$)[11]. Additionally, it has been rigorously proven that random fields destroy long-range order in systems with dimensions $d \leq 2$ for discrete order parameter systems (Ising), and $d \leq 4$ for continuous order parameter systems (Heisenberg)[68, 72, 73], if H_i is symmetric under rotations.

In our research, defects are similarly modelled in Heisenberg systems with a spatially-dependent random pinning (RP) field, with randomness constrained to only vary in only the z-direction, $\boldsymbol{H}_i = H_{z,i}\hat{\boldsymbol{z}}$, with $H_{z,i} \sim \mathcal{N}(0, H_\sigma)$. This appears as an additional term in the Heisenberg Hamiltonian (Equation 2.13),

$$\mathcal{H}_{RP,\sigma} = \mu_0 \sum_i^N \boldsymbol{H}_i \cdot \boldsymbol{s}_i. \tag{2.20}$$

Here, $\boldsymbol{H}_i$ is a quenched field that is fixed throughout the simulation and H_σ controls the level of defects in the system. When $J = 1$, a H_σ value of 0.001 approximates a 'clean' system, while $H_\sigma = 0.1$ represents a 'dirty' system with a high defect

[10] There is also weak disorder, where the form of disorder simply modifies critical transition points or the order of the transition.

[11] When $\sigma \gg J$, the preference for disorder is higher than for an ordered state (J), and it has been proven exactly that low temperature states are disordered[71].

level. Research on the 2D XY model demonstrated that small variations of this constrained defect field results magnetic structures with long-range order[74]. Likewise, in our scenario, the $\boldsymbol{H}_i$ should also function as a weak disorder that simply alters the boundaries of the skyrmion lattice phase[12]. Although these random field systems were previously used to study diluted antiferromagnetic systems[75] and Jahn-Teller systems[76, 77], they are typically avoided due to the difficulty in creating a spatially varying external field in real environments[78]. Random anisotropy[42, 79–82], and random bond defects in DMI[18, 20], and exchange strength[83] present a more realistic representation of defects in the crystal field.

2.4 Numerical Methods

In this section, we review two of the main methods used to model equilibrium magnetic states at finite temperature: the Monte Carlo (MC) algorithm and the mean field (MF) approximation. The objective for both methods is to compute macroscopic properties of the ground state configuration, given a particular Hamiltonian. As discussed in Section 2.1.1, a macroscopic property, A, is given as,

$$\langle A \rangle = \sum_{\mu} A_{\mu} p_{\mu}, \tag{2.21}$$

where p_μ is the probability of microstate μ, ($p_\mu = e^{-\beta \mathcal{H}(\mu)}/Z$). This requires the computation of the partition function, Z.

Unfortunately, brute force calculation of Z is not feasible due to the uncountable number of possible configurations, $\mu \in \Omega$. Instead, alternative methods are used to compute Z, including Monte Carlo which uses importance sampling to select the most important configurations, and the mean field approximation which simplifies Z by decoupling the pairwise spin interaction.

2.4.1 Monte Carlo

The Monte Carlo method cleverly selects samples from the entire sample set, which are the most likely to correspond to low energy states: this is known as importance

[12] This is explored further in Chapter 6.

sampling[84, 85]. It does this by generating a Markov chain, where the future state depends on the current one and not on prior history, and each successive state is drawn from a probability distribution that converges to the Gibbs distribution, p_μ.

In our numerical simulation, we want to find the average thermal parameters, $\langle X \rangle$, of the lowest energy state. This can be computed by the usual numerical methods for integration, such as the Simpson's method or via Gaussian quadrature, but the high-dimensional nature of state μ results in a very inefficient algorithm. Instead we use Monte Carlo methods, where the integrand is separated into the random variable (RV) of interest, $X_\mu \in \Omega$, and the probability of its occurrence, $p(\mu)$. Then we draw M independent and identically distributed (IID) RVs from this probability distribution function (pdf), $X_1, ..., X_M \sim p$. The thermal average is thus approximated by the estimator,

$$\bar{X}_M = \frac{1}{M} \sum_i^N X_i \overset{\text{a.s}}{\to} \langle X \rangle, \tag{2.22}$$

which tends to the average value, $\langle X \rangle$, as $M \to \infty$ from the law of large numbers.

Unfortunately, we cannot sample directly from this target distribution in sufficient time[13]. We instead draw dependent samples which in the long run eventually draws from the target distribution, $\pi(x)$. One set of RVs which has been proven to have a pdf that converges to the target distribution is the Markov chain.

A set of IID RVs $X_i \in \Omega$, is a Markov chain if it satisfies the Markov property,

$$\mathbb{P}\left(X_{i+1} = \omega_k | X_i = \omega_j\right), \tag{2.23}$$

where the probability, $\mathbb{P}$, of reaching X_{i+1} only depends on the current state X_i. $\omega_{k,j}$ are the specific states of the RVs X_{i+1} and X_i. This Markov chain is time-homogeneous if the transition probabilities stays constant over i,

$$\mathbb{P}\left(X_{i+1} = \omega_k | X_i = \omega_j\right) = \mathbb{P}\left(X_1 = \omega_a | X_0 = \omega_b\right). \tag{2.24}$$

The probability of going from any state ω_j to any state ω_k can be more succinctly written in matrix form with elements $P_{jk} = \mathbb{P}\left(X_1 = \omega_k | X_0 = \omega_j\right) = P(\omega_j \to \omega_k)$,

[13] The target distribution, p_μ, contains a normalization factor which integrates over the entire sample space.

and in the Markov chain, the pdf is updated successively with this transition matrix,

$$p_1(\omega_k) = \sum_{j,k\in\Omega} p_0(\omega_j)P(\omega_j \to \omega_k). \tag{2.25}$$

Thus after M steps in the chain, the probability distribution is $p_M = p_{M-1}P = p_0P^M$. To use this Markov chain to produce thermal averages, the transition matrix must have a stationary distribution[14], whereby the pdf is invariant under multiplication with the transition matrix, $\pi = \pi P$. This is the target distribution from which we originally wanted to sample.

To generate this target pdf, we first rewrite the stationarity condition as follows,

$$\sum_{j\in\Omega} \pi(j)P(j \to k) = \pi(k) = \sum_{k\in\Omega} \pi(k)P(k \to j), \tag{2.26}$$

where the final equality was obtained by multiplying the distribution by $\sum_{k\in\Omega} P(k \to j) = 1$. This condition is sufficiently satisfied by,

$$\pi(j)P(j \to k) = \pi(k)P(k \to j) \qquad \forall j,k \in \Omega. \tag{2.27}$$

This is known as the detailed balance equation, and states that rate of flow from state j to k is equal to the flow from k to j. The Metropolis-Hastings algorithm[86] uses this equation to generate a suitable Markov chain.

The algorithm works similarly to the simpler MC acceptance-rejection algorithm, in that a newly drawn state, Y, is sampled from $\mathbb{P}(Y = y|X_i = x) = Q(x \to y)$, where Q is the probability of choosing y given x, and this is accepted with probability $A(x \to y)$. If it is accepted, then we set $X_{i+1} = Y$, otherwise it is rejected and the state is unchanged, $X_{i+1} = X_i$. This states that the total probability of transitioning from state x to state y is,

$$P(x \to y) = \mathbb{P}(y \text{ is chosen}|x \text{ AND } y \text{ is accepted}|x) \tag{2.28}$$

$$= Q(x \to y)A(x \to y) \tag{2.29}$$

[14]This requires that transition matrix is both irreducible and aperiodic.

To ensure this satisfies detailed balance, we require that, given acceptance ratio Q, A must satisfy

$$\pi(x)Q(x \to y)A(x \to y) = \pi(y)Q(y \to x)A(y \to x). \tag{2.30}$$

For an efficient algorithm, we preferably want to set the acceptance rate as high as possible in order to accept many states and increase convergence, but since A is a probability it is subject to the limit $A \leq 1$. Therefore $\max\left(A(x \to y), A(y \to x)\right) \leq 1$. Equation 2.30 can be rearranged as,

$$A(x \to y) = \frac{\pi(y)Q(y \to x)}{\pi(x)Q(x \to y)}A(y \to x) = \lambda\pi(y)Q(y \to x), \tag{2.31}$$

where we have introduced the factor $\lambda = \frac{A(y \to x)}{\pi(x)Q(x \to y)}$. From detailed balance, this is equal regardless of the direction of flow, $\lambda = \lambda(x, y) = \lambda(y, x)$. The limit of A can be used to find the maximal value of λ,

$$\lambda \ \max\left(\pi(y)Q(y \to x), \pi(x)Q(x \to y)\right) = 1 \tag{2.32}$$

$$\lambda = \frac{1}{\max\left(\pi(y)Q(y \to x), \pi(x)Q(x \to y)\right)} < \infty. \tag{2.33}$$

Substituting the expression for $\lambda(x, y)$, we find,

$$A(x \to y) = \min\left(1, \frac{\pi(y)Q(y \to x)}{\pi(x)Q(x \to y)}\right). \tag{2.34}$$

This is the Metropolis-Hastings acceptance ratio and it implies that new states with higher probability are preferred, $\pi(y) > \pi(x)$, as well as new states where it is easier to return to the previous state, $Q(y \to x) > Q(x \to y)$. The ratio allows us to use the unnormalised version of the target distribution, $\frac{\pi(y)}{\pi(x)} = \frac{\pi_\mu(y)}{\pi_\mu(x)}$, and we are therefore not required to calculate the partition function, Z. The term within the acceptance ratio is known as the Hastings ratio,

$$\frac{\pi(y)Q(y \to x)}{\pi(x)Q(x \to y)}, \tag{2.35}$$

and this must be specified for different models.

The general algorithm for the Markov Chain Monte Carlo (MCMC) method is thereby given in Algorithm 1. Here, we see that if the Hastings ratio, R, is greater than a random number drawn from the uniform distribution, $\mathbf{U}[0,1]$, then the new state is accepted, otherwise it is rejected.

Algorithm 1 Metropolis Hastings

procedure MC STEP
  $y \sim Q(x \to y)$
  $R = \frac{\pi_\mu(y)Q(y\to x)}{\pi_\mu(x)Q(x\to y)}$
  $U \sim \mathbf{U}[0,1]$
  if $U < R$ **then**
    $x_{i+1} \leftarrow y$ ▷ accept the state
  else
    $x_{i+1} \leftarrow x_i$ ▷ reject the state
  end if
  goto *top*.
end procedure

Applying this to the simple Ising model, a new state, y, is picked by randomly choosing a spin on the lattice and flipping it, thus the probability of picking a new state is $Q(x \to y) = \frac{1}{N}$. This transition matrix is symmetric, $Q(x \to y) = Q(y \to x) = \frac{1}{N}$, and so the Hastings ratio simplifies to,

$$\frac{\pi_\mu(y)}{\pi_\mu(x)} = \exp\left[-\beta(\mathcal{H}(y) - \mathcal{H}(x))\right] = \exp\left[-\beta dE\right], \tag{2.36}$$

where dE is the change in energy from the spin flip.

The Heisenberg model is likewise similar but the transition matrix is the joint probability of choosing a spin on the lattice and choosing a new value for the spin vector, $Q(x \to y) = \frac{1}{N} \cdot \mathbb{P}(\text{choose } \boldsymbol{s}_i = \{s_{ix}, s_{iy}, s_{iz}\}, |\boldsymbol{s}_i| = 1)$. The only limitation in our choice of Q is that it is symmetric so that the Metropolis Hastings algorithm is kept simple, but otherwise we would ideally choose Q such that it samples high probability states[15](for faster convergence). Our program uses the method developed by Hinzke and Nowak[87], which integrates three separate spin kick choices into a single algorithm. These three choices are outlined below, where $\boldsymbol{s}_i^t$ and $\boldsymbol{s}_i^{t+1}$ refers to the spin vectors at lattice position i, at current and future MC steps, t and $t+1$,

[15]This is known as importance sampling.

1. **Spin inversion.** Signs of all spin components are inverted,

$$\boldsymbol{s}_i^{t+1} = \{-s_{ix}, -s_{iy}, -s_{iz}\} \tag{2.37}$$

2. **A random spin on the unit sphere.** There are numerous ways to generate this vector, but an efficient method involves sampling the components from the standard normal distribution, $\mathcal{N}(0,1)$ and then normalizing the result,

$$\{s_{jx} \sim \mathcal{N}(0,1), s_{jy} \sim \mathcal{N}(0,1), s_{jz} \sim \mathcal{N}(0,1)\} \tag{2.38}$$

$$\boldsymbol{s}_i^{t+1} = \frac{1}{|\boldsymbol{s}_j|}\{s_{jx}, s_{jy}, s_{jz}\} \tag{2.39}$$

3. **A random spin within a small cone about the axis of the original spin.** This is achieved by finding a new random vector, $\boldsymbol{s}_j$, as in method 2, adding a small fraction, γ of that vector to the old spin vector, $\boldsymbol{s}_i^t$, and then normalizing,

$$\boldsymbol{s}_i^{t+1} = \frac{1}{|\boldsymbol{s}_i^t + \gamma \boldsymbol{s}_j|}(\boldsymbol{s}_i^t + \gamma \boldsymbol{s}_j) \tag{2.40}$$

There is an equal probability of choosing any of the above three methods. Typically in most MC algorithms only method 2 is used, but the Hinzke-Nowak method results in a faster convergence rate, since it allows the simulation to gradually reach a low energy state through the cone step, but does not let the system freeze in a metastable state by using the spin flip and random spin selection[87]. The full technique is outline in Algorithm 2.

Once the system has relaxed through a sufficient number of steps to reach equilibrium and ensure sampling from the correct target distribution (π), thermally averaged properties of the material can be computed. The two notable properties are the system energy, E and the average magnetization, M. These are trivially calculated by finding the local change in energy/magnetism, dX, after a spin kick, and adding this to the prior value, X^{t-1} to find the global value, X^t. Then these global values are averaged across all MC steps to find the ensemble average $\langle X \rangle$. Other parameters of interest include the square of energy and magnetism, E^2 and M^2, which can be used to compute the variance of the E and M, respectively known as the specific heat and

susceptibility of the material,

$$C_h = \frac{\beta^2}{N}\Big(\langle E^2\rangle - \langle E\rangle^2\Big), \tag{2.41}$$

$$\chi = \beta N\Big(\langle M^2\rangle - \langle M\rangle^2\Big). \tag{2.42}$$

These properties can be used to identify the Curie temperature since they show a point of discontinuity as a function of β as $N \to \infty$. The final property of interest in the skyrmion number of the system: in systems with the DM interaction, N_{sk} can be used to identify phase regions which have skyrmionic structures (see Section 2.5). The full algorithm for the MCMC method is given in Algorithm 3.

Algorithm 2 Heisenberg MCMC: Hinzke-Nowak Spin Kick

procedure SPINKICK
  $U \sim \mathbf{U}_D[0,2]$
  if $U = 0$ **then** ▷ spin flip
    return $\boldsymbol{s}_i^{t+1} = \{-s_{ix}, -s_{iy}, -s_{iz}\}$
  else if $U = 1$ **then** ▷ random spin
    `generate` $\boldsymbol{s}_j = \{s_{j\alpha} \sim \mathcal{N}(0,1) \text{ for } \alpha \in \{x,y,z\}\}$
    return $\boldsymbol{s}_i^{t+1} = \frac{\boldsymbol{s}_j}{|\boldsymbol{s}_j|}$
  else if $U = 2$ **then** ▷ random spin in cone
    `generate` $\boldsymbol{s}_j = \{s_{j\alpha} \sim \mathcal{N}(0,1) \text{ for } \alpha \in \{x,y,z\}\}$
    $\boldsymbol{s}_j = \boldsymbol{s}_i + \gamma \boldsymbol{s}_j$
    return $\boldsymbol{s}_i^{t+1} = \frac{\boldsymbol{s}_j}{|\boldsymbol{s}_j|}$
  end if
end procedure

Algorithm 3 Heisenberg MCMC: Metropolis Hastings

procedure MC STEP
  for `spin_index in N` **do**
    $i \sim \mathbf{U}_D[0, N-1]$
    $\boldsymbol{s}_i^t \leftarrow$ `Find old spin at index,` i
    $\boldsymbol{s}_\alpha \leftarrow$ `Get a new random spin` ▷ Spin Kick, Alg. 2
    $dE \leftarrow$ `Find local energy change`
    $U \sim \mathbf{U}[0,1]$
    if $dE < 0$ **then**
      $\boldsymbol{s}_i^{t+1} \leftarrow \boldsymbol{s}_\alpha$ ▷ accept the state
    else if $\exp(-\beta dE) < U$ **then**
      $\boldsymbol{s}_i^{t+1} \leftarrow \boldsymbol{s}_\alpha$ ▷ accept the state
    else
      $\boldsymbol{s}_i^{t+1} \leftarrow \boldsymbol{s}_i^t$ ▷ reject the state
    end if
  end for
end procedure

2.4.2 Mean Field

The mean field (MF) approximation decouples the pairwise spin-spin interactions to an interaction with an average field of all spins. For example, in the Ising model at zero external field, the energy from the interaction between spin s_i and its z neighbouring spins can be decoupled as follows,

$$\mathcal{H}_i = -\frac{1}{2}J\sum_{j\in i_{nn}} s_i s_j = -\frac{1}{2}(Jzs_i)\cdot\frac{1}{z}\sum_{j\in i_{nn}} s_j, \tag{2.43}$$

where $\frac{1}{z}\sum_{j\in i_{nn}} s_j$ can be interpreted as the average interaction from the nearest neighbours. This is a local magnetisation density that varies across the lattice, but MF replaces this with a global density, $\frac{1}{N}\sum_j^N s_j$. Consequently, the partition function simplifies to a system of non-interacting spins, with spins interacting with a global mean field generated by their neighbours. From this, we obtain a self-consistency equation for magnetisation.

$$m = \tanh\left(\beta(Jzm + \mu H)\right) \tag{2.44}$$

This is known as the Curie-Weiss (CW) method[88, 89] and its full derivation is given in the Appendix, Chapter A.1. For more complex Heisenberg systems, CW does not result in a partition function that can be analytically computed, and therefore we introduce the Hubbard Statonovich (HS) method that is more generisable to complex systems.

The Hubbard-Stratonovich Transformation

For simplicity, we consider a system with isotropic exchange, DMI and and external field,

$$\mathcal{H} = -\frac{1}{2}\sum_i^N\sum_{j\in i_{nn}}\left[J_{ij}(\boldsymbol{s}_i\cdot\boldsymbol{s}_j) + \boldsymbol{D}_{ij}\cdot(\boldsymbol{s}_i\times\boldsymbol{s}_j)\right] - \mu\sum_i^N \boldsymbol{H}_i\cdot\boldsymbol{s}_i). \tag{2.45}$$

This leads to the partition function as follows,

$$\begin{aligned} Z &= \sum_{\boldsymbol{s}_k} \exp\Big(-\beta(-\frac{1}{2}\sum_{i}^{N}\sum_{j\in i_{nn}}\left[J_{ij}(\boldsymbol{s}_i\cdot\boldsymbol{s}_j)+\boldsymbol{D}_{ij}\cdot(\boldsymbol{s}_i\times\boldsymbol{s}_j)\right]-\mu\sum_{i}^{N}\boldsymbol{H}_i\cdot\boldsymbol{s}_i)\Big), \\ &= \mathrm{Tr}_{\boldsymbol{s}_k}\left(\exp\left(\beta(\frac{1}{2}(\boldsymbol{s}_i\cdot\mathbb{I}\cdot\boldsymbol{s}_j)-\mu\boldsymbol{H}_i\cdot\boldsymbol{s}_i)\right)\right). \end{aligned} \tag{2.46}$$

Here, the sum is taken over all spin configurations, $\boldsymbol{s}_k$, and $\mathbb{I}$ is a general matrix containing the symmetric isotropic exchange and antisymmetric DMI,

$$\mathbb{I}_{ij} = \begin{bmatrix} J_{ij}^{xx} & D_{ij,z}^{xy} & -D_{ij,y}^{xz} \\ -D_{ij,z}^{yx} & J_{ij}^{yy} & D_{ij,x}^{yz} \\ D_{ij,y}^{zx} & -D_{ij,x}^{zy} & J_{ij}^{zz} \end{bmatrix}. \tag{2.47}$$

The spin-spin interaction, $\boldsymbol{s}_i\cdot\mathbb{I}\cdot\boldsymbol{s}_j$, prevents any further simplification of Z, and therefore we use the Hubbard Stratonovich (HS) transformation to decouple this.

HS uses a mathematical identity of Gaussian integrals to transform the nearest neighbour interaction term into a simpler coupling between non-interacting spins and a new auxiliary field, $\boldsymbol{x}_i$. The Gaussian identity can be summarised as follows,

$$\frac{1}{\sqrt{|\boldsymbol{A}|}}e^{\frac{1}{2}\boldsymbol{b}_i(\boldsymbol{A}^{-1})_{ij}\boldsymbol{b}_j} = \int_{-\infty}^{\infty}\prod_{i}^{N}\frac{\exp\left[-\frac{1}{2}\boldsymbol{x}_i\cdot A_{ij}\cdot\boldsymbol{x}_j+\boldsymbol{x}_i\cdot\boldsymbol{b}_i\right]}{\sqrt{2\pi}}d\boldsymbol{x}_i, \tag{2.48}$$

where $\boldsymbol{A}$ is a real symmetric matrix with determinant $|\boldsymbol{A}|$, $\boldsymbol{b}_i$ is an arbitrary vector at i, which is coupled to a fluctuating auxiliary field $\boldsymbol{x}_i$. Applying this results in the following partition function,

$$Z = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{|\beta\mathbb{I}|}}\int_{-\infty}^{\infty}\prod_{i}^{N}\exp\Big(-\frac{1}{2}\boldsymbol{\psi}_i\cdot(\beta^{-1}\mathbb{I}_{ij}^{-1})\cdot\boldsymbol{\psi}_j\Big)\mathrm{Tr}_{\boldsymbol{s}_k}\exp\Big((\boldsymbol{\psi}_i+\mu\beta\boldsymbol{H}_i)\cdot\boldsymbol{s}_i\Big)d\boldsymbol{\psi}_i. \tag{2.49}$$

Z now describes the interaction of non-interacting spins, $\boldsymbol{s}_i$, with a fluctuation field,

$\boldsymbol{\psi}_i$. Up until this point, the partition function remains in its exact form with no approximations made. We rewrite this in terms of an action, $\mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \mu\boldsymbol{H}_i)$, as follows,

$$\begin{aligned} Z &= \frac{\beta^{\frac{3}{2}}}{\sqrt{|\mathbb{I}|}} \int_{-\infty}^{\infty} \exp\Big(-\beta\mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \mu\boldsymbol{H}_i)\Big) d\Omega \\ \mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \mu\boldsymbol{H}_i) &= \frac{1}{2}\boldsymbol{\phi}_i \cdot \mathbb{I}_{ij}^{-1} \cdot \boldsymbol{\phi}_j - \frac{1}{\beta} \ln\Big[\mathrm{Tr}_{\boldsymbol{s}_k} \exp\Big(\beta\left(\boldsymbol{\phi}_i + \mu\boldsymbol{H}_i\right) \cdot \boldsymbol{s}_i\Big)\Big]. \end{aligned} \tag{2.50}$$

Further simplification requires the saddle point approximation, which assumes that the integral is dominated by the maximum value of the integrand, and therefore approximates the partition function by the maximum of $\exp(-\beta\mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \mu\boldsymbol{H}_i))$. This is equivalent to minimising the action $\mathcal{L}$, which is found by setting $\frac{\partial\mathcal{L}}{\partial\boldsymbol{\phi}_i}$ to zero. After a rearrangement, a set of self-consistent equations is found for the average local magnetic moment, $\bar{\boldsymbol{m}}_i$. A full derivation is provided in the Appendix, Section A.2.2, but the main results are summarised below,

$$\bar{\boldsymbol{m}}_i = \mathbf{L}(\beta\mu|\bar{\boldsymbol{H}}_i^e|)\frac{\bar{\boldsymbol{H}}_i^e}{|\bar{\boldsymbol{H}}_i^e|}, \tag{2.51}$$

$$\mu\bar{\boldsymbol{H}}_i^e = J_{ij}\bar{\boldsymbol{m}}_j - \boldsymbol{D}_{ij} \times \bar{\boldsymbol{m}}_j - \mu\boldsymbol{H}_i, \tag{2.52}$$

where $\mathbf{L}(x) = \coth(x) - x^{-1}$ is the Langevin function, and μ is set to 1. $\bar{\boldsymbol{H}}_i^e$ is the effective field acting on each mean-field spin, $\bar{\boldsymbol{m}}_i$, including all interactions specified in the mean-field Hamiltonian[16]

$$\mathcal{H}_{MF} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\in i_{nn}}\left[J_{ij}(\bar{\boldsymbol{m}}_i \cdot \bar{\boldsymbol{m}}_j) + \boldsymbol{D}_{ij} \cdot (\bar{\boldsymbol{m}}_i \times \bar{\boldsymbol{m}}_j)\right] - \mu\sum_{i}^{N}\boldsymbol{H}_i \cdot \bar{\boldsymbol{m}}_i \tag{2.53}$$

Equations 2.51 and 2.52 are a set of N coupled equations which must be solved iteratively as follows:

1. The system is initialized in a well-defined state: either the uniform state at high field, or a paramagnetic state at high temperatures.

2. At each field and temperature point on the phase diagram, we perform the following steps:

[16]The MF Hamiltonian can be trivially derived as the Curie-Weiss model, whereby atomistic spins are replaced by their mean field equivalent, $\boldsymbol{s}_i = \bar{\boldsymbol{m}}_i + \delta\boldsymbol{s}$, and terms beyond the first-order are neglected. Derivation of the self-consistent equations for $\bar{\boldsymbol{m}}_i$, however, requires the field-theoretic Hubbard-Stratonovich method.

(a) Calculate the effective field for each lattice spin, $\bar{\boldsymbol{H}}_i^e$ (Equation 2.52).

(b) Find the new magnetic moment, $\bar{\boldsymbol{m}}_i$, using $\bar{\boldsymbol{H}}_i^e$ (Equation 2.51).

(c) Update the old spin with a small fraction of the newly calculated moment[17], $\boldsymbol{s}_i^{t+1} = \boldsymbol{s}_i^t + \alpha(\bar{\boldsymbol{m}}_i - \boldsymbol{s}_i^t)$

(d) Calculate the error, $err = ||\boldsymbol{s}_i^{t+1} - \boldsymbol{s}_i^t||^2$

(e) If this error if below a specified tolerance value, then end the simulation. Otherwise, repeat from step a.

3. Store the spin configuration of the ground state for future analysis.

The MF step has been outlined with more clarity in Algorithm 4.

Algorithm 4 Heisenberg MF: Iterative Step

while `err` $>$ `tol` **do**
  `err` $= 0$
  for $i \in N$ **do**
    `calculate` $\bar{\boldsymbol{H}}_i^e$ ▷ Equation 2.52
    `calculate` $\bar{\boldsymbol{m}}_i(\bar{\boldsymbol{H}}_i^e)$ ▷ Equation 2.51
    $\boldsymbol{s}_i^{t+1} = \boldsymbol{s}_i^t + \alpha(\bar{\boldsymbol{m}}_i - \boldsymbol{s}_i^t)$ ▷ calculate new spin
    `err` $+= \left(\boldsymbol{s}_{i,\gamma}^{t+1} - \boldsymbol{s}_{i,\gamma}^t\right)^2$ for $\gamma \in \{x, y, z\}$ ▷ find `err`
  end for
end while

The computational resources required by the algorithm significantly depends on the convergence of the method. Despite this, MF is generally faster compared to MC, and therefore investigation into magnetic materials typically commence with an MF exploration to identify key regions of the phase diagram. Accurately identifying the boundaries of these regions, however, requires a comprehensive Monte Carlo study.

Initial Conditions

In our simulations, we run hysteresis loops, whereby the system starts in a perfectly uniform spin-down state at an external field of $H = -1$, and then the field is gradually increased until the system reaches the spin-up uniform state. When running these loops, the initial uniform state must include a small variation from the perfectly aligned state when the Hamiltonian has no term preferring canted spins (i.e.

[17] New spin only uses a fraction of the new moment in order to control the rate of convergence. If α is too high, the system would jump sporadically about the true minima without settling, but with too low an α, simulations take a long time to converge.

the Dzyaloshinskii-Moriya interaction). A perfectly aligned state can lead to constrained dynamics[18], and the dipolar term in Equation 2.18 works only partially.

When starting from a perfectly aligned state, only the first term contributes to the energy minimisation, and directly competes with isotropic exchange. The second term, which prefers in-plane spins, reamins at zero, with no other terms to cant spins away from $(0, 0, 1)$ such that it is no longer zero. Our implementation adds a random tilt in a cone of 0.0001° to the initial uniform state, and a random variation to the uniaxial axis in a cone of 0.1°, in case the system relaxes back to the perfectly uniform state.

2.5 Skyrmion Number

The skyrmion number of a spin system is the simplest method of detecting topologically non-trivial spin structures. In the continuum limit, it is calculated as follows,

$$N_{sk} = \frac{1}{4\pi} \iint d^2 r \boldsymbol{m} \cdot \left(\frac{\partial \boldsymbol{m}}{\partial x} \times \frac{\partial \boldsymbol{m}}{\partial y} \right), \tag{2.54}$$

where N_{sk} describes the number of times m can be wrapped around the sphere S^2. It is possible to simply discretise this on a square lattice as follow,

$$N_{sk} = \frac{1}{4\pi} \sum_x \sum_y m_{x,y} \cdot \left(\frac{m_{x+1,y} - m_{x,y}}{\Delta x} \times \frac{m_{x,y+1} - m_{x,y}}{\Delta y} \right) \Delta x \Delta y. \tag{2.55}$$

But a naïve discretisation leads to small numerical errors and, therefore, N_{sk} is no longer quantised.

Berg and Luscher developed an analytic equivalent of Equation 2.54 for lattice models[90]. In this method, the lattice is first triangulated, and then the solid angles are calculated for each triangle[90, 91]. Our implementation only considers opposite triangles, $(\boldsymbol{s}_i, \boldsymbol{s}_1, \boldsymbol{s}_2)$ and $(\boldsymbol{s}_i, \boldsymbol{s}_3, \boldsymbol{s}_4)$ to avoid double counting (see Figure 2.9 for triangles in a square and hexagonal lattice)[92]. The solid angle is best described as the area on the sphere which is enclosed by the three vectors $(\boldsymbol{s}_i, \boldsymbol{s}_1, \boldsymbol{s}_2)$ (final image in

[18] States where only the m_z component can vary.

Figure 2.9), and can be found as,

$$\exp\left(\frac{i\Omega}{2}\right) = \frac{1}{\rho}(1 + \boldsymbol{s}_1 \cdot \boldsymbol{s}_2 + \boldsymbol{s}_2 \cdot \boldsymbol{s}_3 + \boldsymbol{s}_3 \cdot \boldsymbol{s}_1 + i\boldsymbol{s}_1 \cdot (\boldsymbol{s}_2 \times \boldsymbol{s}_3)), \tag{2.56}$$

where $-2\pi < \Omega < 2\pi$ and ρ is the normalisation factor,

$$\rho = \sqrt{2(1 + \boldsymbol{s}_1 \cdot \boldsymbol{s}_2)(1 + \boldsymbol{s}_2 \cdot \boldsymbol{s}_3)(1 + \boldsymbol{s}_3 \cdot \boldsymbol{s}_1)}. \tag{2.57}$$

The lattice topological charge is given by sum of the spherical areas of the triangles, (s_i, s_1, s_2) and (s_i, s_3, s_4) over all lattice points, i,

$$Q = \sum_i q(i) = \sum_i \frac{1}{4\pi}(\Omega_{i,1,2} + \Omega_{i,3,4}). \tag{2.58}$$

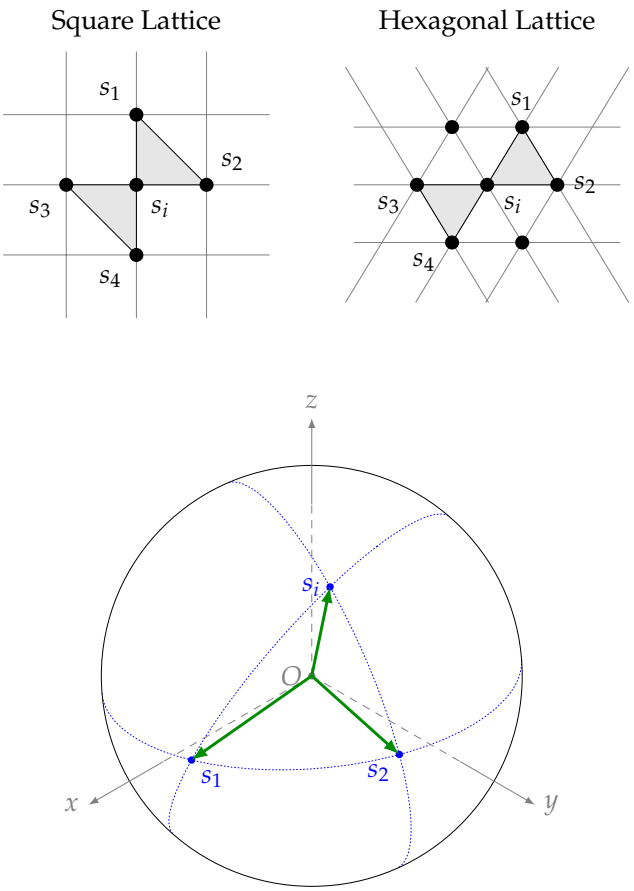

FIGURE 2.9: The top two images show the triangulation method used to calculate the skyrmion number for a square and hexagonal lattice. The solid angle is the spherical area of the triangle enclosed by the three spins, $(\boldsymbol{s_i}, \boldsymbol{s_1}, \boldsymbol{s_2})$, shown in the bottom plot.

Chapter 3

Background in relevant Machine Learning methods

Machine learning (ML) methods are used extensively in our research both in phase classification of materials with defects (Chapter 7), as well as in defect detection and Hamiltonian parameter estimation (Chapter 8). This chapter commences with a brief outline of neural networks, their training procedure and the various layers they are comprised of. We then delve into the specific architectures and ML methods used in our work, including the U-Net architecture, principal component analysis, and variational autoencoders.

3.1 Supervised vs. Unsupervised Learning

ML methods can be broadly categorised into two main groups, each with different objectives and methodologies. **Supervised** learning involves training on a labelled dataset, where input x has a corresponding label y, and the model learns a mapping $f(x)$, that closely predicts the true target, y[93, 94]. Some examples of supervised methods include classification and regression tasks. For example, Chapter 8 involves a regression task to predict the parameters of the magnetic Hamiltonian using the U-Net architecture. Conversely, in **unsupervised** learning, the model operates on an unlabelled dataset, where it learns to uncover patterns, structures, or relationships within the data without explicit guidance[93, 94]. Tasks requiring unsupervised learning include clustering, anomaly detection, and dimensionality reduction.

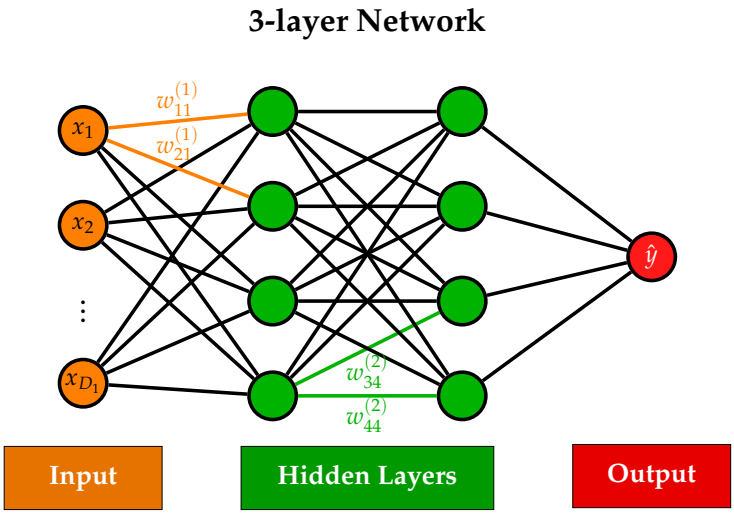

FIGURE 3.1: A simple 3-layer neural network with fully connected hidden and output layers.

Chapter 7 is an example of dimensionality reduction, where a spin configuration is reduced to key variables that retain maximal information about the phase.

3.2 Neural Networks

Artificial neural networks (ANNs) are ML models that loosely emulate the neural pathways found in the brain, where the nodes are the 'neurons' of the system and a collection of nodes form a layer in the network. There are connections between nodes of different layers and these are weighted depending on the importance of the feature detected by the node. ANNs typically consist of several layers of nodes and therefore these models are classified as deep learning (DL) models. They are considered to be **universal function approximators**[95], capable of approximating any any nonlinear function due to their inherent complexity.

A simple 3-layer feed-forward neural network[1] is shown in Figure 3.1. The network is composed of **fully connected** (Dense) layers with connections between all nodes in layers n and $n-1$, effectively modeling the following non-linear function,

$$\hat{y}(\boldsymbol{x}, \boldsymbol{w}^{(n)}) = \sigma\left(w_{1k}^{(3)} h\left(\sum_{j=0}^{D_2} w_{kj}^{(2)} h\left(\sum_{i=0}^{D_1} w_{ji}^{(1)} x_i\right)\right)\right), \tag{3.1}$$

[1] In our notation, an n-layer neural network, contains 1 input layer, $n-1$ hidden layers and 1 output layer. n refers to the hidden and output layers.

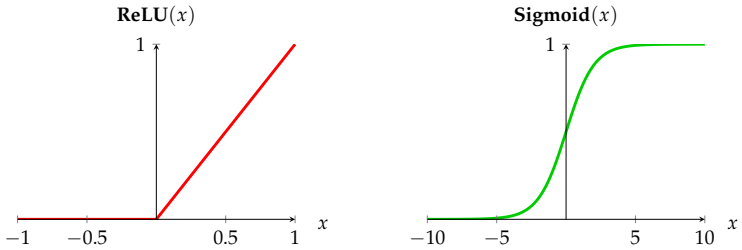

FIGURE 3.2: a) ReLU activaton function, typically used in the hidden layers of ANNs. b) Sigmoid, σ, activation function, used as an output activation function for binary classification.

where $\boldsymbol{x}$ is the 1D input vector with an additional bias of x_0, $\boldsymbol{w}^{(n)}$ is the weight matrix between nodes in layer n and $n-1$ and the weight of the bias is set to 1 ($w_{j0}^{(n)} = 1$). D_q is the number of nodes in the q-th hidden layer. h is the hidden layer activation function, whose main purpose is to add non-linearity to the function[2]. For example, for the rectified linear unit (ReLU) activation ($h(x) = \max(0, x)$), the neuron makes a decision on whether or not to pass on the information contained in the weighted sum, x (see Figure 3.2). σ is the output activation function, which transforms the last layer's features into a format that is compatible with a given problem. For example, in binary classification, σ is the sigmoid activation function (Figure 3.2), giving the probability that input x belongs to class 1,

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \tag{3.2}$$

whereas in a regression task, σ is a linear function, $\sigma(x) = x$.

The training process of a neural network consists of a **forward propagation** step, the evaluation of a **loss function**, and finally a **backpropagation** step to update the weights based on the loss function[93]. Forward propagation is a straightforward progression of input data through the network represented by Equation 3.1. Given the input of the previous layer, a^{l-1}, each layer performs the following steps,

$$z^l = w^{(l)} a^{l-1} \tag{3.3}$$

$$a^l = h(z^l). \tag{3.4}$$

[2]The network must have non-linearity in order to model non-linear decision boundaries. If the activation functions were all linear, then this would be equivalent to a single layer network and the class of potential functions for the network would be severely limited.

This process generates a prediction for the target variable, $\hat{y}_i$. The loss function measures the difference of the predicted value from the true target value (y_i) and can be found using **maximum likelihood estimation**. For regression tasks, the most common loss function is the mean squared error (MSE), given as follow[93],

$$E(\boldsymbol{w}^{(n)}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2. \tag{3.5}$$

For binary classification tasks, the binary cross-entropy (BCE) loss function is commonly used. BCE measures the performance of a classification model whose output is a probability value between 0 and 1. It is defined as,

$$E(\boldsymbol{w}^{(n)}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]. \tag{3.6}$$

Our primary goal is to find the model parameters, $\boldsymbol{w}^{(n)}$, that minimises the error function, $E(\boldsymbol{w}^{(n)})$. The minima of this function occurs when its gradient with respect to $\boldsymbol{w}$ vanishes[3],

$$\nabla E(\boldsymbol{w}^{(n)}) = 0. \tag{3.7}$$

To find the minima we use an iterative approach, known as **gradient descent**, whereby $\boldsymbol{w}^{(n)}$ is updated by taking small steps in the direction of the steepest descent of $E(\boldsymbol{w}^{(n)})$, given by $-\nabla E(\boldsymbol{w}^{(n)})$[93].

$$\boldsymbol{w}^{\tau+1} = \boldsymbol{w}^{\tau} - \eta \nabla E(\boldsymbol{w}^{\tau}) \tag{3.8}$$

Here, τ is the iteration step, and $\eta \geq 0$ is the learning rate[4]. If this is computed using the full dataset, then it is referred to as **batch gradient descent** (BGD). For large datasets, where storing the entire dataset in memory for gradient calculation is impractical, we instead use an online variant, called **stochastic gradient descent** (SGD), where the gradients are updated after every individual datapoint. Though SGD converges faster than BGD and is more likely to escape local minima, it never

[3] Due to the nonlinearity, this is not a convex function of $\boldsymbol{w}^{(n)}$, thus containing multiple minima. The training process may converge to a local minimum rather than the global one, which still provides a satisfactory approximation for the parameters.

[4] If the learning rate is set to a small value, it may take a long time for the algorithm to converge, but a high learning rate may prevent the algorithm from settling to the optimum solution.

fully converges to the optimum, instead oscillating about the minima. An intermediate method is **mini-batch gradient descent** where the dataset is partitioned into batches and the gradient is computed over each batch[93]. This method can take advantage of vectorisation to speed up the computation and reduces the volatility of the oscillations.

To implement the iterative schemes discussed above, the computation of $\nabla E(\boldsymbol{w})$ is essential, and can be found via **backpropagation**[96]. This procedure initially computes the loss gradient with respect to the output, $\frac{\partial E}{\partial \hat{y}_i}$, and on subsequent steps, uses Leibniz's chain rule to recursively determine the gradients with respect to the activations, a^l, and inputs, z^l, propagating the error backward through the network. These gradients indicate the level of contribution of each parameter to the final error. Once $\nabla E(\boldsymbol{w})$ is computed, parameters are updated through gradient descent. This process of forward propagation, backpropagation and parameter update is repeated for multiple epochs until satisfactory convergence of the model parameters $\boldsymbol{w}$ is achieved. Following training, the model's performance is evaluated on an unseen dataset (test set). Additionally, model hyperparameters such as the number of layers or the learning rate can be fine-tuned using a validation dataset before final evaluation on the test set.

3.3 Network Layers

Our discussions so far have only mentioned the Dense layer with fully interconnected nodes, $z^l = w^{(l)} a^{l-1}$. However this transformation can be replaced by a number of other layers, each serving a different purpose. The only requirement is that z is differentiable with respect to its weights and inputs, such that it is compatible with the backpropagation algorithm. The following section discusses a few of the key layers used in our study.

3.3.1 Convolutional Layers

The main drawback of Dense layers is the inefficiency in learning their numerous weights, leading to longer training times and potential overfitting[5]. Convolutional (Conv) layers address this issue by leveraging the spatial locality of input features to substantially reduce the number of trainable parameters[97–102]. They extract key features of an image through a convolution with a small matrix known as the **kernel**. For a 2D image, X, of size $N_x \times N_y$, the convolution with a kernel, K, of size $k_x \times k_y$, applied at a stride of (s_x, s_y) gives the following feature map,

$$Z_{ij} = (K * X)_{ij} = \sum_{m=0}^{k_x-1} \sum_{n=0}^{k_y-1} K_{m,n} X_{m+s_x i, n+s_y j}. \tag{3.9}$$

Here, the size of output Z is Floor$(\frac{N_x-k_x+2P}{s_x}+1)$ in the x-dimension, and Floor$(\frac{N_y-k_y+2P}{s_y}+1)$ in the y-dimension. P refers to padding around the input image and is used to preserve any features on the boundaries. Some example feature maps with specific kernels are shown in Figure 3.3.

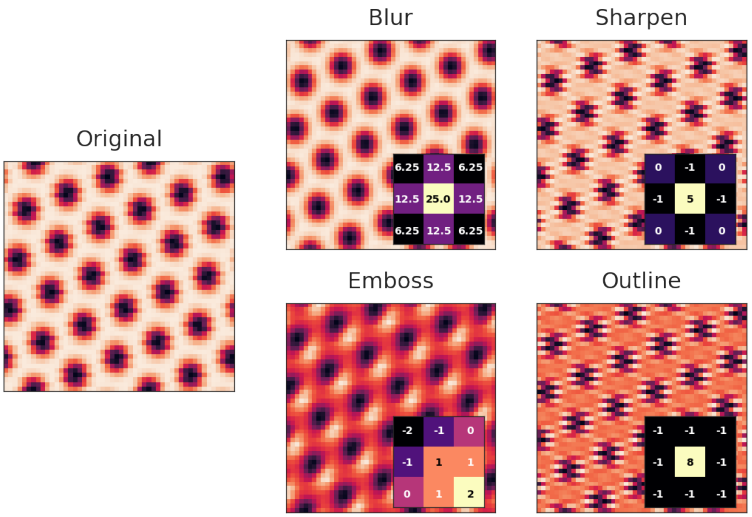

FIGURE 3.3: The output of applying a convolution of the blur, sharpen, emboss and outline kernels (inset) to an image of a skyrmion lattice.

The convolution operation is performed by sliding the kernel window over the image, calculating an elementwise multiplication on the overlapping pixels and summing the result (see Figure 3.4). The size of the kernel determines the number of

[5]Overfitting occurs when a machine learning model learns the training data too precisely, capturing noise and random fluctuations instead of the underlying true relationship, which results in good performance on training data but poor generalisation to new, unseen data.

input features that affect a particular output feature, limiting the spatial locality of these input features. When the size of the kernel matches that of the input, the Conv layer behaves equivalently to a Dense layer, with each output feature representing a weighted combination of all input features. The weights within the kernel matrix are learned via backpropagation.

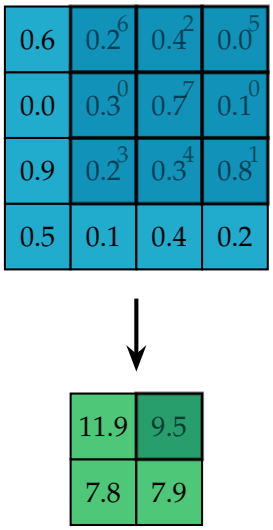

FIGURE 3.4: Demonstration of the convolution operation. The 4×4 matrix is the input data, which is overlaid by a darker blue 3×3 kernel with is values displayed in the top-right corner. The operation is applied with a stride of 1 in the x and y directions, resulting in a 2×2 output matrix in green. The dark green output value shows the results of the sum of the elementwise multiplication of the kernel and the pixels in the top right corner of the input matrix.

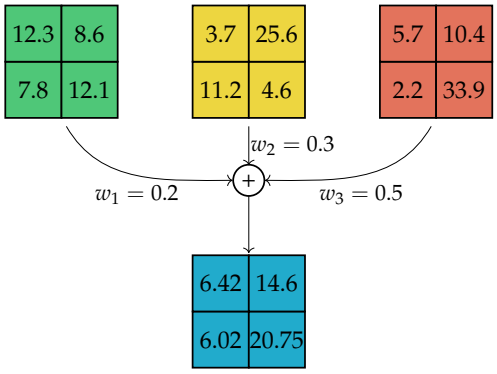

FIGURE 3.5: The feature maps of each channel, obtained by convolving a kernel with a channel of the input, are summed to produce a final output of a filter. A convolutional layer can have several filters.

Our image input may contain more than one channel, for example a coloured image of size $N_x \times N_y$ has three channels representing the red, green and blue pixel values, resulting in an input size of $N_x \times N_y \times 3$. Conv layers implement a unique

kernel for each channel, and the concatenation of these kernels is called a **filter**. A kernel of size $k_x \times k_y$ results in a filter of size $k_x \times k_y \times c_{in}$, where c_{in} is the number of input channels, and $k_x < N_x$, $k_y < N_y$. When the filter is applied to an image, each kernel processes its respective channel and the resulting feature maps are elementwise summed to produce a single output map, as shown in Figure 3.5. Each kernel in a filter has an independent weight, potentially causing one channel to be prioitised over the others. A bias term is then added to the resulting kernel, before it is passed through a non-linear activation function. Conv layers can have multiple filters for extracting different features, with the number of filters being a tunable hyperparameter.

3.3.2 Up-convolutional Layers

The transpose or up-convolution (UpConv) operation is an upsampling technique that can reverse the size reduction that occurs from a Conv layer. However, it is important to note that this does not reverse the convolution operation, but rather performs a backward pass of the Conv layer. Specifically, it involves sliding the input across the kernel and performing an elementwise multiplication, before overlapping elements are summed. Figure 3.6 shows an example transpose convolution operation, applying the same kernel from Figure 3.4 to its output. The output of UpConv is different to the input of the Conv layer (Figure 3.4) since UpConv does not reverse the Conv operation. The UpConv layer is commonly used in fully convolutional networks such as the U-Net in order to produce an output matrix of identical size as an input image. For more in-depth discussions on the U-Net architecture, please refer to Section 3.4.

3.3.3 Pooling Layers

One limitation of Conv layers is their tendency to encode the precise position of features in the input, and minor changes such as rotations or translations can drastically alter the feature map. Downsampling mitigates this issue by reducing the resolution of the input features while preserving essential information, thereby decreasing sensitivity to small variations. This can be acheived either by increasing the stride of

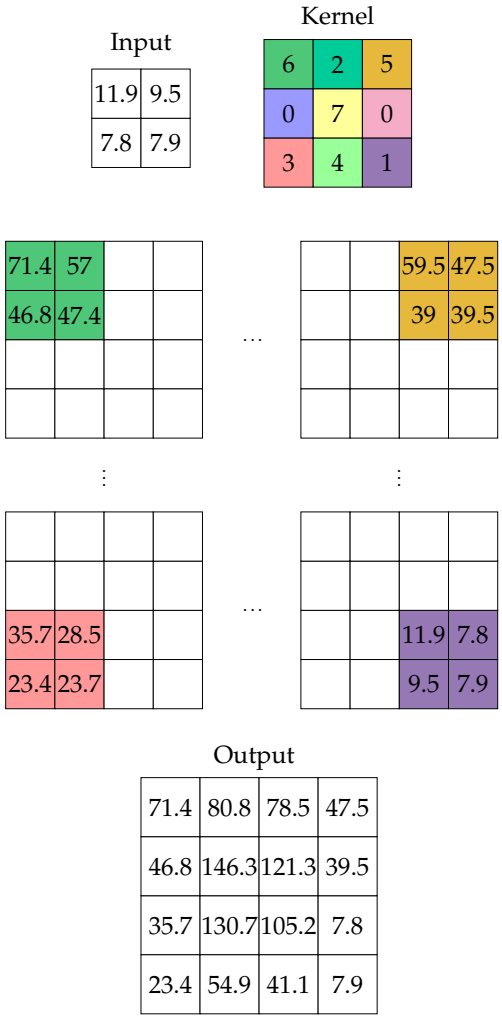

FIGURE 3.6: The transpose convolution operation. Each element of the kernel is element-wise multiplied by the input to produce 3×3 (kernel-sized) intermediate matrices and the overlapping values are summed for the final output.

the Conv layer or through the use of pooling layers. Pooling layers divide the space into non-overlapping regions called pooling regions (F) and performs an operation to reduce the patch to a single value[103].

$$\text{Pool}(X)_{i,j,k} = f(X_{is_x+m,js_y+n,k}; n, m) \tag{3.10}$$

Here, (s_x, s_y) is the stride and (m, n) is the size of F in the x and y dimensions. The most common pooling layer is MaxPool, where $f(X; n, m) = \max_{m,n}(X)$. This operation retains the presence of a feature without also learning its exact position, thereby enhancing the translation invariance of the model.

3.3.4 Skip Connections

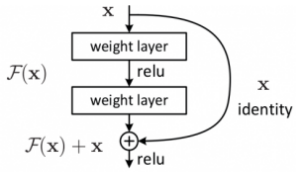

FIGURE 3.7: A residual block in ResNet with a skip connection. Taken from [101].

As the depth of a network increases, its ability to capture abstract and intricate features from the input is enhanced. However, beyond a certain depth threshold, the model performance begins to degrade, likely due to vanishing gradients and low convergence rates in deep layers[101]. Skip connections provide an alternative pathway for data to traverse from preceding layers to subsequent ones, bypassing intermediate layers. This is in contrast to the sequential pipeline of traditional networks. More generally, the output of neurons in layer l, incorporating information from layer $l-k$, where $k < l$, is given by,

$$a_l = h(W^{l-1,l}a_{l-1} + b_{l-1} + W^{l-k,l}a_{l-k}) \tag{3.11}$$

where h is the activation function, $W^{l-i,l}$ denotes the weights between the $l-i$-th and l-th layers, and a_{l-i} and b_{l-i} represent the outputs and biases of the $l-i$-th layer. This was first implemented in the ResNet architecture[101], where the weight $W^{l-k,l}$ corresponds to the identity matrix (see Figure 3.7). With convolutional layers, skip connections are implemented by either concatenating earlier p feature channels from layer $l-k$ to the q channels in layer l, leading to a layer output with $p+q$ channels, or through the addition of the p feature channels in $l-k$ to an equal number of p channels in l.

3.4 The U-Net Model

The U-Net architecture was designed specifically for image segmentation tasks[44], where instead of classifying a single object in a image, it simultaneously identifies multiple objects[44, 105–108], such as the ones shown in Figure 3.8. In Chapter 8, we

FIGURE 3.8: An example input image with its corresponding label matrix containing the class of each object, taken from the COCO dataset[104].

reuse this architecture, as a general mapping function between a 2D input and 2D output, to learn the underlying defect field of a material, given an input spin texture.

One of the earliest models to tackle image segmentation was the **fully convolutional network** (FCN)[105], which substituted the Dense layers of a traditional convolutional neural network with an up-convolutional layer, enabling pixel-wise classification. This modification resulted in two aditional benefits: firstly, the model was no longer restricted to fixed-size inputs which are necessary for the matrix multiplication in the forward-pass step for Dense layers, and secondly, it reduces the computational resources required for training, due to the decreased number of parameters in Conv layers as opposed to Dense ones. In Chapter 8, we use FCN as a benchmark model to test the effectiveness of our U-Net variant in Hamiltonian parameter estimation.

The U-Net is an encoder-decoder model, that enchances the network structure of the FCN by introducing a symmetric architecture between the successive Conv and UpConv layers[44]. These pathways, called the contracting and expanding paths respectively, form a 'U'-shaped architecture that gives the model its name, see Figure 3.9[44]. In the contracting path, three blocks of two 3×3 Conv layers with a ReLU activation are followed by a downsampling using a 2×2 MaxPool operation with a stride of 2. MaxPool halves the resolution of the features, while the Conv layer immediately following it doubles the number of feature channels. The expanding path reverses this structure with three blocks of a transposed 2×2 convolution to halve the number of feature channels, followed by a skip connection (via concatenation) to features channels from the encoder path, and finally two additional 3×3 Conv layers. Upsampling is a sparse operation, so concatenation improves the prediction

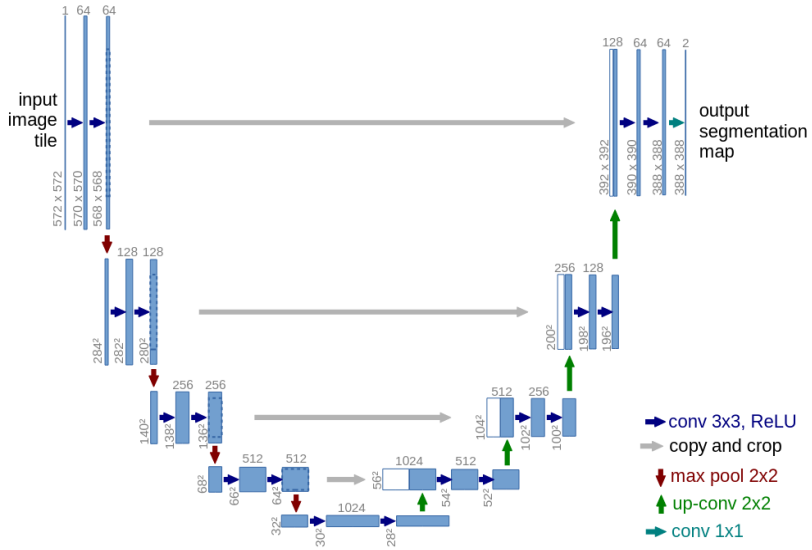

FIGURE 3.9: The U-Net architecture. The blue boxes correspond to the feature channels computed at each layer while the white boxes in the decoder path represent the high-resolution feature channels taken from their respective encoder layer. The number of feature channels is denoted at the top of the box and the $x - y$ dimensions are at the side of the boxes.

accuracy by adding the high-resolution details from the encoding path with the currently learned features in the decoder path. The final layer uses a 1×1 convolutional layer to map all prior feature channels to a matrix indicating the pixel class. We implement a variant of this network in Chapter 8, where the final output contains the predictions of a site-dependent random pinning field.

3.5 Dimensionality Reduction

The following section summarises the two dimensionality reduction techniques used for phase classification in Chapter 7.

3.5.1 Principal Component Analysis

Principal Component Analysis (PCA) is a dimensionality reduction technique that reduces the redundancy in the feature vector while preserving as much information as possible from the original dataset, thereby resulting in a more concise representation of the data[109]. It accomplishes this by linearly transforming a high-dimensional dataset into a lower-dimensional subspace, spanned by a set of orthogonal axes, called principal axes, along which the data exhibits the greatest variance[109–112].

This transformation is achieved by diagonalising the covariance matrix of the dataset[109]. A dataset of m examples with n features, $\boldsymbol{X} \in \mathbb{R}^{n \times m}$, may contain several highly-correlated features that could potentially be reduced to a single variable. The covariance between any two features, $(\boldsymbol{x}_i, \boldsymbol{x}_j)$, captures how these move in relation to one another,

$$cov(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{1}{m} \sum_{k=1}^{m} (\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{jk} - \bar{\boldsymbol{x}}_j) \tag{3.12}$$

where $\bar{\boldsymbol{x}}_i$ and $\bar{\boldsymbol{x}}_j$ denote the means of features i and j. The covariance matrix encodes the pairwise covariances between all features in the dataset,

$$\Sigma_{\boldsymbol{X}} = cov(\boldsymbol{X}, \boldsymbol{X}) = \begin{pmatrix} var(\boldsymbol{x}_1) & \dots & cov(\boldsymbol{x}_1, \boldsymbol{x}_n) \\ \vdots & \ddots & \vdots \\ cov(\boldsymbol{x}_n, \boldsymbol{x}_1) & \dots & var(\boldsymbol{x}_n) \end{pmatrix}, \tag{3.13}$$

where $var(\boldsymbol{x}_i) = \frac{1}{m} \sum_{k=1}^{m} (\boldsymbol{x}_{ik} - \bar{\boldsymbol{x}}_i)^2$ is the variance of feature i, and $\Sigma_{\boldsymbol{X}}$ is symmetric, with $cov(\boldsymbol{x}_i, \boldsymbol{x}_j) = cov(\boldsymbol{x}_j, \boldsymbol{x}_i)$. If the features are mean-centered, the covariance simplifies to $\Sigma_{\boldsymbol{X}} = \frac{1}{m} \boldsymbol{X}\boldsymbol{X}^T$. A large magnitude of a diagonal term in this matrix signifies the importance of this feature (high variance), whereas a large magnitude of an off-diagonal term suggests a high redundancy due to these features. Ideally, the optimised covariance matrix would have off-diagonal terms set to zero, producing uncorrelated features, and the diagonal terms would be ordered in terms of the features exhibiting the greatest variance. This can be achieved by finding an matrix, $\boldsymbol{P}$, that linearly transforms the dataset $\boldsymbol{X}$ to $\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{X}$, such that the first basis vector

aligns with the axis of maximal variance and subsequent basis vectors are found, subject to the constraint of orthonormality.

The covariance of $\boldsymbol{Y}$ can be written in terms of the covariance of $\boldsymbol{X}$ as follows,

$$\Sigma_{\boldsymbol{Y}} = \frac{1}{m}\boldsymbol{Y}\boldsymbol{Y}^T = \frac{1}{m}(\boldsymbol{P}\boldsymbol{X})(\boldsymbol{P}\boldsymbol{X})^T = \frac{1}{m}\boldsymbol{P}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{P}^T = \boldsymbol{P}\Sigma_{\boldsymbol{X}}\boldsymbol{P}^T \tag{3.14}$$

Since $\Sigma_{\boldsymbol{X}}$ is a symmetric matrix it can be diagonalised by an orthogonal matrix composed of its eigenvectors, $\Sigma_{\boldsymbol{X}} = \boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^T$, where each column in $\boldsymbol{C}$ corresponds to an eigenvector of $\Sigma_{\boldsymbol{X}}$, and $\boldsymbol{D}$ is a diagonal matrix with the respective eigenvalue of each eigenvector. Selecting $\boldsymbol{P}$, such that each row ($\boldsymbol{p}_i$) is an eigenvector of $\Sigma_{\boldsymbol{X}}$ ($\boldsymbol{P} = \boldsymbol{C}^T$), diagonalises $\Sigma_{\boldsymbol{Y}}$[6],

$$\Sigma_{\boldsymbol{Y}} = \boldsymbol{P}\Sigma_{\boldsymbol{X}}\boldsymbol{P}^T = \boldsymbol{P}(\boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^T)\boldsymbol{P}^T = (\boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{D}(\boldsymbol{P}\boldsymbol{P}^T) = (\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{D}(\boldsymbol{P}\boldsymbol{P}^{-1}) = \boldsymbol{D} \tag{3.15}$$

The principal components are therefore given by the eigenvectors of $\Sigma_{\boldsymbol{X}}$ with the variance along $\boldsymbol{p}_i$ given by the i-th eigenvalue, found via the decomposition of $\Sigma_{\boldsymbol{X}}$[7]. Finally, by retaining only the first d principal components ($d < n$), the higher-dimensional dataset can be projected to a lower-dimensional subspace,

$$\hat{\boldsymbol{Y}} = \hat{\boldsymbol{P}}\boldsymbol{X} \in \mathbb{R}^{d \times m}, \tag{3.16}$$

where $\hat{\boldsymbol{P}} \in \mathbb{R}^{d \times n}$ is the projection matrix.

3.5.2 Variational Autoencoders

An autoencoder is as a latent variable model, mapping a D-dimensional datapoint, denoted as $\boldsymbol{x}_i \in \mathbb{R}^D$, to lower-dimensional latent variables, represented by $\boldsymbol{z}_i \in \mathbb{R}^M$, where $M \leq D$. Although these latent variables cannot be directly observed, they can be inferred from observable variables, for example intelligence can be inferred

[6] Uses the property that the inverse of an orthogonal matrix is equal to its transpose, $P^{-1} = P^T$.

[7] In an alternative derivation, principal components can be found by maximising the variance of the data projected onto the principal component, $\boldsymbol{p}_1^T\Sigma_{\boldsymbol{X}}\boldsymbol{p}_1$, under the constraint that $\boldsymbol{p}_1$ is a unit vector, $(\boldsymbol{p}_1^T\Sigma_{\boldsymbol{X}}\boldsymbol{p}_1 - \lambda(1 - \boldsymbol{p}_1^T\boldsymbol{p}_1)$. Consequently, this also results in $\boldsymbol{p}_1$ being the eigenvector of $\Sigma_{\boldsymbol{X}}$ with the greatest eigenvalue. Subsequent principal components can be similarly found, with additional constraints on the orthonormality of basis vectors.

from the scores of different exams. Autoencoders learn to extract these latent variables from the observed data, denoted as $p(\boldsymbol{z}|\boldsymbol{x})$ (the encoder), and subsequently reconstruct the full features from the compressed latent information, represented by $p(\boldsymbol{x}|\boldsymbol{z})$ (the decoder).

Given a dataset of N independent and identically distributed samples, $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^{N}$, the objective is to determine the parameters, θ^*, that approximate the true distribution, $p(\boldsymbol{X})$. The common approach is maximum likelihood estimation (MLE), where the likelihood of the data, marginalized over the latent variables, is maximized,

$$\hat{\theta} = \arg\max_{\theta \in \Theta} p_\theta(\boldsymbol{X}) = \arg\max_{\theta \in \Theta} \sum_{i=1}^{N} \log p_\theta(\boldsymbol{x}_i). \tag{3.17}$$

However, this method encounters two main challenges[43],

1. The computation of the marginal distribution, $p_\theta(\boldsymbol{x}_i)$, involves solving the integral
$$p_\theta(\boldsymbol{x}_i) = \int p_\theta(\boldsymbol{x}_i, \boldsymbol{z}) d\boldsymbol{z} = \int p_\theta(\boldsymbol{x}_i|\boldsymbol{z}) p_\theta(\boldsymbol{z}) d\boldsymbol{z}, \tag{3.18}$$
which is often intractable[8].

2. The integral must be computed over all N datapoints, thus we cannot optimize for speed or RAM usage with mini-batch GD or SGD.

To address the intractability issue, we use variational inference (VI), which reframes the inference problem as an optimization task. In VI, we seek to approximate the true posterior distribution, $p_\theta(\boldsymbol{z}|\boldsymbol{x})$, using a variational approximation, $q_\phi(\boldsymbol{z}|\boldsymbol{x})$, parameterized by ϕ, drawn from a family of approximations denoted as $\mathcal{Q}$. Our objective is to find the ϕ that selects the $q_\phi(\boldsymbol{z}|\boldsymbol{x})$ closest to $p_\theta(\boldsymbol{z}|\boldsymbol{x})$, measured by the Kullback-Leibler (KL) divergence,

$$D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x_i})||p_\theta(\boldsymbol{z}|\boldsymbol{x}_i)) = \int q_\phi(\boldsymbol{z}|\boldsymbol{x_i}) \log \frac{q_\phi(\boldsymbol{z}|\boldsymbol{x_i})}{p_\theta(\boldsymbol{z}|\boldsymbol{x}_i)} d\boldsymbol{z} \tag{3.19}$$

Rewriting $\log p_\theta(\boldsymbol{x}_i)$ in terms of $D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x_i})||p_\theta(\boldsymbol{z}|\boldsymbol{x}_i))$, we find that maximising the log-likelihood is equivalent to maximising the **evidence lower bound** (ELBO), $\mathcal{L}(\phi, \theta; \boldsymbol{x}_i)$, with respect to the model parameters, θ and variational parameters, ϕ,

[8] Requires exponential time to compute.

for each datapoint $\boldsymbol{x}_i$,

$$\max_{\theta} \sum_{i=1}^{N} \max_{\phi} \mathcal{L}(\phi, \theta; \boldsymbol{x}_i) \tag{3.20}$$

$$\mathcal{L}(\phi, \theta; \boldsymbol{x}_i) = \mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)}[\log p_\theta(\boldsymbol{x}_i|\boldsymbol{z})] - D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)||p_\theta(\boldsymbol{z})) \tag{3.21}$$

The Stochastic Gradient Variational Bayes (SGVB) estimator

Maximising the ELBO requires the gradients of $\mathcal{L}(\phi, \theta; \boldsymbol{x}_i)$ with respect to θ and ϕ, which are summarised as follows,

$$\nabla_\theta \mathcal{L}(\boldsymbol{x}_i) = \mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)}[\nabla_\theta \log p_\theta(\boldsymbol{x}_i, \boldsymbol{z})] \tag{3.22}$$

$$\nabla_\phi \mathcal{L}(\boldsymbol{x}_i) = \mathbb{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)}[(\log p_\theta(\boldsymbol{x}_i, \boldsymbol{z}) - \log q_\phi(\boldsymbol{z}|\boldsymbol{x}_i))\nabla_\phi \log q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)] \tag{3.23}$$

These gradients involve intractable expectations over the latent space, and therefore must be estimated using Monte Carlo. This entails generating m samples of latent variables, $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$, from $q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)$ and computing the following averages,

$$\nabla_\theta \tilde{\mathcal{L}}(\boldsymbol{x}_i) = \frac{1}{m} \sum_{j=1}^{m} \nabla_\theta \log p_\theta(\boldsymbol{x}_i, \boldsymbol{z}_j) \tag{3.24}$$

$$\nabla_\phi \tilde{\mathcal{L}}(\boldsymbol{x}_i) = \frac{1}{m} \sum_{j=1}^{m} (\log p_\theta(\boldsymbol{x}_i, \boldsymbol{z}_j) - \log q_\phi(\boldsymbol{z}_j|\boldsymbol{x}_i))\nabla_\phi \log q_\phi(\boldsymbol{z}_j|\boldsymbol{x}_i) \tag{3.25}$$

Though these estimators are unbiased, the variance of the **score function estimator**, $\nabla_\phi \tilde{\mathcal{L}}(\boldsymbol{x}_i)$, is high, reducing the efficiency of the Monte Carlo algorithm and therefore requiring more steps for convergence[113]. Kingma et al. proposed an alternative estimator, the **Stochastic Gradient Variational Bayes** (SGVB) estimator, using the **reparametrisation trick**, which exhibits lower variance[43].

For various differentiable parametric families, we can generate samples $\tilde{\boldsymbol{z}} \sim q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)$ using a two-step generative process:

1. Generate samples from a simple distribution, $\mathcal{E} \sim p(\mathcal{E})$, i.e. $\mathcal{N}(0, I)$
2. Apply a differentiable deterministic function, $\tilde{\boldsymbol{z}} = g_\phi(\mathcal{E}, \boldsymbol{x})$, to $\mathcal{E}$. Then $\tilde{\boldsymbol{z}}$ follows the distribution $q_\phi(\boldsymbol{z}|\boldsymbol{x}_i)$.

For instance, if the parametric approximation was a univariate Gaussian, $q_\phi(z|x_i) \sim \mathcal{N}(\mu, \sigma)$, then drawing $\mathcal{E} \sim \mathcal{N}(0,1)$ and transforming it by $g_\phi(\mathcal{E}) = \mu + \sigma\mathcal{E} = z$, yields $z \sim \mathcal{N}(\mu, \sigma)$ as desired.

Using the reparametrisation trick, we can express the derivative with respect to ϕ as follows:

$$\nabla_\phi \mathbb{E}_{q_\phi(z|x_i)}[f(z)] = \nabla_\phi \mathbb{E}_{p(\mathcal{E})}[f(g_\phi(\boldsymbol{\mathcal{E}}, x_i))] \tag{3.26}$$

$$= \mathbb{E}_{p(\mathcal{E})}[\nabla_\phi f(g_\phi(\boldsymbol{\mathcal{E}}, x_i))], \tag{3.27}$$

hence the estimator for the derivative becomes:

$$\nabla_\phi \bar{\mathcal{L}}(x_i) = \frac{1}{m}\sum_{j=1}^{m} \nabla_\phi f(g_\phi(\boldsymbol{\mathcal{E}}, x_j)) \tag{3.28}$$

This estimator is also unbiased for $\nabla_\phi \mathcal{L}(x_i)$, but it exhibits lower variance compared to the score estimator, thereby enhancing the efficiency of the learning process. The full algorithm for the auto-encoding variational bayes (AEVB) is given in Algorithm 5.

Algorithm 5 The AEVB algorithm : Taken from [43]

1: Initialize θ, ϕ
2: **while** θ, ϕ not converged **do**
3: Select random minibatch of M points, $\boldsymbol{X}^M$
4: Take random samples, $\mathcal{E} \sim p(\mathcal{E})$
5: Calculate gradients, $\boldsymbol{g} \longleftarrow \nabla_{\theta,\phi}\bar{\mathcal{L}}(\theta, \phi; \boldsymbol{X}^M, \mathcal{E})$
6: Update θ, ϕ with $\boldsymbol{g}$ (using SGD or Adagrad)
7: **end while**
8: **return** θ, ϕ

Until now, we've maintained a general approach without specifying the exact functional forms of $q_\phi(z|x_i)$, and $p_\theta(x_i|z)$. In our study, convolutional neural networks are used to define both the encoder, $q_\phi(z|x_i)$ and the decoder, $p_\theta(x_i|z)$. Figure 3.10 depicts the architecture, where the input spin texture ($\boldsymbol{x}$) is fed into the CNN encoder to output the means ($\boldsymbol{\mu}$) and standard deviations ($\boldsymbol{\sigma}$) of the latent space. Latent variable $\boldsymbol{z}$ is sampled from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma})$ using the reparametrisation trick, and then input into the CNN decoder to reproduce the spin texture ($\hat{\boldsymbol{x}}$). By training such a model in Chapter 7, we can classify the majority of phases observed in materials with defects.

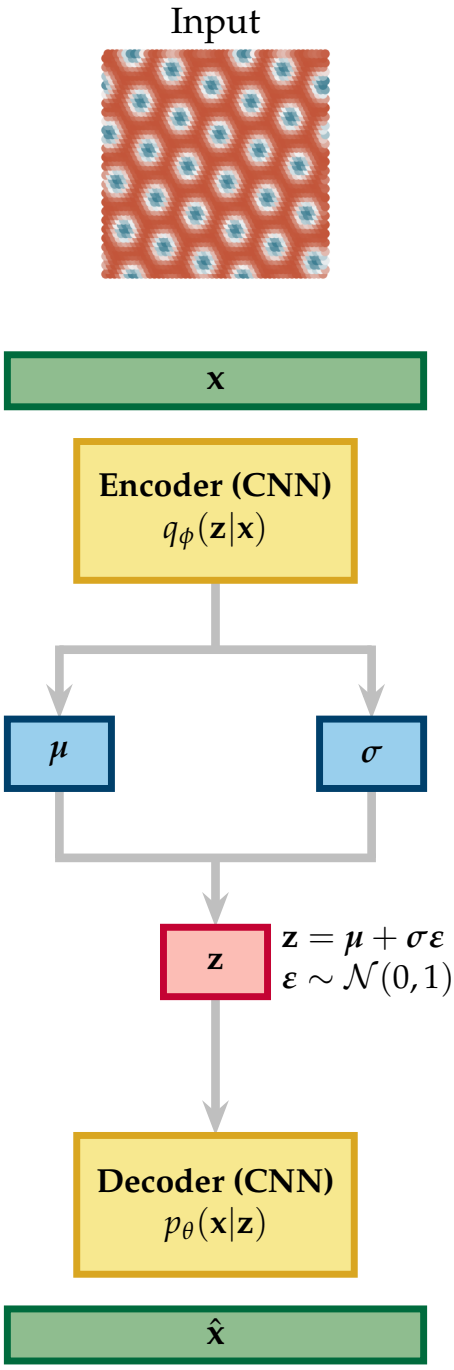

FIGURE 3.10: VAE network architecture, using convolutional neural networks for the encoder and decoder.

3.6 Conclusion

In conclusion, we have introduced the fundamental concepts behind neural networks, focusing on the U-Net network, which we repurpose in Chapter 8 for Hamiltonian parameter estimation and defect detection. Additionally, we discuss dimensionality reduction techniques, such as PCA and VAE, used in Chapter 7 to capture

signatures of phase transitions directly from the spin configuration. In the upcoming chapter, we delve into the topic of skyrmions, providing an overview of the current research and highlighting challenges in using skyrmions for storage solutions as a result of material defects.

Chapter 4

Background of Skyrmion Research

4.1 Skyrmions and their Discovery

In 1962, Tony Skyrme introduced a groundbreaking concept that fundamentally reshaped our understanding of particles — the skyrmion[114]. Driven by the quest to unify meson and baryon particles, Skyrme devised a field-theoretical model of the meson, from which a topological knot emerges, embodying a localised nucleon particle. Imagined as a "hedgehog"-like structure, formed by wrapping a three-dimensional vector field around a unit sphere, the skyrmion's stability was solely attributed to a unique property of the field — its topology. Initially hailed for its promise, especially in explaining the low-energy behaviours of nucleons, skyrmions eventually yielded to the advancements of quantum chromodynamics. However, the concept of topology, once confined to the annals of mathematics, emerged as a tangible property stabilising particles throughout condensed matter physics, from chiral nematic liquid crystals[115] to Bose-Einstein condensates[116–119] and quantum Hall magnets[120–123]. This thesis focuses specifically on skyrmion particles emerging from the spin field of magnetic materials.

A resurgence of interest in skyrmions within magnetic systems occurred in 1989, inspired by the emergence of Abrikosov vortices within type-II superconductors. Bogdanov and Yablonskii envisioned a lattice of particles, similar to vortices[52], sustained by the delicate interplay between the isotropic and a newly-introduced inhomogeneous exchange, described within Landau's free energy expansion. This additional exchange, known as the Dzyaloshinskii-Moriya interaction (DMI)[49], favours an orthogonal alignment of magnetic moments, encouraging the formation

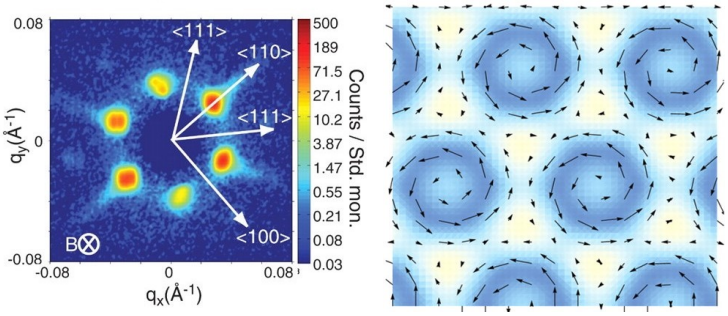

FIGURE 4.1: a) Small-angle neutron scattering (SANS) experiment reveals a novel six-fold FFT image which could indicate the presence of a multi-Q structure. b) Mean field simulations of a proposed three-Q structure produces a skyrmion lattice. Taken from [11].

of a diverse array of canted spin configurations including helices, cones, and — most importantly — skyrmions.

Unfortunately, the presence of skyrmions in real materials is constrained by Neumann's principle which dictates that a material's free energy must adhere to the symmetries of its underlying crystal lattice. Consequently, skyrmions are observable in only a limited selection of materials, where inversion symmetry is broken. Nearly two decades after its theoretical proposition, the experimental confirmation was achieved by Mühlbauer et al. in the low-symmetry ($P2_13$) MnSi crystal[11]. They observed an anomalous six-peak pattern in reciprocal space that defied explanation using known spin structures. By simulating a superposition of three coplanar spin spirals, Mühlbauer successfully reproduced the skyrmion lattice in real space and correspondingly accounted for the six-peak pattern in reciprocal space[11]. Lorentz transmission electron microscopy (LTEM) of a thinned sample reveals the presence of a real-space skyrmion lattices[12], unequivocally establishing the existence of skyrmions.

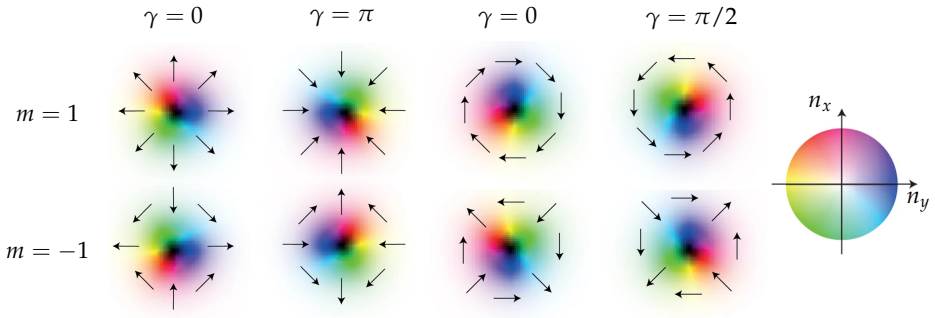

FIGURE 4.2: Various types of skyrmions classified by their helicity, γ, and skyrmion number, m. Taken from [124].

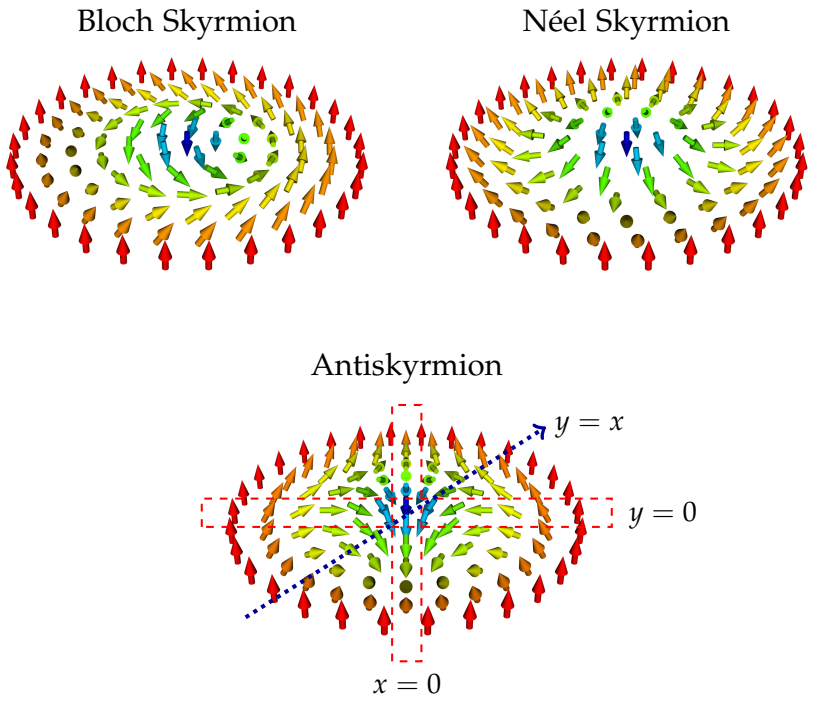

FIGURE 4.3: Examples of commonly-found skyrmion textures: a) Bloch skyrmion, b) Néel skyrmion, c) Antiskyrmion, with a Néel wall along $x = 0$ and $y = 0$, and a Bloch wall on the $y = x$ axis.

Structurally, a skyrmion is a whirling spin texture characterized by a boundary 180° out-of-phase with the central-most spin, and a continuous rotation in spins from perimeter to core. Mathematically, this can be expressed, in terms of polar coordinates $\boldsymbol{r} = (r\cos(\psi), r\sin(\psi))$, as follows,

$$\boldsymbol{m}(\boldsymbol{r}) = \begin{pmatrix} \cos(\theta(\psi))\sin(\phi(r)) \\ \sin(\theta(\psi))\sin(\phi(r)) \\ \cos(\phi(r)) \end{pmatrix} \tag{4.1}$$

This general definition encompasses several skyrmion varieties[124], with differing helicities and skyrmion numbers (Figure 4.2). Among these, most investigations focus on Bloch and Néel skyrmions, so named, due their resemblance to Bloch/Néel domain walls (Figure 4.3). Other variants include the antiskyrmion[125, 126], characterized by a Néel wall along the $x = 0$ and $y = 0$ axes, and a Bloch wall along the $y = x$ axis, a biskyrmion consisting of a pair of bound skyrmions of opposite chirality[127, 128], and the skyrmionium, featuring a skyrmion embedded within another skyrmion with an opposite topological charge[129, 130].

Skyrmions are classified into distinct topological classes based on their skyrmion number (N_{sk}), representing the number of times a spin texture wraps around the unit sphere. At the continuum limit, the skyrmion number is calculated as follows,

$$N_{sk} = \frac{1}{4\pi} \iint d^2 r \boldsymbol{m} \cdot \left(\frac{\partial \boldsymbol{m}}{\partial x} \times \frac{\partial \boldsymbol{m}}{\partial y} \right). \tag{4.2}$$

This integrand sums the total solid angle formed by all spins, mapping the 2D field onto a 3D sphere. Configurations belonging to the same topological class deform smoothly between each other[6], for example in the Bloch to Néel skyrmion transformation. However, when dealing with configurations of different classes, deformation is subject to an energy barrier[6, 7]. This **topological protection** ensures the stability of the skyrmion structure, making them reliable candidates for data particles in non-volatile memory[6, 7].

4.2 Applications in Storage Technology

The primary driving force behind skyrmion research lies in its potential application in innovative spintronic-based storage technology. The demand for efficient data storage solutions has surged in recent years, particularly with the proliferation of Deep Learning applications and expansion of IoT technology. According to a 2020 forecast by Statista[131], the global accumulated data volume, which stood at 64 ZB in 2020, is predicted to soar to 181 ZB by 2025, with the data generation rate continuing to surge.

However, current storage technologies are facing challenges in keeping up with this ever-growing demand for data storage. While significant progress has been made in miniaturising memory cells in transistor-based CMOS[132], with sizes reaching 3nm, these efforts are impeded by quantum tunnelling effects[133, 134]. Moreover, transistors, serving as volatile memory, demand continuous power to uphold data integrity, and substantial power for read/write operations to drive electrons within the transistor circuits[135]. Spintronics presents an alternative approach to creating memory devices, by leveraging electron spin rather than its charge[136–

138]. This approach benefits from lower power consumption since less energy is required to alter spin orientation compared to traditional electron manipulation[135]. Moreover, the inherent property of ferromagnetic materials, to remain magnetised even after the removal of a field, enables non-volatility of data[135]. The pivotal breakthrough in spintronics came with Giant Magnetoresistance (GMR)[139, 140], which facilitated the development of spin-valve read heads, resulting in a 10^3-fold increase in magnetic disk technology efficiency[138]. However, the 10nm magnetic domains of disk technology are reaching their size limit due to challenges in stabilising small magnetic bits against thermal fluctuations[138].

Skyrmion-based devices emerge as a promising replacement for magnetic disks, with the potential to deliver higher data density owing to their smaller sizes in certain materials[4, 5, 135], ranging from 5-200nm[141, 142], and significantly accelerated read/write speeds, of 10ns[135], in contrast to the mechanical read-head used in conventional disk technology.

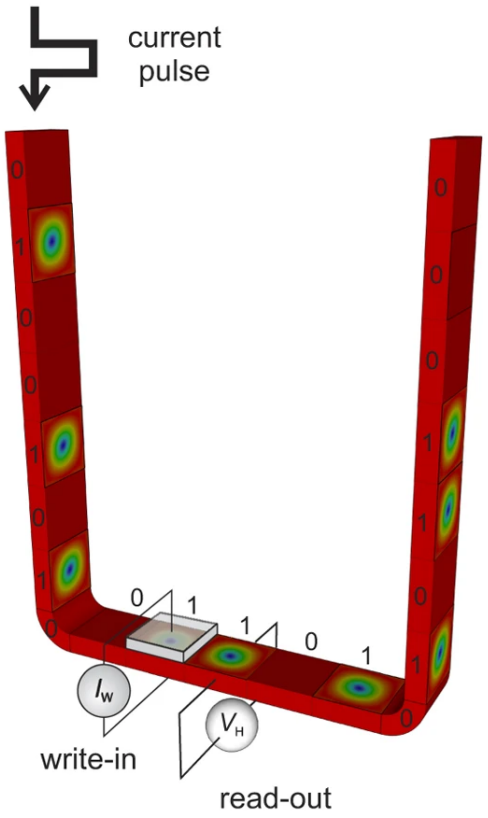

FIGURE 4.4: Depiction of a skyrmion-based racetrack device, where teh presence of a skyrmion encodes a binary 1, and its absence is 0. Taken from [143].

The domain wall (DW) racetrack, first explored in 2007[144, 145], is the leading methodology for memory devices in spintronics, whereby magnetic domains encode binary data traversing a nanowire via spin-polarised current[144–146]. Although promising, DWs are highly sensitive to defects leading to unpredictable motion, reduced speeds, and tracking challenges[147–149]. Conversely, skyrmion-based systems offer multiple advantages[143, 150, 151]. Skyrmions, smaller than domains[135, 141], allow for higher data density and require lower spin currents[1][1–3, 153–156]. Furthermore, owing to their topological stability, skyrmions exhibit greater resilience to defects[124, 157], mitigating the risk of data loss. In skyrmion-based racetracks, binary data is encoded by the presence or absence of a skyrmion and their effective utilisation for data storage offers improvements in writing (via nucleation)[2], reading (via detection)[3], and deletion (via annihilation)[4][22, 177]. Finally, optimising the racetrack itself, by tuning layers[178] (to enhance skyrmion stability and facilitate their formation[22, 32, 35, 179]) or introducing defects (to guide[38, 180, 181] and accelerate skyrmion[36, 37]), provides further advantages.

Unfortunately, the heightened and uncontrollable level of defects in multilayer heterostructures poses a significant obstacle to realising a skyrmion racetrack device. These defects can significantly impede the motion and trajectory of skyrmions[20, 179, 182], even trapping them when the driving current is insufficient[35, 169, 179] — an issue encountered in other device prototypes as well. For example, in a skyrmion-based synapse[183–185], designed to model the weights of a neural network, defects can alter the size and behaviour of skyrmions, leading to unpredictable motion and erroneous weight assignments[185]. Moreover, defects can have detrimental effects on logic gates, amplified by their constricted geometry, resulting in skyrmion annihilation and incorrect logic operations[180, 186, 187]. Though skyrmions offer exciting opportunities, for the realisation of ultradense energy-efficient storage technology, a

[1] $10^{-6} Am^{-2}$ for skyrmions, compared to $10^{-11} Am^{-2}$for DWs[152].

[2] Writing methods include current injection (via spin-transfer torque or spin-orbit torque)[22, 158, 159], applying a localised field (electric or magnetic)[160–162], heating[163], and ion radiation[164, 165], along with the exploitation of defects and geometrical constriction[160, 166].

[3] Reading can be accomplished using magnetoresistance, via magnetic tunnel junctions[167, 168], microscopy imaging[12], or from electrical Hall measurements[169, 170].

[4] Skyrmions can be deleted via magnetic fields[171, 172], electric fields[162, 173, 174] and spin-polarised currents[175, 176].

comprehensive study of defects is essential for advancing skyrmion-based technologies.

4.3 Stability of Skyrmions

One of the primary challenges in skyrmion research, pivotal for the feasibility of skyrmion-based devices, revolves around ensuring their stability under varying external conditions. The energetic stability of a skyrmion hinges on a delicate balance between intrinsic interactions among a material's magnetic moments and external factors like magnetic field strength and temperature. While the former can be influenced by material composition, safeguarding skyrmions against environmental fluctuations across a broad range of fields and temperatures is crucial to prevent their degradation. Furthermore, for the practical implementation of skyrmion-based devices, it is imperative that skyrmions maintain stability even at room temperature (RT)[5] while maintaining a compact size.

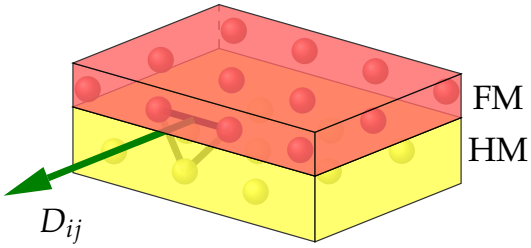

FIGURE 4.5: Interfacial Dzyaloshinskii-Moriya interaction from the symmetry breaking between a ferromagnetic layer and a heavy metal substrate.

However, identifying bulk crystals that meet the stringent criteria for skyrmion-device fabrication poses significant hurdles. Early B20 helimagnets such as MnSi[11] and $Fe_{0.5}Co_{0.5}Si$[12, 13] only host skyrmions within a limited range of fields and temperatures, typically below 40K. Although other B20 crystals like MnGe and FeGe with higher Curie temperatures up to 278 K have been identified, alternative strategies are required to bolster the stability of the skyrmion phase. Early investigations suggest that thinning the material[12, 14, 15] and inducing tensile strain through epitaxial coupling to a substrate can suppress the conical state and broaden the skyrmion phase to encompass a wider temperature and field spectrum. Interfacial

[5] Most consumer storage technology can withstand much higher temperatures (almost 80°)[8–10]

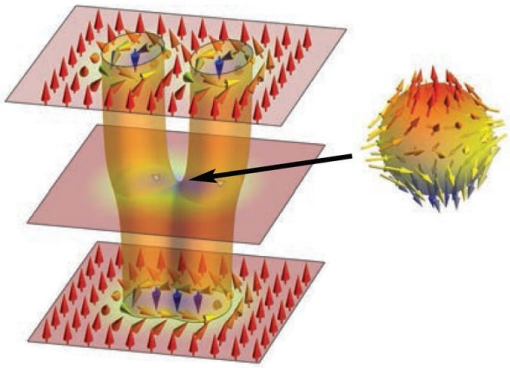

FIGURE 4.6: A depiction of the merging of two skyrmions. This occurs via a Bloch point which creeps up along the tube zipping up skyrmions to form a helical state[191].

DMI, induced by inversion symmetry breaking at the interface between a heavy metal (HM) and ferromagnetic layer (FM), exhibits greater strength compared to bulk DMI[188, 189] (see Figure 4.5). This enhancement, attributed to the substantial spin-orbit coupling from the HM layer[27], further improves the stability of skyrmions[23, 27, 190].

Recent ab-initio and experimental investigations highlight the role of HM substrate imperfections in strengthening spin-orbit coupling and subsequently improving skyrmion stability[192, 193]. Thus, defects emerge as key contributors to skyrmion stability enhancement. Understanding how defects stabilise skyrmions involves delving into the skyrmion decay process. During this process, a topological defect known as a Bloch point emerges at the intersection of two skyrmion tubes, before "zipping" up and merging the skyrmions to form spin spirals[191]. Introducing a defect at the core of the Bloch point halts its progression and prevents the unwinding of the skyrmion, thereby extending its lifespan[17, 194, 195]. Systems with high levels of defects exhibit a broader skyrmion phase under external environmental fluctuations. Defects can be intentionally introduced through doping[13, 17, 40] or by applying mechanical stress[16] to shift atoms from their original lattice sites, consistently enhancing skyrmion stability.

In the absence of defects, improving skyrmion stability relies on rapidly driving the system away from its equilibrium configuration[145], either through cooling protocols[13, 15, 17, 39, 40, 196] or rapid non-thermal excitations within the skyrmion pocket[197, 198]. These processes yield metastable skyrmions, where the particle's

topology forms an energy barrier against thermal agitation, preventing its unwinding to the energetically-favourable spin spiral. However, achieving rapid cooling rates in clean materials necessitates specialised cryogenic equipment, increasing material manufacturing costs[196]. Therefore, leveraging defects holds promise for reducing device manufacturing costs while enhancing data reliability.

4.4 Impact of Material Defects

Thin films and multilayer heterostructures dominate current skyrmion research[15, 18, 20, 22–34, 199, 200], both from ease of tunability for device applicability[23, 190], and their seamless integration to the industry-standard pipeline for semiconductor manufacturing. Unfortunately, a preference for magnetron sputtering[18, 201], over epitaxial methods[202], driven by the pursuit of faster material growth, inadvertently leads to a heightened level of defects in resulting thin films. Pits, dislocations and impurities can arise from a substrate's surface imperfections ultimately impacting the intrinsic interactions of magnetic moments[203].

The inherent topological stability of the skyrmion serves as its primary defense mechanism against small defects, propelling the skyrmion-based racetrack to the forefront of spintronic research over its domain-wall counterpart[143–145, 150, 151]. Essential to the functionality of a viable racetrack device is the unimpeded movement of skyrmions—fast yet predictable—under a preferably low spin-polarized current[1–3, 153–156]. In such systems, skyrmions exhibit high mobility in the presence of small and sparsely-distributed defects, distorting around defects[152, 204, 205], while experiencing only a marginal reduction in their speed[152, 160]. Conversely, domain walls are significantly hindered by defects, resulting in a notable decrease in speed and necessitating a considerable increase in current density to overcome these kinks[206, 207]. Unfortunately, further investigation reveals that larger or denser distributions of defects, often encountered in sputtered films, can impede the rapid motion of skyrmions[20, 204, 208–210], leading to stochastic trajectories[20, 211] with markedly reduced velocities[20, 205, 212–214] and increased transverse motion[204, 213, 214], sometimes even causing skyrmion pinning[20, 204, 205, 213]. Consequently, collisions between skyrmions and other pinned skyrmions,

large defects, or material boundaries, may result in their annihilation[204, 205, 208], jeopardising the non-volatility of data. Even the prototypical logic gates is vulnerable to defects, with skyrmions blocked behind defects within its confined geometry, resulting in erroneous outcomes[215].

Yet, controlled defects offer a promising avenue for predictably manipulating the properties[19, 20, 24, 216, 217], dynamics[218], and nucleation of skyrmions[160, 217, 219, 220], contingent upon a comprehensive understanding of their characteristics[221]. Ab-initio studies categorise atomic defects into two primary types: attractive or repulsive, determined by their chemical composition[6][228]. Strategically positioned attractive defects can serve as stable and predictable nucleation sites for skyrmions[227, 229]. These defects not only reduce the size of skyrmions[226] but can also minimise the spacing between them, thereby enhancing data density in storage devices. Skyrmions exhibit a high resilience to the defects, owing to their topological protection, which favours structural deformation[152, 204, 205] over annihilation (up to a particular defect level) when in close proximity to defects or other anchored skyrmions[222]. This enhanced stability promotes data preservation across a broader range of external conditions compared to defect-free systems, including at zero field using curvilinear defects[230]. Additionally, repulsive defects can be used to engineer skyrmion dynamics, guiding skyrmions along specific trajectories[225, 226, 231–233], altering their speeds[226, 232], trapping them[223, 224, 233–235], and regulating their passage within the confined geometry of logic gates[231, 236]. In the subsequent chapters, we exclusively examine the equilibrium behavior of skyrmion states under the influence of defects, leaving the exploration of dynamics for future research.

[6] Atomistically, these can be modeled as a pinning field incorporated into the various interaction strengths, for example fluctuations in the exchange strength, DMI, or anisotropy[20, 204, 209, 210, 213, 214, 218, 220–227].

Chapter 5

Validation of the mean-field model based on a 2D magnet Fe_3GeTe_2

The key purpose of this chapter is to validate the mean-field model, by comparing its predictions with experimental data obtained from the two-dimensional van der Waals magnets $Fe_{3-x}GeTe_2$, with Fe deficiency only, which was published earlier in [237, 238]. The author of this thesis supported the experimental work in these publications by performing the mean-field modelling as highlighted below. Below we present a brief summary of the results; for more details, the reader is referred to the original article[238].

5.1 Two-dimensional van der Waals magnet Fe_3GeTe_2

Two-dimensional (2D) van der Waals (vdW) magnets are atomic monolayer magnets that have been recently shown to sustain skyrmion structures. These materials are defined by atomically flat surfaces with no dangling bonds [239–242], making them well-suited to stacking into heterostructures, in comparison to typical sputtered multilayer skyrmion systems. The Fe_3GeTe_2 (FGT) [243–246] has been developed recently and is currently studied as prospective skyrmion host. In FGT, skyrmions are primarily stabilised by a combination of the dipolar interaction and strong out-of-plane anisotropy [247, 248]. However, our observation of monochiral Néel-type domain walls in thinner samples indicates that the dipolar interaction may not be sufficient for sustaining skyrmions, and the presence of some form of interfacial DMI may be necessary [237].

For the development of viable skyrmion-based devices, it is crucial to control material parameters to engineer specific skyrmion properties, such as size and stability. Previous studies on bulk skyrmion materials have demonstrated that modifying stoichiometry or substituting particular elements in the composition (e.g., $Mn_{1-x}Fe_xGe$, $Fe_{1x}Co_xSi$) can effectively modulate the Curie temperature (T_c) and skyrmion sizes. Recent investigations on Fe_3GeTe_2 (FGT) flakes have yielded promising results, including the discovery that skyrmion stability increases with additional layers of WTe_2 or Co/Pd in heterostructures. Furthermore, alterations to stoichiometry and chemical substitution have been shown to modify bulk material properties in FGT. Notably, Fe deficiency has been found to decrease magnetocrystalline anisotropy in FGT. However, the direct impact of these compositional changes on topological spin textures in FGT remains largely unexplored. This study aims to bridge this knowledge gap by systematically investigating the effects of Fe deficiency on the formation, stability, and characteristics of skyrmions and related spin textures in FGT, providing valuable insights for future spintronic applications.

In [238], we investigated both bulk and exfoliated flake samples of $Fe_{3-x}GeTe_2$, with Fe deficiency x between 0.03 and 0.37. Using a combination of magnetometry and real-space scanning transmission x-ray microscopy (STXM), we explored the effect that this altered composition has on the formation and stability of all observed spin textures, including stripe domain, skyrmion and skyrmionium states. Supporting mean-field and micromagnetic simulations confirm the vital role that the temperature varying uniaxial anisotropy plays in altering the magnetic phase diagrams of the material.

5.2 Magnetic Phase Diagram

In this section, we examine the material properties and magnetic phase diagrams of four bulk single crystals of $Fe_{3-x}GeTe_2$ with Fe deficiencies of x = 0.03, 0.10, 0.27 and 0.37. Detailed information on sample preparation and characterization can be found in the original article[238]. Figure 5.1a illustrates the sample magnetization as a function of temperature in an external out-of-plane field of 3 mT, revealing a clear reduction in Curie temperature (T_c) with increasing Fe deficiency. We also observe

an additional step in the M-T diagram, indicating a second transition temperature, T^*. While previous studies have attributed this to antiferromagnetic ordering or heavy fermion behavior, we propose that it may result from interlayer ordering between individual van der Waals (vdW) layers, occurring at a lower temperature than the intralayer ordering at T_c.

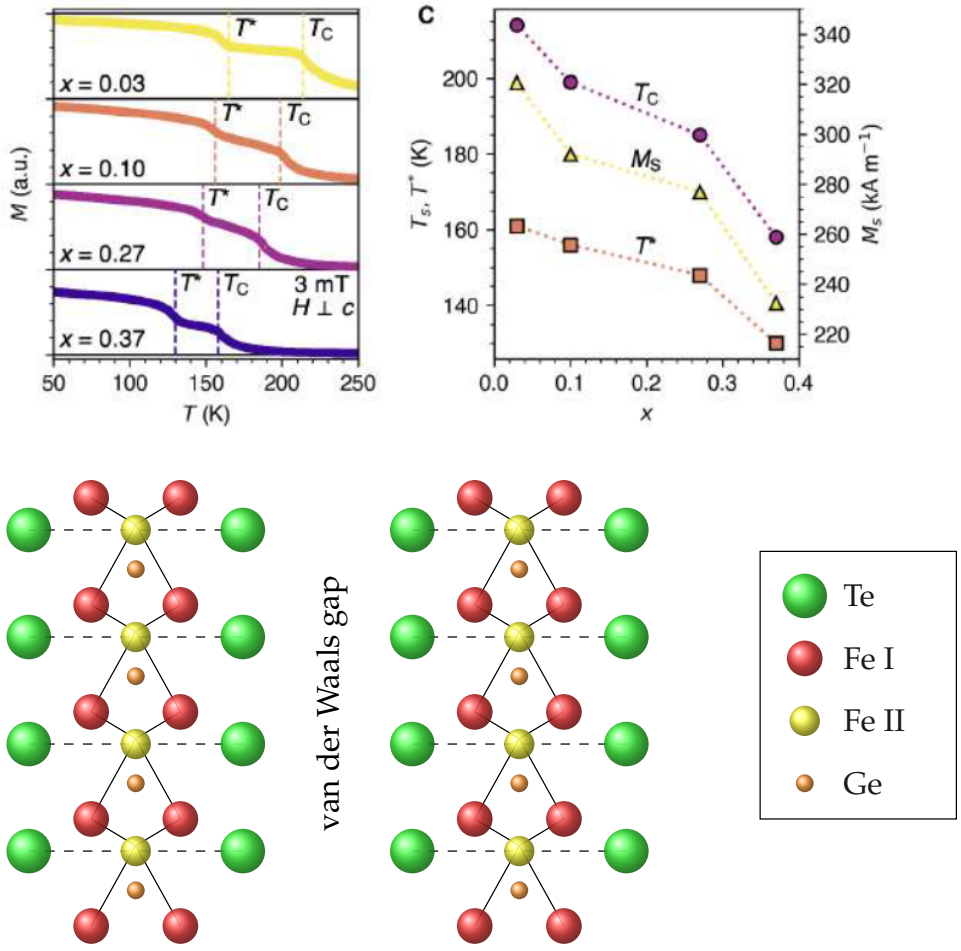

FIGURE 5.1: a) Magnetisation as a function of temperature for each Fe deficiency with an external out-of-plane field of 3 mT, with two critical temperatures, T_c and T^*. b) T_c, T^*, and M_S as a function of Fe deficiency, x. c) Side view of a two layers in Fe_3GeTe_2.

Figure 5.1b displays the behavior of T_c, T^*, and the saturation magnetization (M_s) as a function of Fe deficiency. The observed reduction in these properties can be attributed to the decrease in magnetic Fe ions within the crystal structure. As illustrated in Figure 5.1c, the crystal is composed of FGT layers separated by a vdW gap. The structure contains two types of Fe atoms: Fe I, which are positioned closer to the Te atoms, and Fe II, located midway between the Te layers. Previous studies have shown that Fe deficiency primarily occurs in the Fe II sites, while the Fe I sites

remain fully occupied.

From the bulk crystals with x = 0.03, 0.27 and 0.37, thin exfoliated flakes were mechanically cleaved and stamped onto Si_3N_4 membranes. Exposure to ambient atmospheric conditions resulted in 6 nm of oxidation on both sides of the samples. To prevent further oxidation, the flakes were capped with hexaboron nitride (hBN). Magnetic phase diagrams were generated using a field-sweep (FS) protocol. For each sample, at various temperatures, we initially established a saturated monodomain state at -250 mT. The applied field was then incrementally increased up to +250 mT, with Scanning Transmission X-ray Microscopy (STXM) imaging performed at each field value. Figure 5.2a-o presents a selection of X-ray micrographs acquired following this procedure, showcasing different temperatures, magnetic fields, and flakes with varying Fe compositions, x.

Each row corresponds to different temperature value normalised by the Curie temperature, T_c, of a given sample. The T_c of each flake sample was determined as the temperature at which real-space magnetic contrast could no longer be observed, with values slightly lower than those in the bulk: 207, 180 and 146 K for the x = 0.03, 0.27 and 0.37 compositions, respectively. The reductions of T_c are consistent with decreases seen in thinner flakes of 2D magnets [243], but we also cannot discount a temperature offset of a few Kelvin in the thermocouple measurement.

The results for the x = 0.03 flake, obtained at 203 K, or 0.98 T_c, are shown in Figure 5.2a. The images reveal the formation of a dense disordered array of skyrmions (labeled Sk) for both positive and negative applied fields, as well as stripe domains (labeled SD) at 0 mT. At decreasing temperatures, shown in Figure 5.2(b,c), the characteristic size of the stripe domains increases, and skyrmion formation is no longer observed. Finally, at 150 K (0.72 T_c) and below, only uniform switching between the positive and negative monodomain states (labeled MD$\pm$) is observed.

The high temperature results for the two flakes with greater Fe deficiency, x = 0.27 and x = 0.37, shown in Figure 5.2(f,g) and Figure 5.2(k,l), reveal a similar behavior, with dense skyrmion formation only observed close to T_c, and the characteristic stripe domain size increasing with decreasing temperature. However, at lower temperatures, there is a significant difference in the crossover to monodomain switching behavior. In comparison to the x = 0.03 flake, this occurs at a comparatively lower

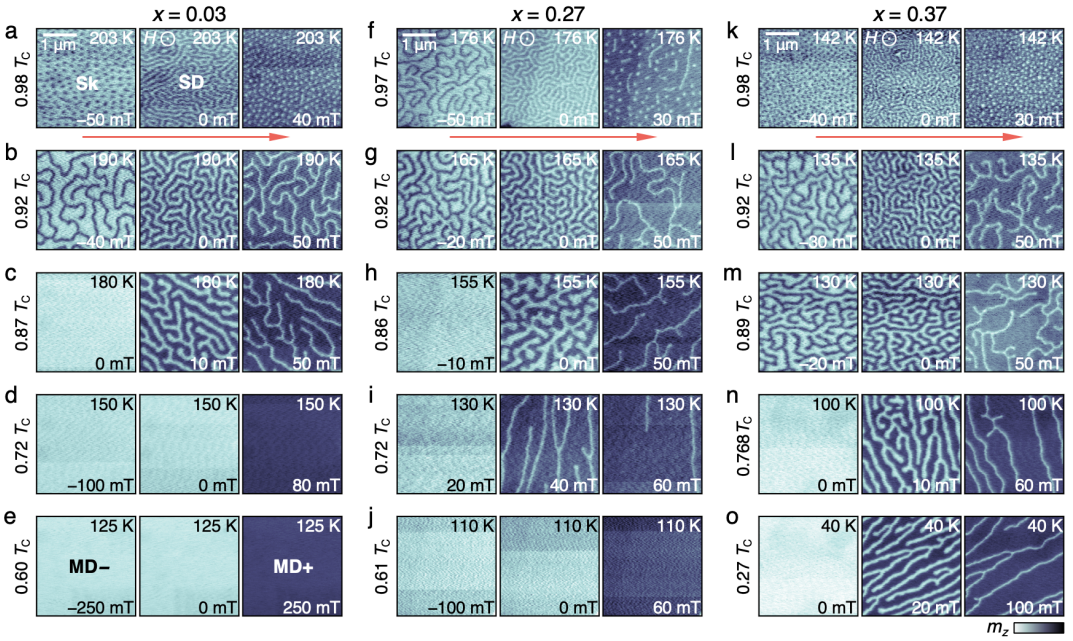

FIGURE 5.2: Scanning transmission microscopy measurements following the field-sweep procedure. a-o) X-ray micrographs of the $Fe_{3-x}GeTe_2$ (FGT) flake samples measured as a function of temperature and applied magnetic field for the x=0.03 (a-e), 0.27 (f-j) and 0.37 (k-o) flakes respectively. The images were taken as a function of increasing out-of-plane applied magnetic field, starting in the saturated state at -250, as indicated by the orange arrows. The temperature as a fraction of the Curie temperature T_c is labeled. The color map indicates the out-of-plane component of the magnetisation, m_z. Reproduced with permission from [238].

temperature in the $x = 0.27$ sample (0.61 T_c), revealed in Figure 5.2(h-j). Furthermore, in the $x = 0.37$ sample, monodomain switching behavior was not observed down to the base temperature of the STXM instrument at 30 K – instead, we observed the formation of stripe domains across the full investigated temperature range, as shown in Figure 5.2m-o.

The overall behavior is visible from the magnetic phase diagrams of each flake sample presented in Figure 5.3(a-c), which plot the observation of each magnetic state as a function of the applied magnetic field at each temperature when following the FS procedure. In cases where both stripe and skyrmion states coexisted, these were included in the skyrmion regions for clarity. The temperature is plotted on both an absolute scale, and as a fraction of T_c. The phase diagram of the $x = 0.03$ sample, with the crossover to the monodomain switching behavior, is similar to those previously reported for FGT flakes with x close to 0 [237]. However, the results of the $x = 0.37$ flake, where real-space spin textures were observed across the

full temperature range, is more reminiscent of typical dipolar-stabilised skyrmion bubble systems [249], as well as the multilayer skyrmion hosts [250].

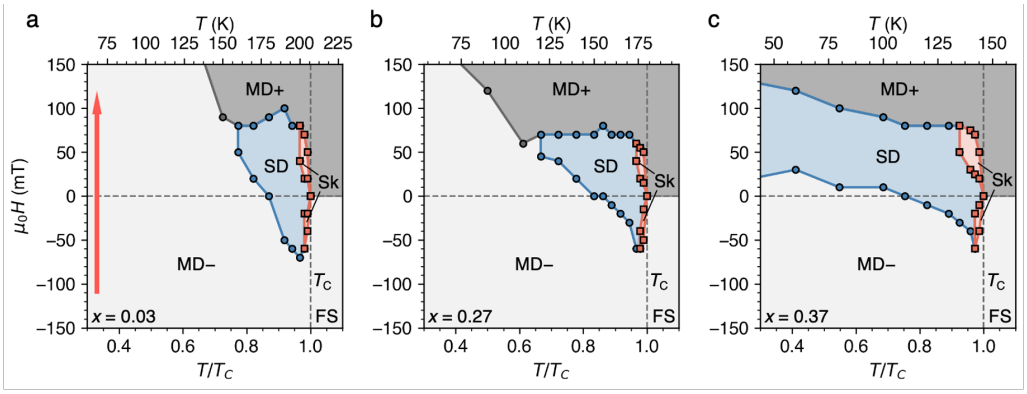

FIGURE 5.3: Composition dependent magnetic phase diagrams. a-c) Magnetic phase diagrams following the field-sweep procedure determined by x-ray microscopy of the three $Fe_{3-x}GeTe_2$ (FGT) flakes as a function of temperature and applied field. Results are shown for each composition x = 0.03, 0.27 and 0.37, as labeled. Red arrow indicates the measurement path of each field-temperature protocol. The extent of the skyrmion (Sk) stripe domain (SD) and uniformly magnetised monodomain (MD$\pm$) states is shown by the colored red, blue and grey regions, respectively. Markers indicate the measured phase boundary points of the Sk (red squares) and SD (blue circle) states. The vertical dashed lines indicate the measured values of T_c. The field was increased stepwise from -250mT to 250mT, in steps of 10mT. Reproduced with permission from [238].

5.3 Mean-Field modelling

To perform temperature-dependent simulations to validate the experimental phase plots shown in Figures 5.2 and 5.3 above, we employed the mean-field model discussed in Chapter 2. To include the essential physics of FGT samples, the Hamiltonian given in Equation 2.13 includes a uniaxial anisotropy energy term with coupling constant J_k, and also the dipolar energy term with coupling constant J_{demag}. In the mean-field formulation, this is given as follows,

$$\begin{aligned}\mathcal{H}_{\text{MF}} = &-\frac{1}{2}J_{ex}\sum_{\langle ij\rangle}(\mu\boldsymbol{m}_i\cdot\mu\boldsymbol{m}_j) - \frac{1}{2}J_{dm}\sum_{\langle ij\rangle}\boldsymbol{D}_{e,ij}\cdot(\mu\boldsymbol{m}_i\times\mu\boldsymbol{m}_j)\\ &- J_k\sum_i(\boldsymbol{m}_i\cdot\boldsymbol{e}_i)^2 - \mu\sum_i\boldsymbol{H}\cdot\boldsymbol{m}_i\\ &- \frac{1}{2}J_{\text{demag}}\sum_{ij}-\frac{\mu\boldsymbol{m}_i\cdot\mu\boldsymbol{m}_j}{r_{ij}^3} + 3\frac{(\mu\boldsymbol{m}_i\cdot\hat{\boldsymbol{r}}_{ij})(\mu\boldsymbol{m}_j\cdot\hat{\boldsymbol{r}}_{ij})}{r_{ij}^3}.\end{aligned} \tag{5.1}$$

The simulations in this section considered a two-dimensional hexagonal lattice of 30×30 spins, with periodic boundary conditions in the plane.

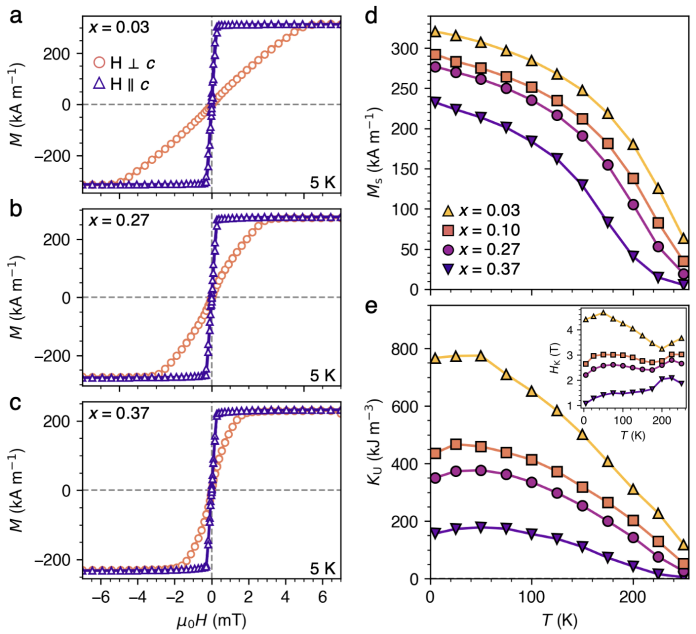

FIGURE 5.4: Composition dependence of magnetisation reversal and uniaxial anisotropy. a-c) Measurements of the magnetisation M versus applied field $\mu_0 H$, at 5 K, of each $Fe_{3-x}GeTe_2$ bulk crystal, with compositions x = 0.03, 0.27 and 0.37, respectively. Measurements were acquired with the magnetic field H applied both parallel (purple triangles) and perpendicular (orange circles) to the c crystalline axis. d) Extracted values of the saturation magnetisation M_S as a function of temperature T for each FGT composition, x = 0.03, 0.10, 0.27 and 0.37. e) Extracted values of the uniaxial anisotropy K_U of each FGT sample as a function of T. Reproduced with permission from [238].

The coupling constants were determined based on experimental magnetometry measurements conducted on bulk single crystals of FGT. Figure 5.4(a-c) shows M of FGTs of different compositions, measured at 5 K, as a function of magnetic field H applied both parallel and perpendicular to the film axis. From these plots, we extracted M_S, needed for calibrating J_{demag} constant in (5.1) to match experimental data, and also estimated values of K_U for each sample, needed for calibrating J_k constant in Eq. (5.1).

The parameters for each composition are plotted as a function of temperature in Figure 5.4(d,e). As expected, both M_S and the anisotropy constant K_U decrease with increasing Fe deficiency (x). However, the anisotropy field, calculated as $H_K = 2K_U/M_S$, varies with temperature irregularly (inset of Figure 5.4e), suggesting that the change in K_U between samples is not proportional to M_S.

Further analysis of the data in Figure 5.4 reveals a scaling behaviour that allows for the estimation of the compositional dependence of the model interactions in Eq. (5.1), via simple scaling laws. Starting from the mean field Hamiltonian, with shorthand symbols to simplify the notation,

$$\mathcal{H}_{\mathrm{MF}} = -\frac{1}{2}J_{ex}\mu^2\sum_{ex} -\frac{1}{2}J_D\mu^2\sum_{D} - J_K\sum_{K} -\frac{1}{2}J_{\mathrm{demag}}\mu^2\sum_{\mathrm{demag}}, \tag{5.2}$$

this equation can be made dimensionless by normalising by the anisotropy strength, J_K,

$$\frac{\mathcal{H}_{\mathrm{MF}}}{J_K} = -\frac{1}{2}\frac{J_{ex}\mu^2}{J_K}\sum_{ex} -\frac{1}{2}\frac{J_D\mu^2}{J_K}\sum_{D} - \sum_{K} -\frac{1}{2}\frac{J_{\mathrm{demag}}\mu^2}{J_K}\sum_{\mathrm{demag}}. \tag{5.3}$$

Two different systems, a and b, are equivalent if the following conditions hold,

$$\frac{J^a_{ex}\mu^2_a}{J^a_K} = \frac{J^b_{ex}\mu^2_b}{J^b_K}, \qquad \frac{J^a_D\mu^2_a}{J^a_K} = \frac{J^b_D\mu^2_b}{J^b_K}, \qquad \frac{J^a_{\mathrm{demag}}\mu^2_a}{J^a_K} = \frac{J^b_{\mathrm{demag}}\mu^2_b}{J^b_K}. \tag{5.4}$$

If the interaction strengths are known for system a, we can obtain the parameters for system b by rearranging,

$$J^b_{ex} = \frac{\mu^2_a J^b_K}{\mu^2_b J^a_K}J^a_{ex}, \qquad J^b_D = \frac{\mu^2_a J^b_K}{\mu^2_b J^a_K}J^a_D, \qquad J^b_{\mathrm{demag}} = \frac{\mu^2_a J^b_K}{\mu^2_b J^a_K}J^a_{\mathrm{demag}}, \tag{5.5}$$

These mean-field parameters are related to the experimentally observed parameters as given below,

$$J^b_{ex} = \frac{M^2_{S,a}K^b_U}{M^2_{S,b}K^a_U}J^a_{ex}, \qquad J^b_D = \frac{M^2_{S,a}K^b_U}{M^2_{S,b}K^a_U}J^a_D, \qquad J^b_{\mathrm{demag}} = \frac{M^2_{S,a}K^b_U}{M^2_{S,b}K^a_U}J^a_{\mathrm{demag}}, \tag{5.6}$$

where K_U and M_S are the experimentally-observed maxima for the anisotropy and magnetisation-temperature curves from Figure 5.4. These values are listed in Table 5.1. Starting from a set of model parameters for the $x = 0.03$ system, we can calculate parameters for each composition system, scaled by K_U and M_S, yielding the model parameters in Table 5.2. Note that these are dimensionless parameters, where the physical dimensions have been remove by additional normalisation, $\frac{\mathcal{H}_{\mathrm{MF}}}{\lambda}$, where $\lambda = 10^{-20}/\sqrt{3}$ J. Here, λ determines the absolute energy scale.

Fe deficiency, x	T_c **[K]**	$K_{U,max}$ **[kJm^{-3}]**	$M_{S,max}$ **[kAm^{-1}]**
0.03	213	790	320
0.27	185	400	280
0.37	156	200	240

TABLE 5.1: Parameters obtained from experimental anisotropy and magnetisation vs. temperature measurements.

Fe deficiency, x	J_E^*	J_D^*	$J_{\mathbf{demag}}^*$	J_K^*	T_c^*
0.03	0.700	0.100	0.350	2.500	2.25 ± 0.05
0.27	0.463	0.066	0.231	2.500	1.95 ± 0.05
0.37	0.315	0.045	0.158	2.500	1.75 ± 0.05

TABLE 5.2: Parameters of the model Hamiltonian.

We performed simulations of each FGT system, following a field-sweep procedure starting at negative applied fields, with a summary of the results shown in Figure 5.5. Additional data can be found in the supporting information in [238]. Simulations 1, 2 and 3 correspond to parameters selected to model the x = 0.03, 0.27 and 0.37 experimental systems, respectively. The visualisations of the simulations in Figure 5.5(a-c) show a good agreement with the experimental behavior in Figure 5.2. The formation of skyrmions at high temperatures is reproduced, while the observed crossover from stripe domain formation to uniform magnetisation switching is evident in simulations 1 and 2.

The results are better visible in the simulated magnetic phase diagrams in Figure 5.5(d-f). Comparison to Figure 5.3(a-c) shows a reasonable agreement to the experimentally determined phase diagrams. In particular, specific features such as the applied field asymmetry of both the high temperature skyrmion pockets and the stripe domain formation (due to the field asymmetry) are reproduced. The main features that are poorly replicated are the switching fields at low temperatures, which might be explained by the presence of thermal fluctuations or defects allowing easier switching in the experimental system.

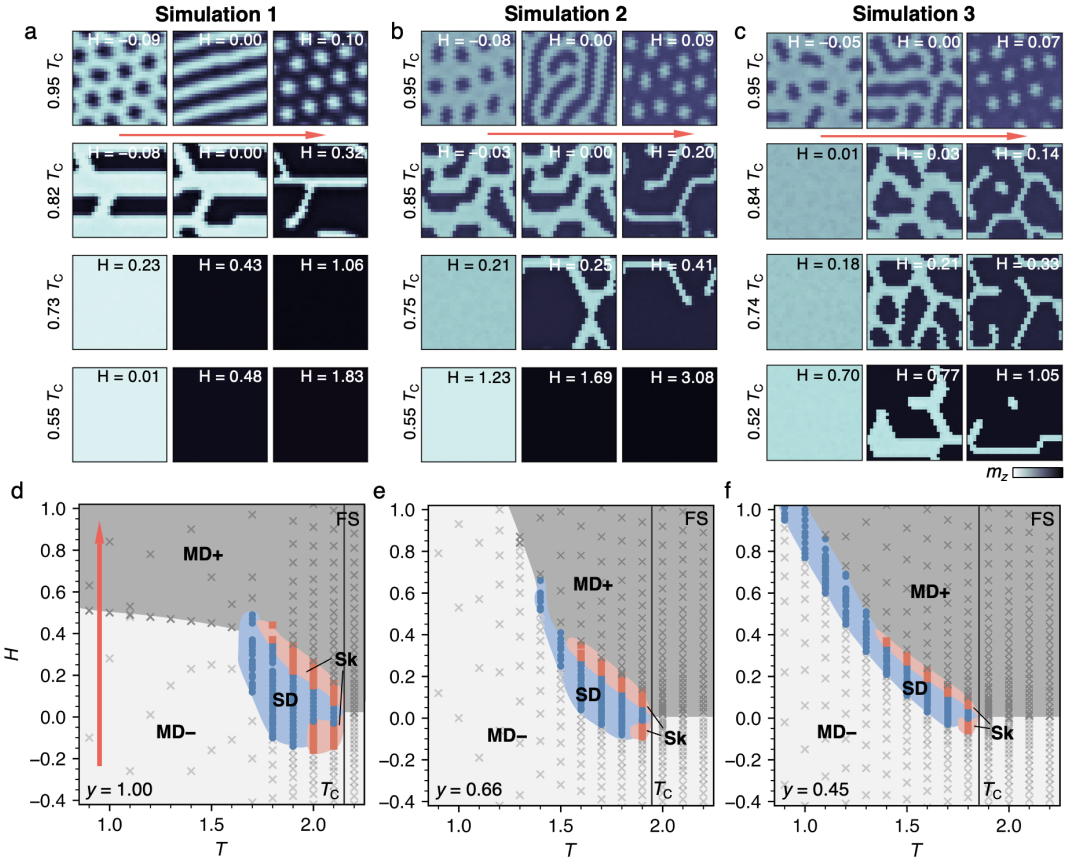

FIGURE 5.5: Mean-field simulations modeling $Fe_{3-x}GeTe_2$ flakes. Simulations 1-3 correspond to systems based on parameters for the x = 0.03, 0.27 and 0.37 experimental compositions, respectively. a-c), Selected visualisations of simulated states acquired following a field sweep procedure starting from negative applied fields. The color map indicates the out-of-plane component of the magnetisation, m_z. d-f) Simulated magnetic phase diagrams of the three systems. The extent of the skyrmion (Sk) stripe domain (SD) and uniformly magnetised monodomain (MD±) states is shown by the colored red, blue and gray regions, respectively. Markers indicate the sampled positions in phase space. The simulated values of T_C are indicated by the vertical lines. Reproduced with permission from [238].

5.4 Conclusions

The mean-field modelling qualitatively reproduces all of the observed features in the X-ray microscopy results of the magnetic spin textures hosted in exfoliated flakes of $Fe_{3-x}GeTe_2$. Namely, the model recovers the observed formation of stripe domain, skyrmion and composite skyrmion states, and the underlying transition paths between them.

Chapter 6

Effect of Defects on the Formation of Ordered Skyrmion Lattices

6.1 Motivation

In experimental studies, it has been hypothesised that defects can stabilise the skyrmion region over a broader range of external fields and temperature. This hypothesis is examined in the following chapter, where we investigate the impact of defects—modeled as a random pinning (RP) field—on the formation of spin textures during a field sweep. Our computational methods, based on mean field algorithms, offer a more comprehensive analysis of defects, leading to insights beyond those obtained from single snapshots in real materials. Our study confirms that the broadening of the skyrmion region can be attributed to defects, resulting in a highly disordered arrangement of skyrmions (DSkL) in 'dirty' systems compared to the ordered skyrmion lattice (OSkL) observed in 'clean' systems. Furthermore, we hypothesise that there is a disorder-driven phase transition from the OSkL to DSkL and we determine the critical defect level necessary to disrupt the regular ordering of skyrmions across various skyrmion sizes.

6.2 Method

We simulated a hypothetical material with no anisotropy or dipolar interactions, using parameters $J_{ex} = 0.5$ and $J_{dm} = 0.5$. For a specific defect level, we generated the RP field[1] and applied the Monte Carlo (MC) method in Section 6.3 and the mean

[1] For further details, see Chapter 2, Section 2.3.

field (MF) algorithm[2] for subsequent analysis. Both algorithms were run through a field sweep at a fixed temperature from $H = -1.0$ to $H = 1.0$, with fine intervals of $dH = 0.01$, on a 2D 50×50 hexagonal lattice, initiating from a well-defined uniform state, with all spins aligned to $-z$. In subsequent field steps, the final configuration of the previous field served as the input for the current field step. At the end of each simulation at every field step, we recorded the final spin configuration and calculated key macroscopic properties such as skyrmion number and average magnetisation. Additionally, we generated reciprocal-space FFT images to mimic experimental small-angle neutron scattering (SANS) output.

With mean-field modelling, this process was repeated for 81 different temperatures between $T = 0.0$ and $T = 2.0$, generating a full $H - T$ phase diagram, with 1000 realisations of the underlying RP field. For MC modelling, the next section highlights the challenges encountered in using this algorithm, and consequently only a single hysteresis loop was run for only one realisation of the RP field, using 1000 equilibrium steps and 10000 MC steps. Macroscopic properties were thermally averaged over all MC steps [3].

6.3 Analysing defects using Monte Carlo simulations

We begin our analysis by examining the equilibrium spin configurations generated using the MC method, which provides a more accurate representation of the impact of thermal fluctuations on magnetic phases. The evolution of the spin configurations along hysteresis loops for $T \in \{0.2, 0.5, 1.0\}$, for a single realisation of a 'clean' system ($H_\sigma = 0.001$), is shown in Figure 6.1. Note that in these figures, the skyrmion number is quoted as a long-run average over the MC steps ($\frac{1}{N}\sum_i N_{sk,i}$), which can result in a non-quantised value as skyrmions nucleate/annihilate along the MC path. In 'clean' system, skyrmions with varying shapes and sizes, approximately arranged in a hexagonal lattice[4], are noticeable at lower temperature, but this becomes increasingly obscured at higher temperature due to the thermal noise. By $T = 1.0$, the presence of skyrmions becomes undetectable, though the thermally

[2] This is outlined in Algorithm 4, Chapter 2.
[3] See Chapter 2, Section 2.4.1 for further detail.
[4] A perfectly ordered hexagonal lattice would show six peaks in reciprocal space.

averaged skyrmion number gives a non-zero value for these spin textures. Further analysis, given below, reveals that the algorithm for N_{sk} fails in the presence of significant defects, such as those produced by the thermal noise.

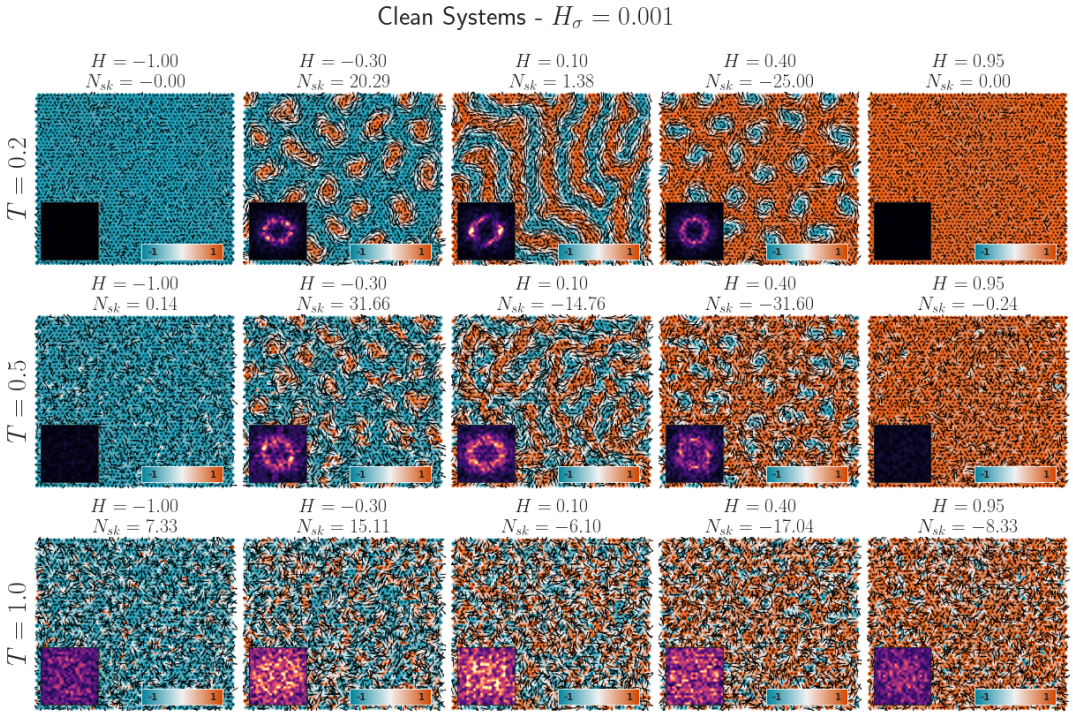

FIGURE 6.1: Evolution of the spin configuration along a hysteresis loop for $T \in \{0.2, 0.5, 1.0\}$ in a single realisation of a 'clean' system ($H_\sigma = 0.001$). The inset shows the reciprocal space of the real-space image, found via fast Fourier transform (FFT).

Figure 6.2 compares the N_{sk} calculation for a single skyrmion (A) and a spin defect point (B). The diagram shows the intermediate step in this algorithm, computing the sum of the solid angle formed by adjacent spins on two opposing triangles at each lattice point, and the skyrmion number is the total sum of these solid angles across the full texture[5]. For a skyrmion with a core in $-z$, the solid angle is at its maximum at the center of the skyrmion before gradually decreasing to 0 near the boundary. Conversely, a spin defect, such as those produced by MC noise in Figure 6.1, has a very different signature to the skyrmion, featuring a large negative solid angle at the center that is now surrounded by a positive solid angle near the boundary. Unfortunately, though these defects are not skyrmions, they are also counted towards the skyrmion number calculation, leading to an erroneous output for N_{sk}. Consequently, reliably detecting skyrmions at high temperatures is challenging both

[5]Further details are given in Chapter 2, Section 2.5

from the real-space MC-generated configuration and from its subsequent N_{sk} calculation.

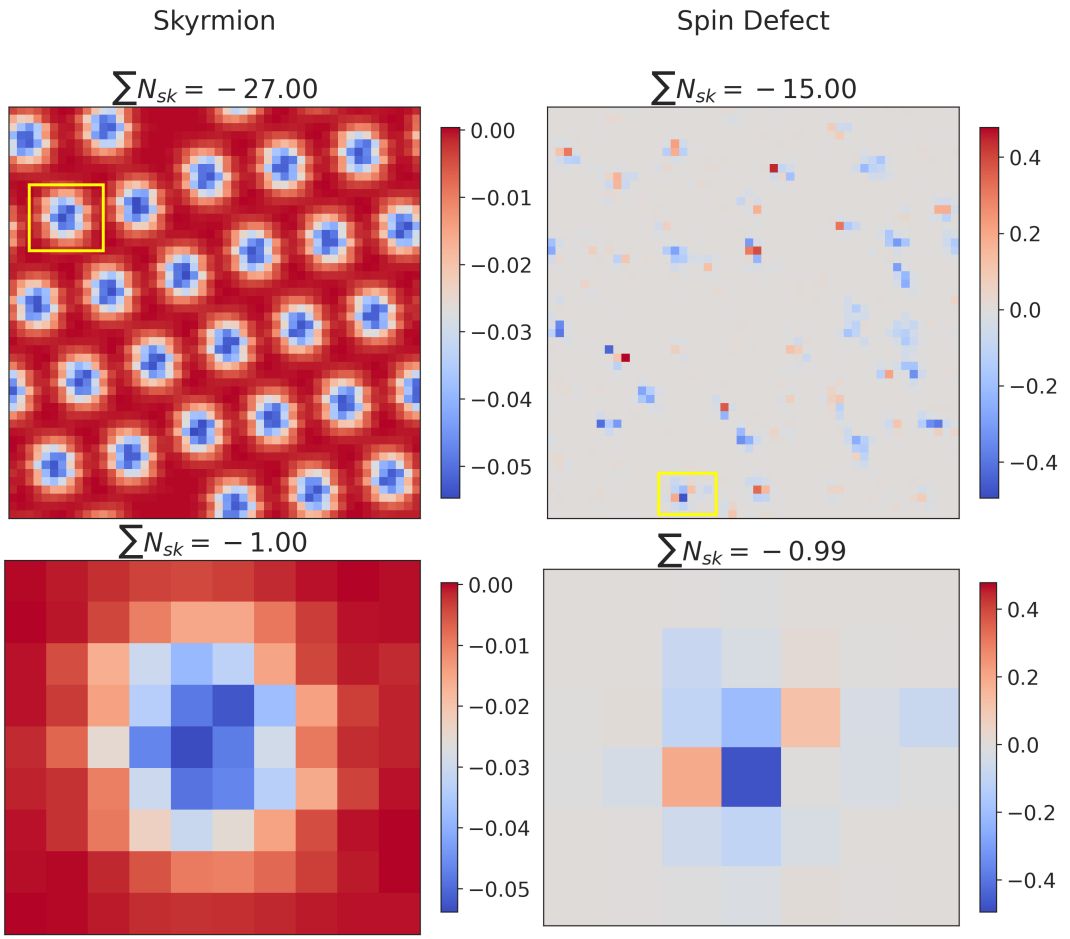

FIGURE 6.2: Comparison of the N_{sk} calculation for a skyrmion and a spin defect. The top row shows the full configuration, while the bottom row shows a single skyrmion/spin defect that is outlined in yellow. The color represents the summed solid angle at each lattice point, formed by the adjacent spins of the two opposing triangles.

When considering a 'dirty' system with $H_\sigma = 0.1$, it similarly exhibits a skyrmion phase at the same field strength as the 'clean' system ($H = -0.3$), but now skyrmions are smaller and more numerous in the presence of defects. These skyrmions are arranged in a less ordered formation, as indicated by the larger area covered by the FFT ring pattern compared to the 'clean' system, potentially caused by the pinning of certain skyrmions to particular defect locations[36, 251]. Moreover, a greater number of skyrmions persist at lower fields, effectively separating the long meandering domains at $H = 0.1$. Encouragingly, we will encounter a similar impact of defects

using a mean field model, despite this model averaging over the thermal fluctuations, allowing the use of MF models over MC.

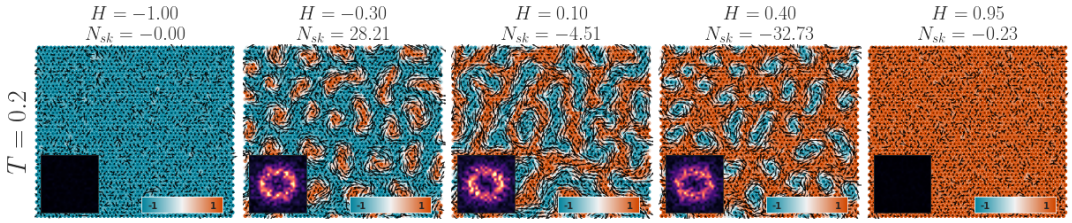

FIGURE 6.3: Evolution of the spin configuration along a hysteresis loop for $T = 0.2$ in a single realisation of a 'dirty' system ($H_\sigma = 0.1$).

Performing a systematic study of defects using only the Monte Carlo method would be challenging, largely due to the difficulty of detecting skyrmions, even in 'clean' systems, and this is further compounded by the slow speed of the algorithm. A single run, with 10000 MC steps and 40 field steps in the hysteresis loop, required approximately $\sim$2500s for only one realisation of the defect field. The algorithm's efficiency is further impacted by the multiple runs, necessary for improved accuracy of the thermal averages, and the several realisations for each defect level, required for a systematic study of defects. In other words, the computation is of order $\mathcal{O}(\text{MCRuns} \times \text{NumRealisations})$. In contrast, due to the deterministic nature of the mean field model, it is of order $\mathcal{O}(\text{NumRealisations})$, and takes only $\sim$150s for a single realisation of a 'clean' system and $\sim$70s for a 'dirty' system. Consequently, in subsequent analysis, we opt for the mean field model to conduct further studies into the effects of defects on equilibrium magnetic states.

6.4 Effect of random pinning fields on the $H - T$ phase diagrams

An analysis of the impact of defects on spin configurations, requires a comparative study involving clean ($H_\sigma = 0.001$) and dirty ($H_\sigma = 0.1$) systems, using the MF model.

6.4.1 Magnetic behaviour of 'Clean' Systems - $H_\sigma = 0.001$

We initially consider the 'clean' case, which serves as our baseline for comparison. At $T = 1$, a 'clean' system transitions through five distinct phase, as shown in Figure 6.4. In order, the phases are as follows: 1. **Unf(-)**: *uniform state*, where spins are aligned in $-z$, 2. **SkL(+)**: an ordered hexagonal *skyrmion lattice*, with cores in $+z$, 3. **Helical**: *helical* state, with spin spirals, 4. **SkL(-)**: *skyrmion lattice*, with cores in $-z$, 5. **Unf(+)**: *uniform state*, with spins aligned in $+z$.

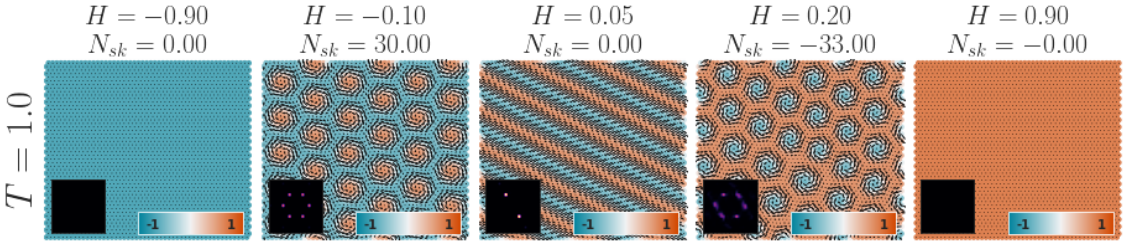

FIGURE 6.4: Evolution of the spin configuration along a hysteresis loop for $T = 1.0$ in a single realisation of a 'clean' system ($H_\sigma = 0.001$), generated by the mean field model. The inset shows the reciprocal space of the spin texture.

Experimentally, the aforementioned phases can be identified via their peaks in the reciprocal space (inset in Figure 6.4), which reflect variations in the long-range ordering of structures in the spin patterns. The hexagonal ordering of skyrmions results in six peaks for both skyrmion core orientations (SkL(+) and SkL(-)). Conversely, the long-range order of spin spirals, which orient in a single direction, are characterised by only two peaks. In certain realisations, however, defects can pin the spin spiral branches, causing a change in their orientation, leading to a smearing of the two peaks (see Figure 6.5). Later, when considering the impact of larger defects (Section 6.4.2), differentiating the phases using reciprocal space patterns becomes even more challenging.

The transition boundaries between the different phases can be found via the macroscopic properties of the system, namely the skyrmion number, N_{sk}, and average z-magnetisation, M_z (see Figure 6.6). These properties are averaged over 1000 realisations of the RP field to account for the small level of defects in our system, and N_{sk} is further normalised by the maximum number of skyrmions in the lattice. The behaviour of N_{sk} distinctly delineates phases of differing topological classes,

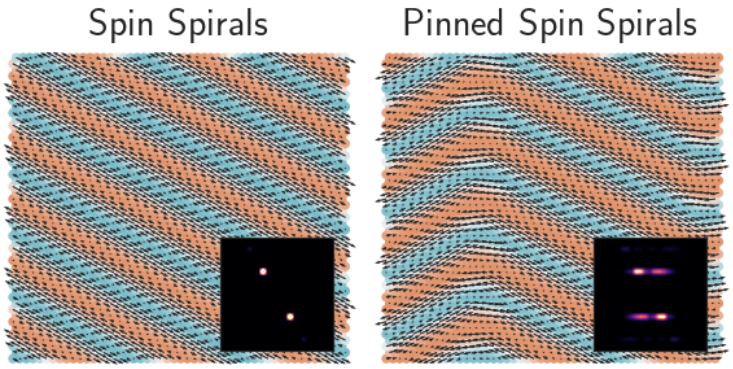

FIGURE 6.5: Spin configuration for an unpinned and pinned spin spiral state, taken at different realisations of $H_\sigma = 0.001$, together with its reciprocal space pattern.

showing a sharp jump at each transition, indicative of a first-order transition.

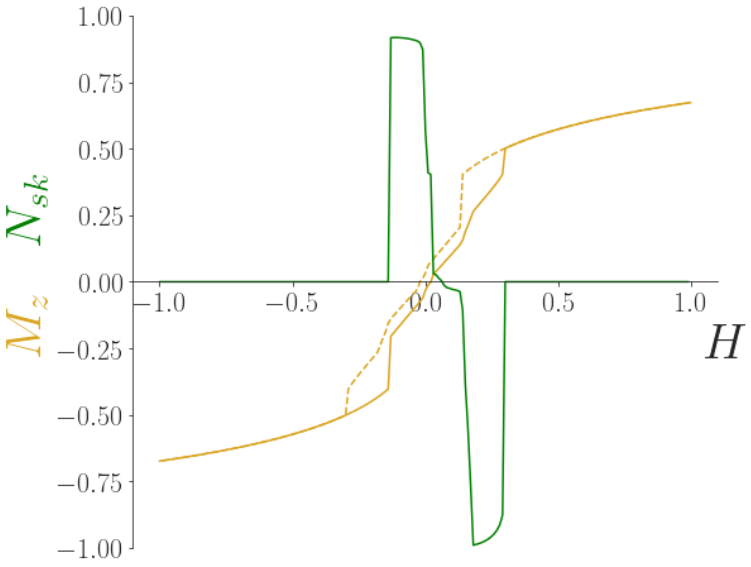

FIGURE 6.6: Spin configurations for an example realisation (first row) and averaged macroscopic properties, N_{sk} and M_z, (second row) along a hysteresis loop in a 'clean' system ($H_\sigma = 0.001$) at $T = 1.0$.

When comparing systems with low defect levels ($H_\sigma = 0.001$) and no defects ($H_\sigma = 0.0$ - Figure 6.7), we observe minor additional impacts resulting from the existence of small defects. One notable effect is the presence of non-zero N_{sk} in the helical phase. This occurs because defects, even small ones, pin the spin spiral branches, causing them to separate and form ends that are counted as half-skyrmions by our N_{sk} algorithm.

Another effect is a decrease in the magnitude of the jump at the transitions exiting both SkL phases. This reduction is caused by a gradual decrease in N_{sk} over the

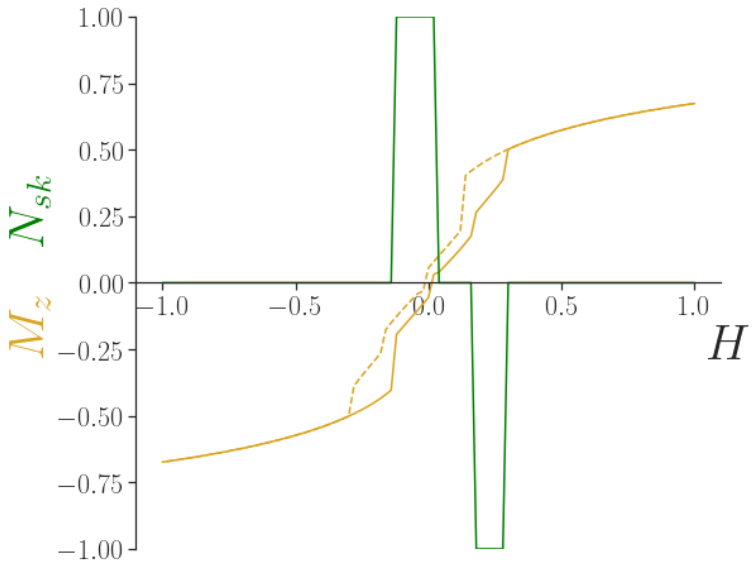

FIGURE 6.7: Averaged macroscopic properties, N_{sk} and M_z, in a defect-free material, $H_\sigma = 0.0$.

SkL phases, indicating that initially a few skyrmions annihilate before the remainder. Conversely, in defect-free materials, N_{sk} remains constant throughout the SkL phases (Figure 6.7). A similar effect occurs at the SkL(+)-Helical boundary, where the jump in N_{sk} splits into two separate jumps, suggesting that, on average, half the skyrmions annihilate before the other half. In both scenarios, this phenomenon is likely induced by the non-zero RP field, which shifts the local energy landscape near defect points, destabilising skyrmions[228, 252]. Consequently, it appears that a small level of frozen defects smears the first-order transition, a phenomenon noted in other studies[252].

The final macroscopic parameter of interest is M_z (see Figure 6.6). Like N_{sk}, this property similarly shows discontinuities at the phase transitions, albeit with a significantly smaller magnitude of jump. This effect is particularly notable at the SkL(+)-Helical boundary, where the smearing of the first-order transition obscures the transition. In defect-free systems, M_z more clearly delineates the SkL(+)-Helical transition, although the jump magnitude remains small (Figure 6.7). However, M_z does display a unique range of values in each of the five phases; in contrast, N_{sk} fails to differentiate between the non-topological phases—the uniform state and the pure helical state—yielding $N_{sk} = 0$ for both. Moreover, the peaks in the FFT fail to differentiate between SkL(+) from SkL(-), or between Unf(-) from Unf(+). Hence, it

appears that all these parameters struggle to differentiate between all phases in this system, a problem that we further explore in Chapter 7.

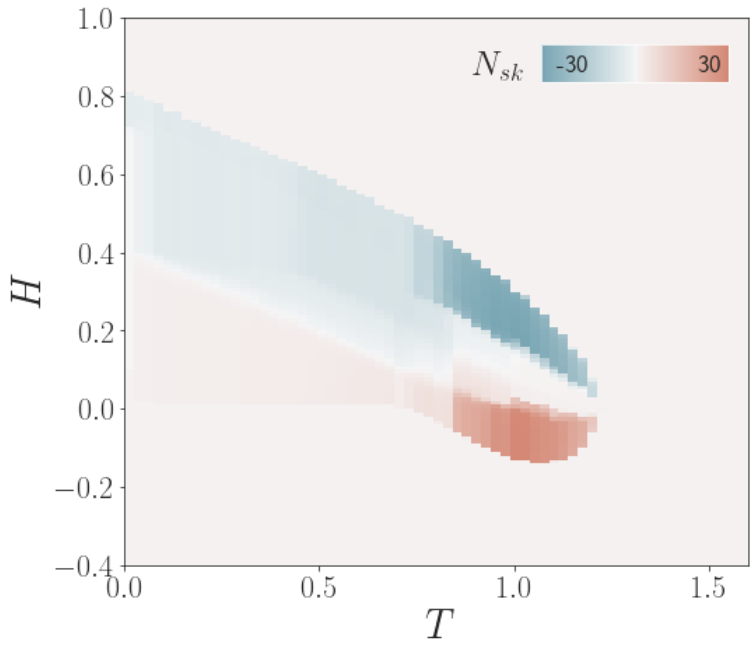

FIGURE 6.8: Phase diagram for a 'clean' system with $H_\sigma = 0.001$, with skyrmion number, N_{sk}, averaged over 200 realizations, as a function of temperature, T, and applied field, H.

Generating hysteresis loops for a range of temperatures in $T \in [0.01, 2]$ produces the phase diagram shown in Figure 6.8, where the colour represents the average N_{sk}. This diagram reveals two pockets of SkLs close to the Curie temperature ($T_c = 1.2\ J/k_b$), one at positive fields and the other at negative fields, with a sharp boundary to the non-topological phases. We further observe that the pocket at positive field expands to low temperatures with fewer skyrmions, but the pocket at negative field is is suppressed at these temperatures.

The spin configurations at $T = 0.5$ reveal that, after the initial transition into a meandering domain state, these wormy structures coalesce to a few isolated skyrmions, randomly distributed within the system. Upon stabilisation into the isolated skyrmion state (Sk(-)), subsequent increases in the external field lead to the gradual shrinking of the skyrmions until their eventual annihilation. The key feature of this Sk(-) state is the absence of peaks in reciprocal space, despite the existence of skyrmions, potentially causing it to be overlooked when analysing solely based on experimental SANS data.

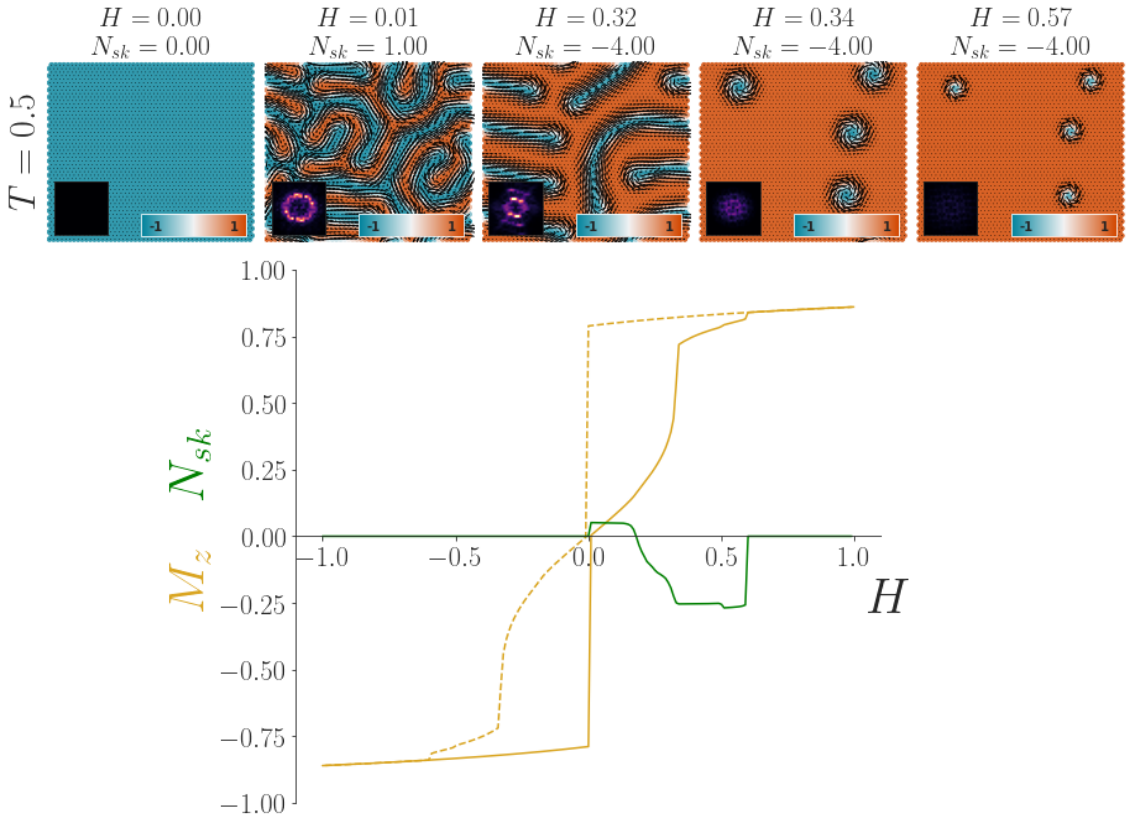

FIGURE 6.9: Spin configurations for an example realisation (first row) and averaged macroscopic properties, N_{sk} and M_z, (second row) along a hysteresis loop in a 'clean' system ($H_\sigma = 0.001$), at $T = 0.5$.

6.4.2 Magnetic behaviour of 'Dirty' Systems - $H_\sigma = 0.1$

We now consider the 'dirty' system, characterised by a substantial level of defect, $H_\sigma = 0.1$. The evolution of spin configurations at $T = 1.0$, illustrated in Figure 6.10, show a more gradual proliferation of skyrmions in the system, starting from a single skyrmion (Sk(+)) at the nucleation point ($H = -0.28$), to a disordered array of skyrmions (DSkL(+)) of varying shapes and sizes ($H = -0.12$). Both the diversity in skyrmion shapes and sizes in the vicinity of defects, as well as the subsequent disordering of the skyrmion lattice, have been documented in numerous prior studies[19, 20, 24, 152, 216].

One important characteristic of highly defective materials is the significant difference in the peaks in reciprocal space (inset of 6.10) from those observed in clean systems. Specifically, a ring pattern emerges for both the DSkL(+) and meandering domain state, contrasting sharply with the six peaks observed in SkL(+) and the

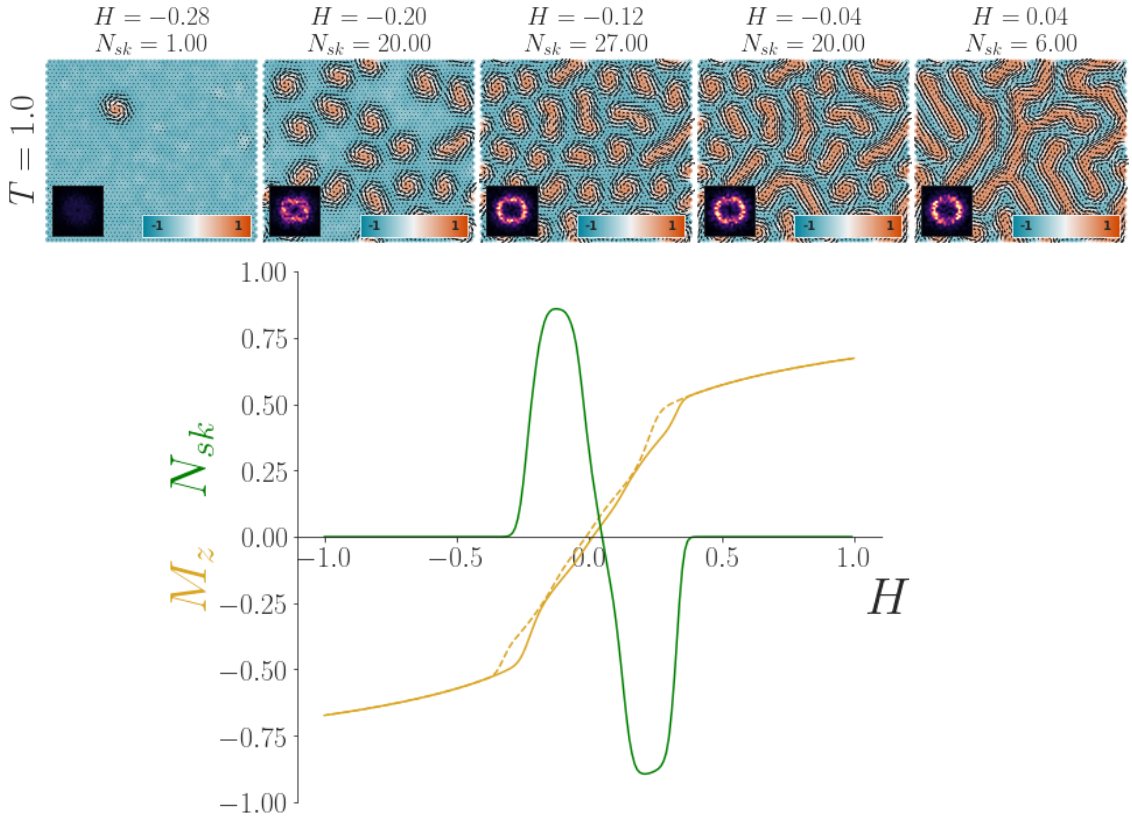

FIGURE 6.10: Spin configurations for an example realisation (first row) and averaged macroscopic properties, N_{sk} and M_z, (second row) along a hysteresis loop in a 'dirty' system ($H_\sigma = 0.1$) at $T = 1.0$.

two peaks observed in the clean helical phase. This ring pattern prevents the differentiation between the meandering domain and DSkL(+) using SANS data, and the presence of skyrmions can be missed in experimental results.

The macroscopic parameters in Figure 6.10 have notably different behaviour from the 'clean' system in Section 6.4.1, now exhibiting a more gradual change in the parameters, as opposed to the discontinuous jump between the various phases. For 'clean' systems, even the introduction of a small level of defects can smear the first order transition of defect-free systems, and this blurring is amplified at higher defect levels. Consequently, by $H_\sigma = 0.1$, the helical/meandering domain phase is undetectable solely from N_{sk}, M_z and even the FFT pattern, despite the distinct lack of skyrmions in the real space configuration at $H = 0.04$. N_{sk} no longer has a constant value of 0, characteristic of a pure spin spiral state, instead defects continuously divide the meandering domains throughout the field sweep, leading to a varying N_{sk}. As a result, now only 3 phase transitions are apparent, two of which are

between the uniform and skyrmion state, and the final transition is from skyrmion with cores in $+z$ to a phase with skyrmion cores in $-z$. Even with a machine learning approach, explored in Chapter 7, identifying macroscopic parameters that can detect the intervening non-topological phase, clearly observable in 'clean' systems, remains a challenging task.

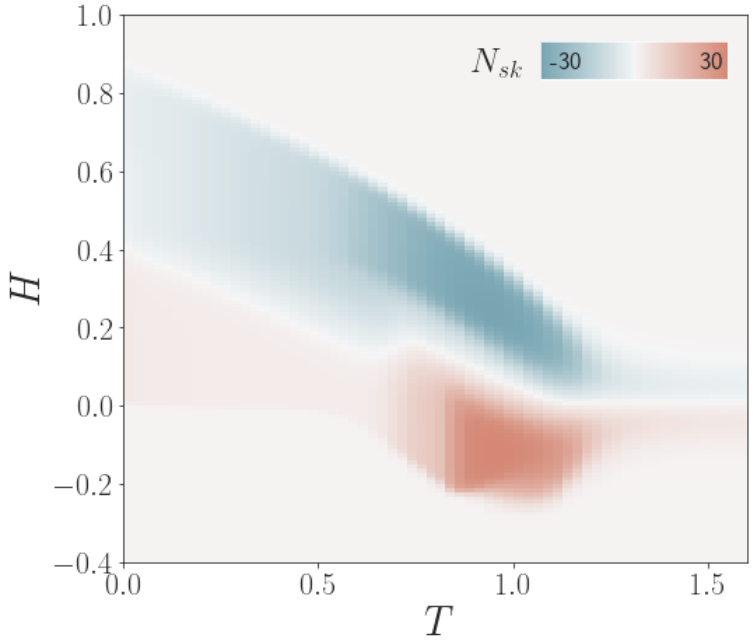

FIGURE 6.11: Phase diagram for a 'dirty' system with $H_\sigma = 0.1$, with skyrmion number, N_{sk}, averaged over 1000 realizations, as a function of temperature, T, and applied field, H.

The complete $H - T$ phase diagram for the 'dirty' system, depicted in Figure 6.11, showcases similar skyrmion pockets seen in the 'clean' case. However now, there is a more pronounced blurring between the boundaries of these regions and the non-topological phases, leading to a broader skyrmion area in the $H - T$ space, with greater variation in N_{sk}. This expansion of the phase region is similarly observed in experimental setups, where the material is subjected to mechanical strain[16], epitaxial stress[253], or impurity doping[13, 17, 40], potentially leading to an increased level of defects[6]. These findings affirm that the underlying defect field is a major contributor to the enlargement of the skyrmion region.

[6]See Chapter 4 for further details.

6.5 Phases along a varying defect level

We investigate system behaviour under varying levels of defect, $H_\sigma \in [0.001, 0.1]$, with the evolution of the spin configurations at negative Zeeman fields illustrated in Figure 6.12. The first spin texture depicts the initial nucleation of the skyrmion state (also called the skyrmion nucleation boundary), and the final configuration is the helical/meandering domain state at $H = 0$. At lower levels of defect, the initial configuration is the full hexagonally-ordered skyrmion lattice with its characteristic six peaks in FFT. This phase remains relatively stable until at higher fields, skyrmions begin merging, likely mediated by a Bloch point[7] [191], before it switches abruptly to a spin spiral state, with two FFT peaks.

At higher defect levels ($H_\sigma = 0.025$), we now only observe a single skyrmion at the nucleation boundary, occurring at a lower Zeeman field, but the full lattice nucleates at the subsequent field step. Thus it appears that defects can shift the uniform-skyrmion transition boundary, stabilising the skyrmion lattice at lower fields. Unfortunately, defects also encourage the annihilation of more skyrmions at lower fields compared to 'cleaner' systems, though some skyrmions continue to persist down to $H = 0$, leading to a mixed state of domains and skyrmions. For higher defect levels ($H_\sigma = 0.05$), the nucleation of skyrmions to the full lattice occurs over a lower but wider range of fields, and the final lattice structure is less ordered and contains fewer skyrmions than the 'clean' system. Consequently, this random distribution of defects, which is common in sputtered thin films, is detrimental for device prototypes where the skyrmions represent data particles[8]. The unpredictability of skyrmion nucleation and annihilation due to defects undermines data reliability, and the reduction in the number of skyrmions decreases the density of stored data. Therefore, our work seeks to identify the critical defect level below which the arrangement of skyrmions remains ordered.

From these preliminary studies, we map a hypothetical phase diagram for our dataset as a function of the Zeeman field, H, and the level of defect, H_σ (Figure 6.13). Our objective is to find a macroscopic property capable of delineating these phases, exhibiting a sharp transition at all phase boundaries for hysteresis loops at low H_σ

[7] See Chapter 4 for more information on the skyrmion decay process.
[8] See Chapter 4 for more information on the impact of defects on skyrmion device prototypes.

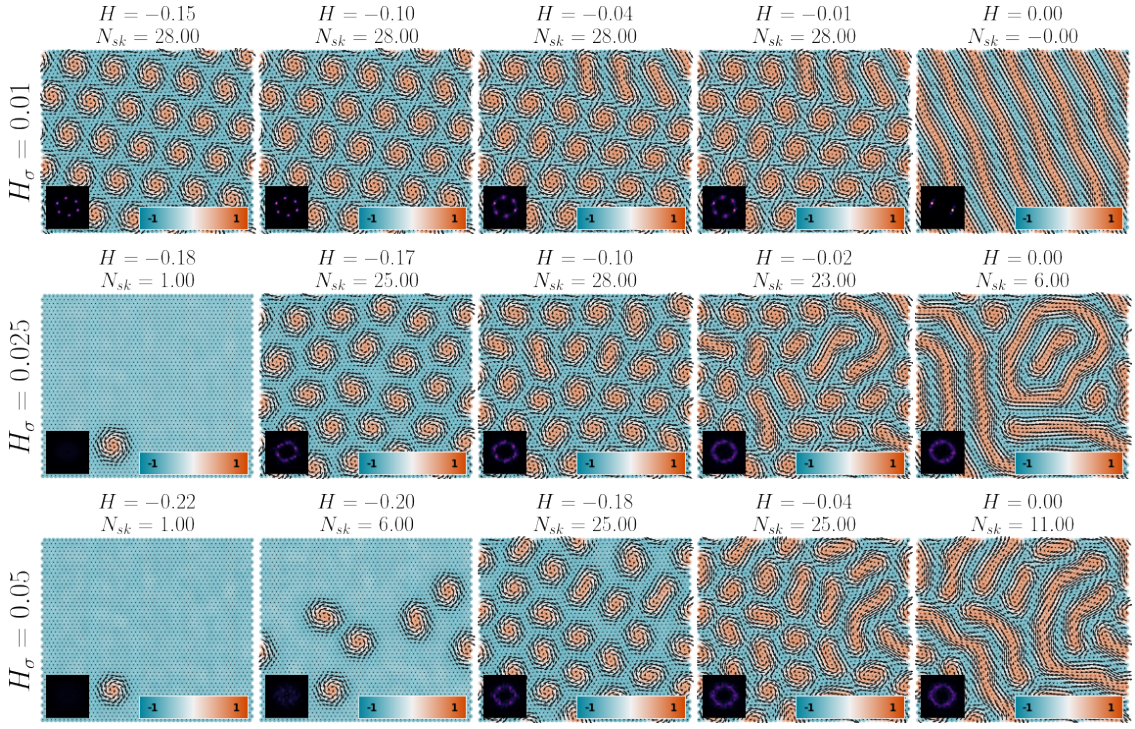

FIGURE 6.12: Evolution of the spin configuration along a hysteresis loop for $T = 1.0$, between the skyrmion nucleation boundary and $H = 0$, for a single realisation with defect levels, $H_\sigma \in \{0.01, 0.025, 0.05\}$.

levels, while the transition would be more gradual for higher H_σ levels. Moreover, there should be a smooth transition along H_σ as the system progresses from the OSkL to the DSkL, or from the clean helical state to a mixed state of meandering domains and skyrmions. To simplify our analysis, we focus solely on the transition boundary between OSkL and DSkL along the skyrmion nucleation boundary, k, rather than at a fixed external field. This simplifies our analysis because 'dirty' systems only exhibit one or a few skyrmions here, making them easier to identify than the mixed states present at higher fields. Additionally, the nucleation boundary provides a more consistent phase path when comparing systems with different Hamiltonian parameters, which may have differing phase regions. The next section introduces the macroscopic variable under consideration: the spin-spin correlation mapping between adjacent spin textures along a hysteresis loop.

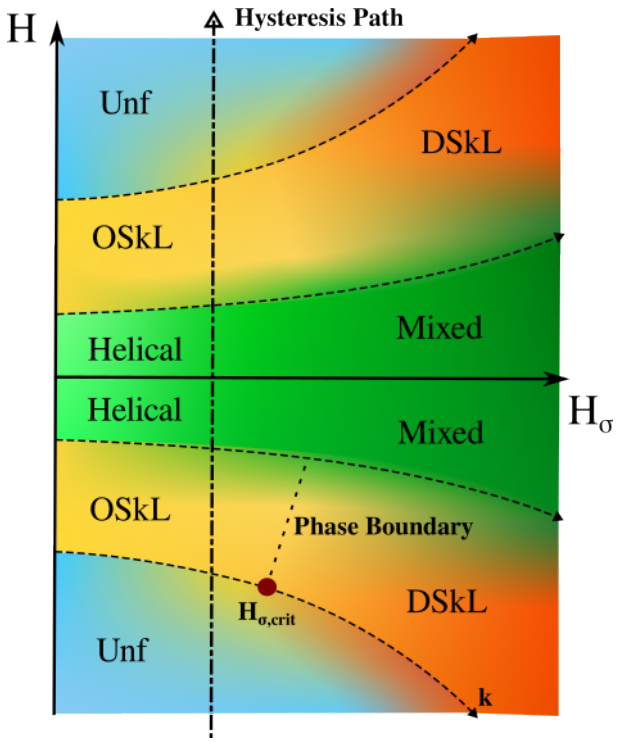

FIGURE 6.13: A hypothetical phase diagram of the spin system as a function of the external field, (H), and the level of defects, H_σ. At low levels of H_σ, we expect the system to pass through the uniform state, the OSkL, a helical state, and then back to a reversed OSkL, and the reversed uniform state. At high H_σ, the system smoothly transitions from the uniform state to a disordered array of skyrmions, which then becomes a mixed state of meandering patterns and skyrmions, before reverting back. k is the skyrmion nucleation boundary, and $H_{\sigma,crit}$ is the OSkL-DSkL transition point at this boundary.

6.6 Mapping the spin-spin correlation along the field sweep

The correlation between two spin configurations provides maximal information about any differences in the configurations. This information is summed across all lattice points to generate a single value characterising the degree of spin change. If we consider the spin texture as a line of spins (an $N \times 3$ vector),

$$S = \begin{bmatrix} s_{1x} & s_{1y} & s_{1z} \\ s_{2x} & s_{2y} & s_{2z} \\ \vdots & \vdots & \vdots \\ s_{Nx} & s_{Ny} & s_{Nz} \end{bmatrix}, \tag{6.1}$$

these textures are first normalised to eliminate dependency on lattice size,

$$\boldsymbol{s} = \frac{\boldsymbol{S}}{\sqrt{\sum_i^N \boldsymbol{S}_i^T \cdot \boldsymbol{S}_i}} \tag{6.2}$$

where $\boldsymbol{s}$ is the normalised line of spin vectors, and the spin-spin correlation is calculated as follows,

$$Q(1,2) = 1 - \sum_i^N (\boldsymbol{s}_{i,1} \cdot \boldsymbol{s}_{i,2}) \tag{6.3}$$

Here, $\boldsymbol{s}_{i,j}$ denotes the normalized spin vector of configuration j at lattice point i. When $\sum_i^N (\boldsymbol{s}_{i,1} \cdot \boldsymbol{s}_{i,2}) = 1$, indicating no spin direction change between configurations 1 and 2, we subtract this from 1, ensuring that $Q(1,2) = 0$ in this scenario. We choose this measure, since this detects all local changes between two spin configurations. Computing Q is equivalent to finding the squared change in length of the spin configuration vector[9],

$$Q = \frac{1}{2} \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|^2 \tag{6.4}$$

To map the nucleation boundary, we compute Equation 6.3 along the hysteresis loop between configurations at adjacent Zeeman fields, $Q(H, H + \delta H)$, as shown in Figure 6.14 for various H_σ (rows) and different realizations (columns). Initially, the hysteresis loop contains consecutive uniform states with no spin changes, resulting in $Q = 0$. The first instance when this value falls below 1 (within a tolerance of 1×10^3) marks the transition from the uniform to the skyrmion state at the nucleation boundary (highlighted by a yellow cross in Figure 6.14). In the 'clean' scenario, the SkL nucleates fully, resulting in a large initial signal in the spin-spin correlation behaviour (6.14(A-C)). Conversely, for the 'dirty' system, the appearance of only a single or a few skyrmions at the nucleation point leads to a much smaller signal (6.14(G-I)). The spin configuration at the nucleation point is plotted below the spin-spin correlation curve. For intermediate levels of defect, such as $H_\sigma = 0.02$, some realisations exhibit a full SkL at the nucleation point (6.14(D,E)), while others only display a single or a few skyrmions (6.14(F)). This behaviour occurs near the critical H_σ, which marks the OSkL-DSkL transition boundary.

We further note that the spin-spin correlation exhibits a varying number and size

[9] $\|\boldsymbol{s}_1 - \boldsymbol{s}_2\| = \sqrt{\|\boldsymbol{s}_1\| + \|\boldsymbol{s}_2\| - 2\boldsymbol{s}_1 \cdot \boldsymbol{s}_2} = \sqrt{2(1 - \boldsymbol{s}_1 \cdot \boldsymbol{s}_2)} = \sqrt{2Q}$

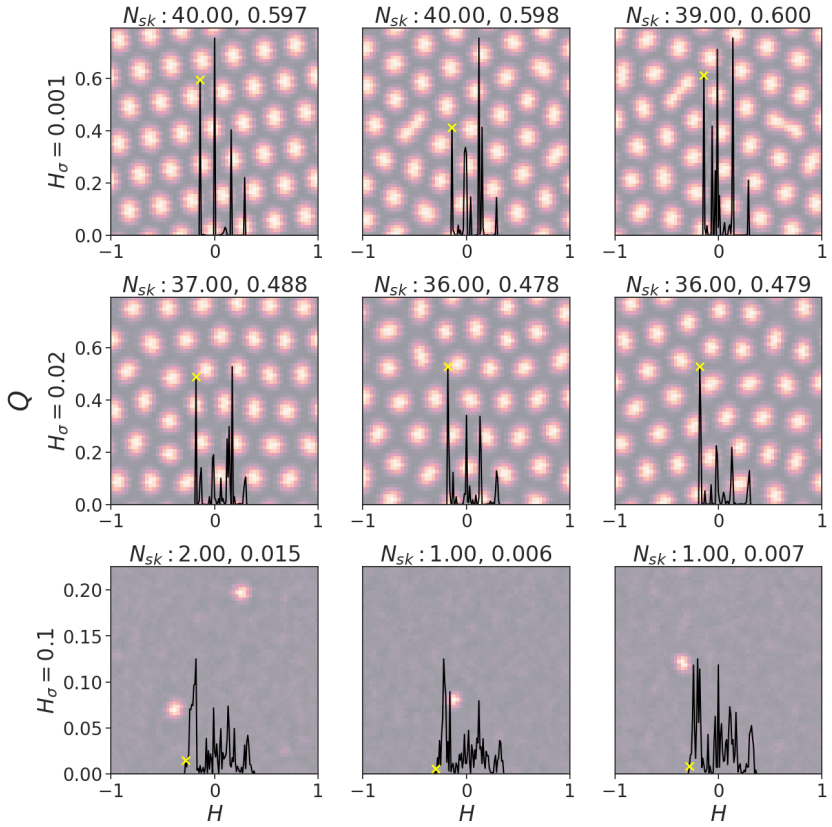

FIGURE 6.14: The spin-spin correlation, Q, between spin configurations at H and $H + \delta H$ along the hysteresis loop for one example in the 'clean' (top row) and 'dirty' (bottom row) cases. The yellow cross indicates the first point at which the Q is no longer 0 (i.e. there is a change in spin direction). In the 'clean' case, the full skyrmion lattice nucleates leading to a large deviation from 0, but in the 'dirty' case, only a single skyrmion appears which results in a much smaller change in Q.

of peaks as the level of defect varies. In the 'clean' case, the four major peaks identify the four transition boundaries between the five phases in the system. The smaller peaks correspond to certain structural changes such as the merging of skyrmions, or the movement of spin spiral branches. As the level of defects increase ($H_\sigma = 0.02$), the amplitude of the major peaks decrease, especially for intermediate skyrmion-helical transition, and the number of minor peaks begin to rise. The initial phase transition to the skyrmion state still exhibits a sharp signal, but, in certain realisations, defects smear the first-order transition, resulting in cases of a single skyrmion rather than the full lattice. Further increases in the defect level significantly diminishes the uniform-skyrmion transition signal and it is difficult to determine if any of the peaks represent a phase change. These large defects have rendered the system more 'fluid', with skyrmions/meandering structures continuously nucleating and

merging along the hysteresis loop, rarely maintaining a stable configuration (Q is rarely 0 here). The skyrmion nucleation boundary (yellow cross) therefore serves as the most consistent reference point for comparison across these multiple different defect levels.

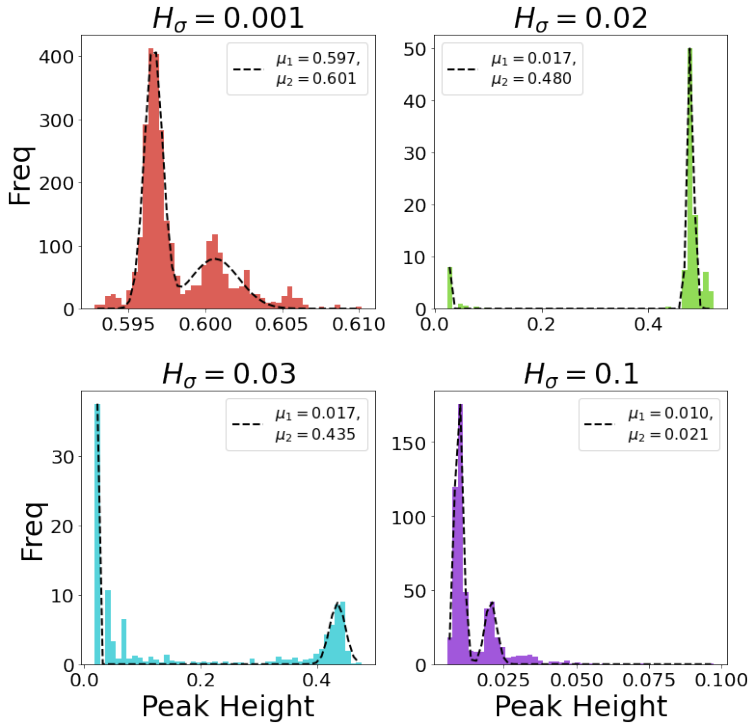

FIGURE 6.15: Histograms of the Q at the nucleation point at various defect levels, $H_\sigma \in [0.001, 0.1]$. The 'clean' and 'dirty' cases have an approximately unimodal distribution centered at 0.6 and 0.015, respectively. In intervening defect levels, the means of the distribution separate even further since skyrmions are likely to nucleate both as a full lattice or singly. The histograms have been fitted with a mixture of two Gaussians, given by the dashed black line.

To identify the OSkL-DSkL transition, we use value of Q at the skyrmion nucleation boundary, Q_{sk}, and present a histogram of these values in Figure 6.15 for all realisations and at various H_σ. This histogram is fitted to a mixture of two Gaussians, modelled as follows,

$$\rho(x, A, \mu, \sigma) = Ae^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{6.5}$$

$$\text{Model}(x) = \rho(x, A_1, \mu_1, \sigma_1) + \rho(x, A_2, \mu_2, \sigma_2) \tag{6.6}$$

Here, A_i, μ_i, and σ_i represent the amplitude, mean, and standard deviations of

distribution i. The histogram for the 'clean' case is almost a unimodal distribution, with the means of both Gaussians approximately coinciding at 0.6. For the two modes, configurations with $\mu_1 = 0.597$ correspond to a clean hexagonal lattice of 40 skyrmions, while $\mu_2 = 0.601$ also contains an ordered lattice, but with two pairs of skyrmions merged to form elongated structures. Similarly, the 'dirty' case also exhibits an almost unimodal distribution, with an average mean of 0.015. The two significant modes correspond to a 7 skyrmions ($\mu_1 = 0.0010$) and 2 skyrmions ($\mu_1 = 0.021$).

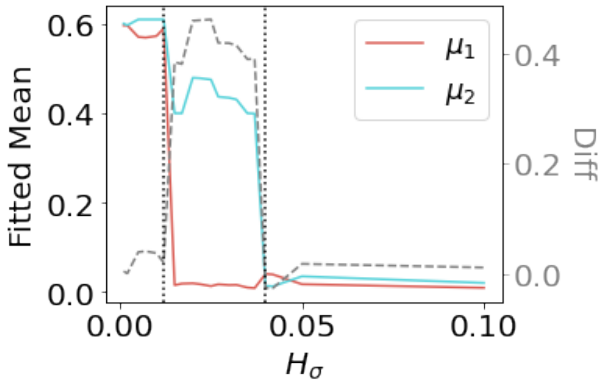

FIGURE 6.16: Evolution of the two means, μ_1 and μ_2, for the fitted mixture of Gaussian model, given as a function of the level of defects, H_σ The grey dashed line plots $\mu_2 - \mu_1$, which is maximal in the region $0.012 < H_\sigma < 0.4$ (dotted vertical lines).

At intermediate H_σ values, the modes of the distributions diverge, as some realisations initially nucleate the full SkL, while others generate only one or a few skyrmions. For lower defect levels ($H_\sigma = 0.02$), the system tends to relax to the full SkL, resulting in a larger peak centered at 0.48. As the level of defects increases ($H_\sigma = 0.03$), the height of this peak decreases, while the peak at 0.017 rises, indicating a higher likelihood of the system relaxing to a state with a few skyrmions. The evolution of μ_1 and μ_2 as a function of the defect level is illustrated in Figure 6.16. This graph demonstrates that the difference between μ_1 and μ_2 is most pronounced when $0.012 < H_\sigma < 0.4$, which encompasses the location of the OSkL-DSkL transition boundary.

To analyse the distribution splitting behaviour, we partition Q_{sk} into two ranges so that every realisation at all defect levels falls into one of these categories: $0 < \text{bucket}_1 \leq 0.35$ and $0.35 < \text{bucket}_2 \leq 0.7$. We then record the occurrences where

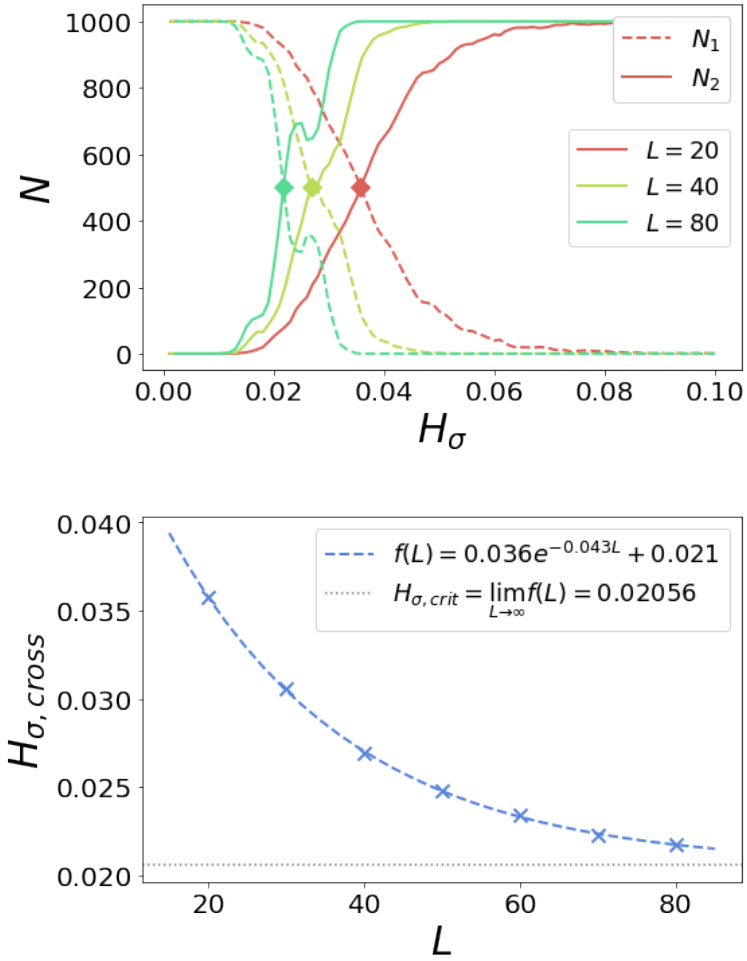

FIGURE 6.17: A) The evolution of the number of realisations where peak height fall in $0 <$ bucket$_1 \leq 0.35$ (N_1) and $0.35 <$ bucket$_2 \leq 0.7$ (N_2) as a function of the defect level, for various lattice sizes, $L \in [20, 80]$. The intersection of N_1 and N_2 occurs at $H_{\sigma,cross}$ which marks the critical defect level at L. B) $H_{\sigma,cross}$ as a function of L, which is fitted to an exponential decay, resulting in a critical defect level of $H_{\sigma,crit} = 0.02056$.

Q_{sk} falls into each range, denoted as N_1 and N_2, for each H_σ. The results are given in Figure 6.17A for various lattice sizes ranging from $L = 20$ to $L = 80$. Here, the intersection of N_1 and N_2 marks the critical H_σ level, signifying when the RP field is sufficiently strong to disrupt the OSkL. As the lattice size (L) increases, the transition in N_1 and N_2 becomes steeper, and the intersection occurs at a lower H_σ. The behaviour of the critical H_σ as a function of L follows an exponential decay (6.17B), where a least squares fit yields the following relationship,

$$f(L) = 0.036e^{-0.043L} + 0.021 \tag{6.7}$$

Extrapolating to $L = \infty$, the critical point is estimated as $H_{\sigma,crit} = \lim_{L\to\infty} f(L) = 0.021$.

The above analysis was conducted for a specific material with Hamiltonian parameters $J_{ex} = 0.5$ and $J_{dm} = 0.5$. To broaden our scope, we extend this investigation across a range of parameters, and since the critical H_σ likely varies with the skyrmion size (proportional to J_{ex}/J_{dm}), we focus our parameter variation solely on J_{ex}. Unfortunately, changing J_{ex} can potentially shift the boundaries of the skyrmion pocket, causing field sweeps at $T = 1$ to no longer intersect with the lower skyrmion pocket. Therefore, the first step is to identify these boundaries as a function of J_{ex}.

We begin by determining the simulation temperature as follows,

$$\text{Simulated } T(J_{ex}) = \frac{1 - T_{\text{l}}(0.5)}{T_{\text{u}}(0.5) - T_{\text{l}}(0.5)}(T_{\text{u}}(J_{ex}) - T_{\text{l}}(J_{ex})) + T_{\text{l}}(J_{ex}), \tag{6.8}$$

Here, $T_{\text{l/u}}(J_{ex})$ represents the lower and upper temperature boundaries of the lower skyrmion pocket at J_{ex}, occurring at the SkL-helical transition and the SkL-paramagnetic transition, respectively.

To identify the upper temperature boundary at J_{ex}, we generate 200 realizations of hysteresis loops at temperatures around $T_{\text{min, u}} < T_{\text{u}}(0.5) < T_{\text{max, u}}$ at $H_\sigma = 0.001$, spanning fields from $H = [-0.1, 0]$, where $T_{\text{max, u}}$ is selected such that the field sweep passes through the paramagnetic state, and $T_{\text{min, u}}$ is through the SkL. We iterate temperature from $T_{\text{max, u}}$ to $T_{\text{min, u}}$, and the temperature of the first occurrence where $N_{sk} > 0$ is recorded as the upper boundary for a particular realisation. Similarly, to determine the lower boundary, we simulate hysteresis loops around $T_{\text{min, l}} < T_{\text{l}}(0.5) < T_{\text{max, l}}$, where $T_{\text{min, l}}$ is in the helical state and $T_{\text{max, l}}$ is through the SkL, and record the first occurrence of $N_{sk} > 10$ when iterating through $[T_{\text{min, l}}, T_{\text{max, l}}]$. We use a threshold of $N_{sk} = 10$ because N_{sk} might not reach zero during the transition out of the skyrmion pocket, since the adjacent phase could be a meandering pattern, where the structure extremities could be counted as half-skyrmions. After identifying the boundaries per realisation, we average these to determine the true upper and lower temperature boundary for a specific J_{ex}.

The findings are given in Figure 6.18, showcasing the upper temperature boundary denoted by the red line and the lower temperature boundary represented by

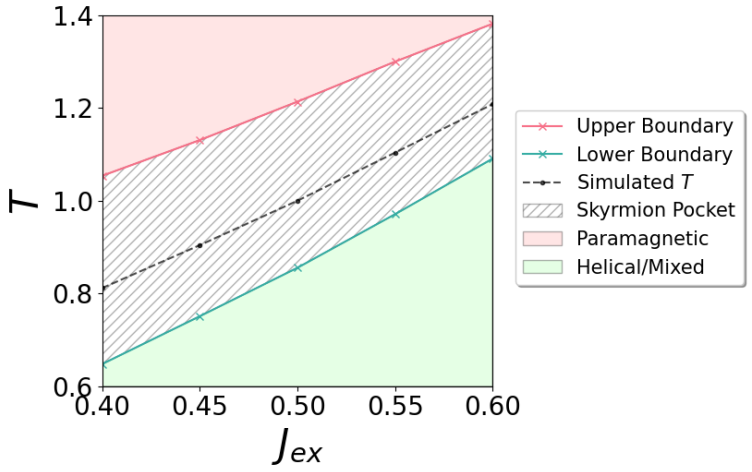

FIGURE 6.18: A rough approximation of the skyrmion pocket boundaries in T for various exchange strengths, $J_{ex} \in [0.4, 0.6]$. The hatched area is the skyrmion pocket region, above which is paramagnetic state at high temperatures, and the state below the pocket contains either helical structures or mixed states with isolated skyrmions. The black dashed line is the simulation temperature at which we will run our analysis for computing $H_{\sigma,crit}$.

the green line, both varying with respect to J_{ex}. The skyrmion pocket lies between these boundaries, with a paramagnetic state existing above $T_{\mathrm{u}}(J_{ex})$, and either a helical state or meandering state below $T_{\mathrm{l}}(J_{ex})$. The black dashed line traversing the skyrmion pocket corresponds to the simulation temperature calculated using Eqn. 6.8.

The analysis conducted in Figure 6.17 has been extended to various exchange strengths, $J_{ex} \in [0.4, 0.6]$, where hysteresis loops were generated at the simulated temperature (Figure 6.18). The resulting $H_{\sigma,crit}$ in Figure 6.19, exhibits a general trend of decreasing with increasing J_{ex}. This observation suggests that larger skyrmions are less robust to material defects, likely because the RP field tends to align spins with the z-axis, and therefore even a small positive defect point coupled with a strong J_{ex} can overcome the negative Zeeman field, leading to the formation of a skyrmion core. Conversely, a smaller J_{ex} necessitates a much larger RP defect point to achieve a similar effect. These observations suggest that thin film heterostructures with a large DMI (smaller skyrmions) may mitigate the impact of defects, although the increased perpendicular magnetic anisotropy, which aligns spins with the RP field, could present challenges. This highlights the possiblity of tuning material parameters to optimise for stability against defects.

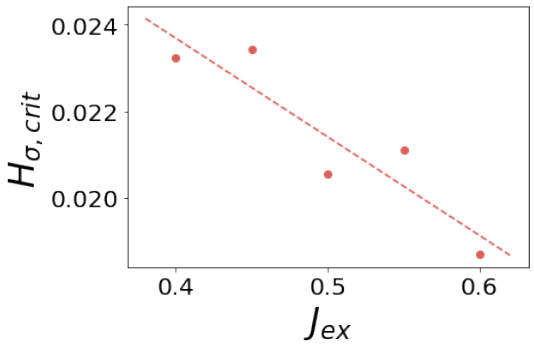

FIGURE 6.19: The critical H_σ as a function of the exchange strength, J_{ex}, and therefore as a function of skyrmion size which is propertional to J_{ex}/J_{dm}. For this analysis, the DMI strength is fixed at $J_{dm} = 0.5$.

6.7 Conclusion

Mean field analysis reveals the significant impact of defects on the formation of ordered skyrmion lattices, resulting in a gradual proliferation of skyrmions within the system and the emergence of a disordered array instead of an instantaneous first-order transition to a regular skyrmion lattice. Such configurations manifest as ring patterns in the small-angle neutron scattering output data, indistinguishable from the FFT of meandering domains or mixed skyrmion states. These differ significantly from the characteristic six peaks of a skyrmion lattice, and therefore may go unnoticed in experimental data. Our research underscores that these materials can still host skyrmions despite the presence of defects, though this may not be always evident from experimental results.

Additionally, we identify a phase transition from an OSkL to a DSkL as a function of the defect level, pinpointing the transition boundary, $H_{\sigma,crit}$, which signifies the minimal level of defect permissible to form an OSkL. This boundary decreases at higher J_{ex}, indicating that larger skyrmions are less resilient to defects, but also highlights that careful material tuning could mitigate the impact of these defects. Subsequent chapters use machine learning techniques to precisely identify these defect locations as well as the phase transition boundaries.

Chapter 7

Unsupervised Classification of Phases in Materials with Defects

7.1 Motivation

The macroscopic properties of Heisenberg systems with DMI, such as skyrmion number, average magnetisation and the reciprocal space peaks, were insufficient to uniquely determine all phases in 'clean' systems. The presence of defects further exacerbated these difficulties, as it blurred all first-order transitions, ultimately obscuring the helical phase consisting of meandering domains. To address these difficulties, we turn to unsupervised dimensionality reduction techniques to explore the space of all transformations of the spin vectors, in order to identify parameters that can classify all phases, particularly the OSkL and DSkL states.

7.2 Method

Our study uses two methods, **principal component analysis** (PCA) and **variational autoencoders** (VAE). PCA is ideal for situations where a linear transformation of the input components is sufficient to cluster the data. In contrast, VAE provides greater flexibility by allowing nonlinear transformations of the underlying features, and we can further enhance this method through the use of CNNs, which exploits the spatial information in the input to reduce the number of model parameters. VAEs become

essential for classifying phases in complex magnetic materials where linear transformations are insufficient. For instance, skyrmion number is a nonlinear transformation that can distinguish topological phases, and VAE may generate similarly complex parameters. Detailed explanations of both methods are given in Chapter 3.

7.3 Classifying Ising Phases

7.3.1 Dataset

We begin by evaluating the effectiveness of these dimensionality reduction techniques on classifying the phases in the well-understood Ising model. Our dataset includes magnetic lattices of varying sizes ($L \in [8, 16, 32, 64]$), where L is the length of the square lattice. The lattice contains binary spin variables ($s_i \in -1, 1$), which is relaxed through a Monte Carlo simulation consisting of 10000 equilibrium steps and 100000 Monte Carlo steps (MCS). Thermal statistics are collected every 1000 MCS, resulting in 100 snapshots of the final spin configuration, and the dataset includes 46 temperature points within the range of $T \in [0.1, 5]$ at $H = 0$, with 100 realisations at each temperature. Representative magnetic configurations collected at different temperatures are depicted in Figure 2.1a, and thermal properties are shown in Figure 2.1b. The dataset is divided into a training set and test set of 24,000 and 6,000 example configurations in each.

7.3.2 Classifying Ising phases using PCA

The following section analyses the effectiveness of PCA on classifying the phases in the Ising model. This model takes as input a matrix $\boldsymbol{X}$ representing the MC spin textures taken at all temperatures, where each row corresponds to a flattened vector of the configuration (with dimensions $L \times L$). The model is fitted using only the training data, and subsequently, both the training and test datasets are projected onto the linear subspace identified by the model. Notably, the test dataset is only transformed by the PCA components extracted from the training set to prevent any potential data leakage into the model.

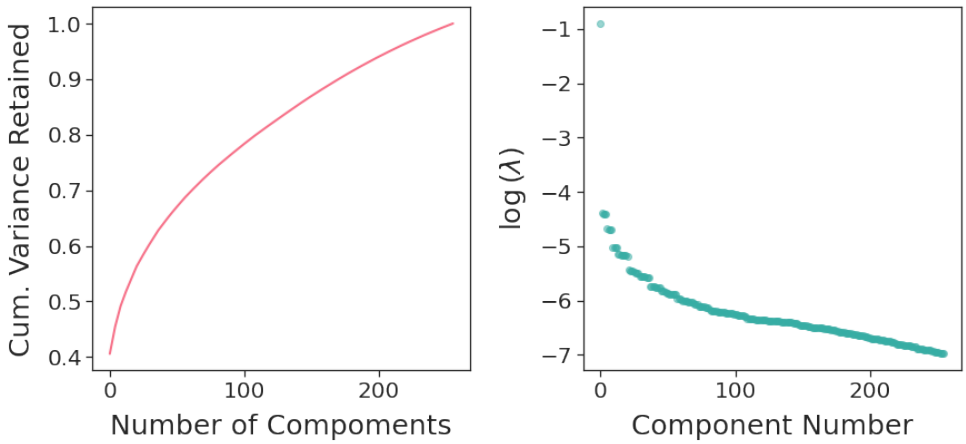

FIGURE 7.1: a) The cumulative sum of variance explained by each additional component of the PCA model, and the b) log eigenvalue of each component of the model for Ising lattice size $L = 64$.

We selected the maximum number of components to fit the PCA model to comprehensively assess the percentage of variance explained by each additional principal component (Figure 7.1). As anticipated, the retained variance sharply increases with the initial few components before levelling off. The eigenvalue of the first principal axis significantly surpasses all others, indicating that the key phase information is captured by the first component.

Figure 7.2A illustrates the test data at $L = 64$, projected onto the subspace spanned by the first two principal components, (p_0, p_1). This plot reveals three distinct regions: a central red cluster representing disordered Ising lattices at high temperatures (paramagnetic), flanked by two blue clusters indicating ordered Ising lattices at low temperatures (ferromagnetic). Figure 7.2B displays the same latent space, but now the color corresponds to the average magnetisation, thereby further distinguishing the spin-up (red) from the spin-down (blue) ordered states. Configurations captured at $T = 2.2$[1], depicted in dark grey, form two clusters positioned between the red (high T) and blue (low T) regions in Figure 7.2B, suggesting that one state contains more spin-up clusters while the other contains more spin-down clusters. A histogram of the transformed dataset along the first principal axis, p_0, displayed above the latent space figure, identifies the three main spin states identified by the ML algorithm. Notably, only the first principal component, characterised by the largest eigenvalue, effectively distinguishes between these different configurations.

[1] This is the temperature closest to $T_c = 2.2691$ in our dataset.

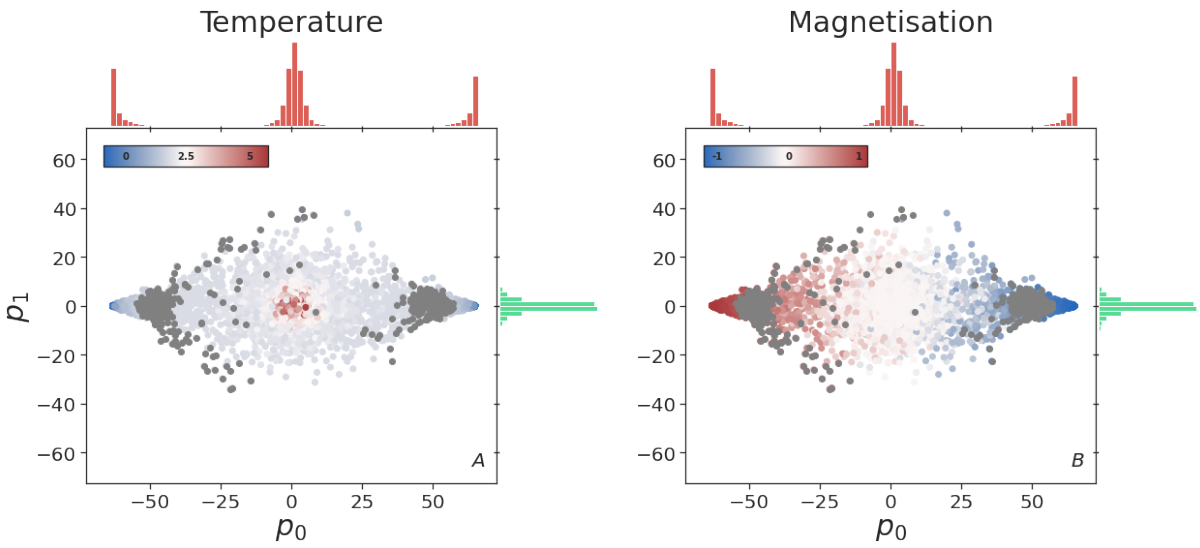

FIGURE 7.2: The test datasets transformed into the subspace spanned first the first two principal components, where the colors represent the A) temperature, and B) average magnetisation. Histograms are shown along the first, p_0, and second, p_1, principal axes.

We examine the weight matrices of the first principal component (p_0) to understand the features learned by PCA for clustering the different phases within its latent space. The weight matrix of the first principal component contains values which are uniformly constant at approximately $w_0 = \frac{1}{N}[1, \ldots, 1]$, thereby implying that each spin variable contributes equally in the transformation, $p_0 = w_0 \cdot X_i$. The average of the principal component, p_k, for each temperature is computed as,

$$\langle p_k \rangle = \frac{1}{n} \sum_{i}^{n} |w_k \cdot X_i|. \tag{7.1}$$

This characteristic resembles the calculation of the thermally-averaged order parameter, $m = \frac{1}{n} \sum_i^n \left(\frac{1}{N} \sum_j^N s_j \right)$, implying that PCA has captured this parameter. Indeed, $\langle p_0 \rangle$ exhibits a comparable power-law behavior to m (see Figure 7.3a) with perfect correlation between the two, as shown in Figure 7.3c.

The second principal component, $\langle p_1 \rangle$, mirrors the behaviour of susceptibility, χ, with a peak at the critical boundary, T_c. By applying a finite size scaling argument, where T_c scales inversely with L, we determine a critical temperature of $T_c = 2.139 \pm 0.042$ through a least-squares fit (Figure 7.3d). This result closely aligns with the true value of 2.2691, underscoring PCA's effectiveness in delineating the phases.

Consequently, PCA, an unsupervised machine learning method, successfully

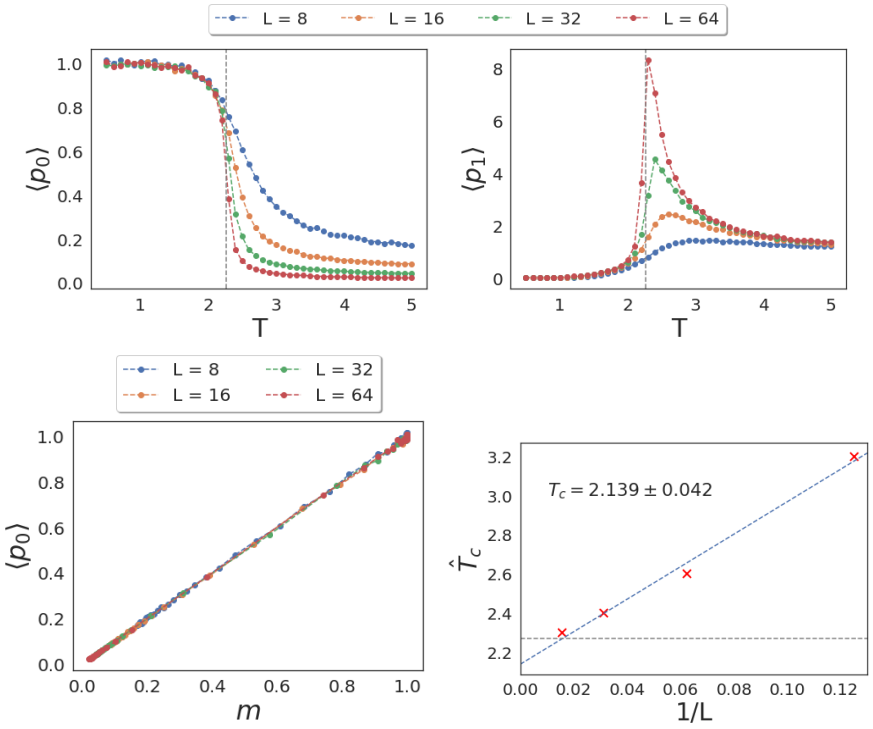

FIGURE 7.3: The averaged principal components at each temperature of the a) first and b) second principal components. c) The average of the first principal component, $\langle p_o \rangle$, as a function of the thermally-averaged order parameter, m, at different length scales. There is a perfect correspondence between the two. d) The temperature at the peak of $\langle p_1 \rangle$ ($\hat{T}_c$) is plotted as a function of the inverse lattice size, $\frac{1}{L}$, and a least squares fitting predicts a critical temperature of $T_c = 2.139 \pm 0.042$.

condenses the spin texture information into a single value, highly correlated with the order parameter, which enables phase identification without additional knowledge of the system's Hamiltonian. However, PCA's limitation to linear transformations excludes non-linear features, potentially restricting its ability to summarise phase information in complex Heisenberg systems. This presumption arises from our observation in Chapter 6 where N_{sk} serves as the most reliable indicator of a phase transition along a field sweep, representing a non-linear transformation of the spin configuration. Therefore, in subsequent sections, we use VAEs, first testing the methodology on the Ising model, and then applying it on the more complex Heisenberg systems, covered in Chapter 6.

7.3.3 Classifying Ising phases using VAE

In this section, we test the VAE method for its effectiveness in classifying the Ising phases[2]. The encoder and decoder of the VAE uses convolutional layers to extract spatial information from the spin textures while minimising computational overhead. It comprises three convolutional layers in the encoder and filter sizes of $\{32, 64, 64\}$, all with a stride of 2 and a `relu` activation function, which is condensed to 16 nodes in a `Dense` layer. The latent space dimension was set to 2, thus the `Dense` layer produces two means ($\mu_{0/1}$) and standard deviations ($\sigma_{0/1}$). Finally, a latent variable, z_i, is sampled and input into the decoder, which reverses the encoder layers, producing a reconstruction of the original spin configuration. The full architecture is shown in Figure 7.4.

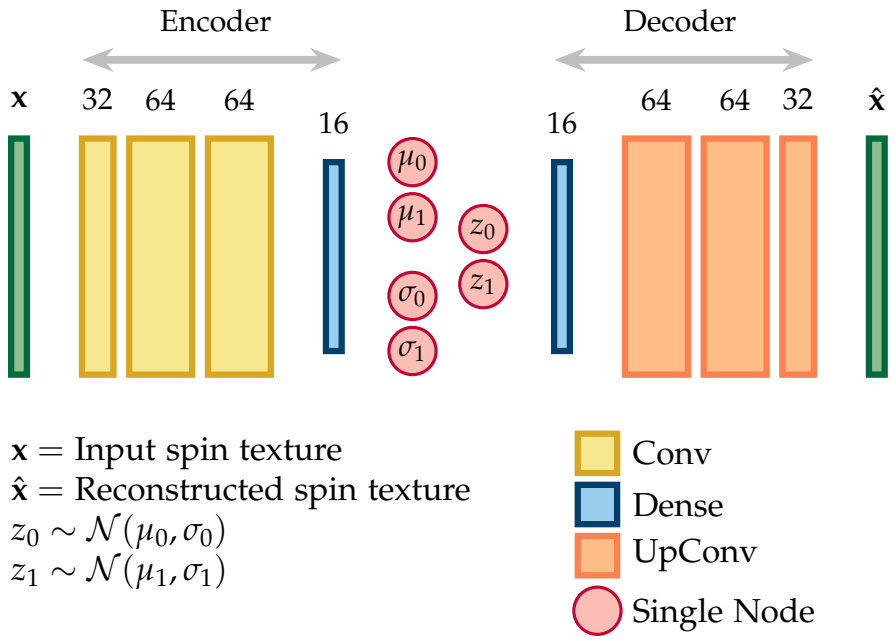

FIGURE 7.4: VAE architecture used to classify phases in the Ising model. The Ising spin configuration is compressed to two dimensions (z_0, z_1) which contain the maximum information about the state.

The resulting latent space of the VAE for the test dataset is illustrated in Figure 7.5, in terms of the means (A,C) and standard deviations (B,D) of z. In the top plots, colour denotes temperature (A, B), while the bottom plots illustrate the average magnetisation of the spin texture (C, D). Similar to PCA, spin textures of the

[2]See Chapter 3 for more information about the variational autoencoder.

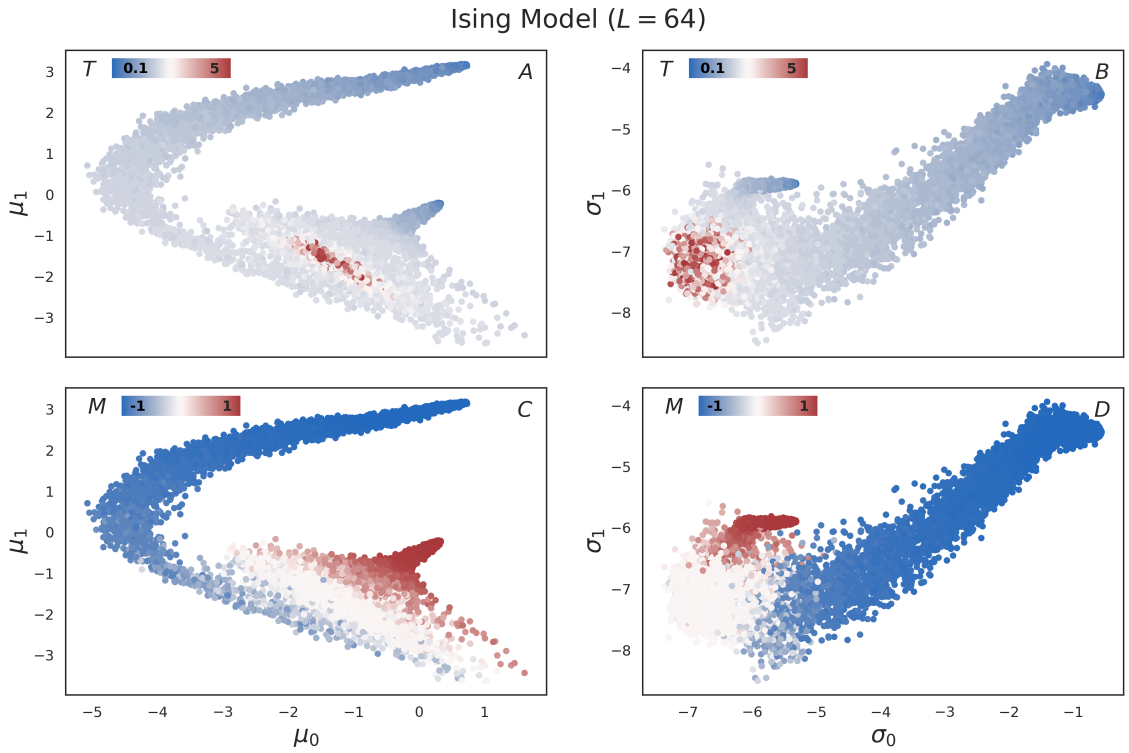

FIGURE 7.5: **Latent space of the Ising model ($L = 64$)**: The test datasets are transformed into the means (A,C) and standard deviations (B,D) of the latent variables, with the colors representing the temperature (A,B) and average magnetisation (C,D).

paramagnetic, and the two ferromagnetic states with spin-up and spin-down configurations, all occupy different regions of the latent space, with the paramagnetic state ($M = 0$) forming a wide elliptical area in the mean space and the two ferromagnetic states extending out as branches from this state. Unlike PCA, however, which orders its principal components such that most of the phase information is condensed in the first component, both latent variables are required to pinpoint a phase. Given only the μ_0 of a configuration, it would be difficult to identify whether it is in the paramagnetic, spin-down or even spin-up phase depending on the value of μ_0. Conversely, these phases can be differentiated exactly using only p_0 from PCA, with a linear relationship between p_0 and the system magnetisation. Nevertheless, from inspection, we see that the VAE is successful in differentiating the Ising phases, with these phases identified via μ_0 and μ_1. If we plot these variables individually, as we do below, we will see a clear discontinuity occuring at the critical temperature.

The temperature-dependent behaviour of the means and standard deviations,

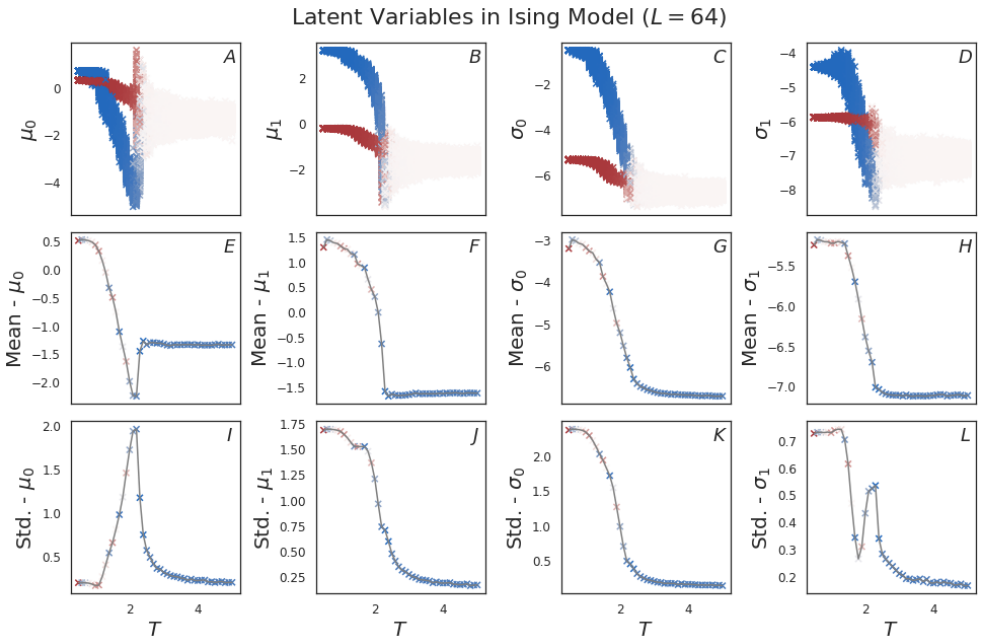

FIGURE 7.6: **Latent variables of Ising model ($L = 64$)**: The individual means (A,B) and standard deviations (C,D) of the encoder are shown as a function of temperature. The second and third row shows the means and standard deviations of $\mu_{0/1}$ and $\sigma_{0/1}$.

given in Figure 7.6, shows a clear distinction between the phases from their distribution. At high temperatures above T_c, the paramagnetic phase occupies a wide range of values in all four latent variables with a unimodal distribution. For illustrative purposes, we depict distributions of only μ_1 in Figure 7.7. Below T_c, this latent variables split into two branches representing either the spin-up or spin-down ferromagnetic state, exhibiting a bimodal distribution, with each branch showing a smaller spread. As seen in Chapter 6, this splitting of the distribution is a characteristic signal of a phase transition.

By taking the means and standard deviations of the four latent variables (7.6(E-L)), we can pinpoint the exact temperature boundary of the two phases. Many of these measures exhibit similarities either to the behaviour of the order parameter, m (F-H, J-K), or the magnetic susceptibility, χ (I). However, due to the stochastic nature of the training process and the non-convexity of the error landscape, these latent variables are not perfectly correlated to m or χ, as observed with the PCA variable p_0 (Figure 7.3c). Additionally, different runs of the training process may identify different behaviours for the latent variables, reducing the comparability of these parameters across different lattice sizes, L. Regardless, using the standard

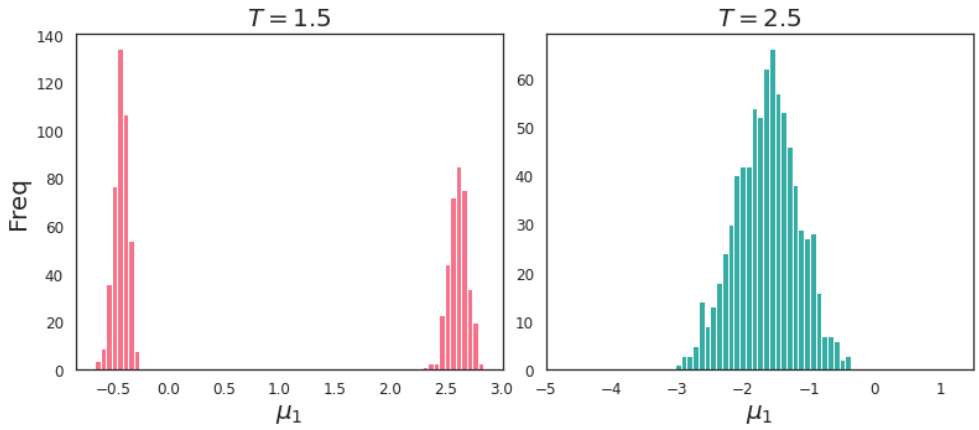

FIGURE 7.7: Distribution of μ_1 with A) a bimodal distribution at temperatures below T_c ($T = 1.5$) and B) a unimodal distribution at temperatures above T_c ($T = 2.5$). C) There is a positive correlation between the mean of μ_1 and the order parameter, m.

deviation of μ_0 (7.6I), we identify the Curie temperature as $T_c = 2.2$ at $L = 64$, which closely approximates the true value of 2.2691.

Despite the increased complexity inherent in VAEs from their non-linearity and stochastic learning process, the model demonstrates its capability of accurately differentiating the Ising phases and identifying the critical transition boundary. In the next section, we use the VAE to classify phases in the more complex magnetic systems covered in Chapter 6.

7.4 'Clean' and 'Dirty' Heisenberg System

In this section, we test the performance of the VAE in clustering the different phases present in both 'clean' and 'dirty' Heisenberg systems. Our previous analysis in Chapter 6 highlighted a clear first-order phase transitions among all states in 'clean' systems, easily identifiable due to the N_{sk} behaviour along the field sweep. However, distinguishing between the phases in 'dirty' systems proved more challenging, as the skyrmions continuously merged into meandering domains, forming a mixed state with a smoothly varying N_{sk}. Our objective is to assess whether the VAE can provide better differentiation of phases in 'dirty' systems.

7.4.1 Model and Dataset

The VAE encoder architecture comprises five zero-padded `Conv2D` layers with filters in $[32, 64, 128, 128, 128]$, employing the `selu` activation function[3]. Once again, we generate 2 values for the means and standard deviations of the latent variables, and once z is sampled from these distributions, the decoder reverses the encoder architecture to generate a reconstruction of the spin texture.

We use the hysteresis data outlined in Chapter 6 at $T = 1.0$ and $H_\sigma = 0.001$. This dataset is divided into training and test sets with a ratio of 80% to 20%, respectively. For simplicity, only the z-component of the spin configuration is considered, where $s_{i,z}$ is set to 0 if less than 0, and to 1 otherwise, at each lattice point i. The reconstruction loss is calculated as the binary cross-entropy (Equation 3.6) between the reconstructed spin texture and the original at each lattice point. The network is trained using the `Adam` optimizer for 4000 epochs, with an `EarlyStopping` callback used to halt training if the validation loss change drops below 0.001. Additionally, a `ModelCheckpoint` callback is implemented to save the weights of the model with the lowest validation loss.

7.4.2 Results

Classifying phases in 'Clean' Systems - $H_\sigma = 0.001$

Figures 7.8(A,C) and (B,D) illustrate the resulting latent space, represented by the means and standard deviations respectively, where the color denotes the external field in the top plots (A, B) and it signifies the skyrmion number in the bottom plots (C, D). N_{sk} partitions the latent space, delineating the five observed phases in the field sweep, particularly highlighting the uniform states, which form isolated clusters separate from the others.

Interestingly, despite the VAE lacking prior knowledge of the field sweep protocol and the dataset being randomised, it effectively encodes the phase progression

[3]The `selu` (Scaled Exponential Linear Unit) activation function can be mathematically defined as follows,

$$\text{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha(\exp(x) - 1) & \text{if } x \leq 0 \end{cases} \tag{7.2}$$

where $\alpha \approx 1.6733$ and $\lambda \approx 1.0507$.

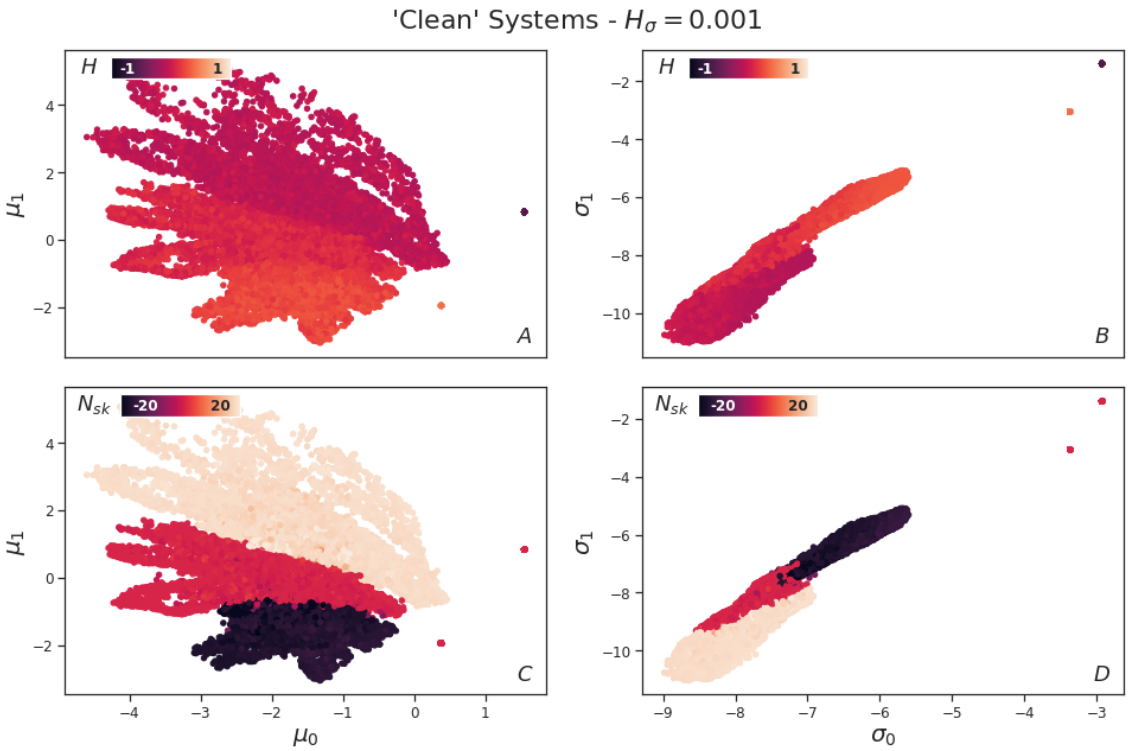

FIGURE 7.8: **Latent space of 'clean' systems**: The test datasets are transformed into the means (A,C) and standard deviations (B,D) of the latent variables, with the colors representing the external field (A,B), and skyrmion number, (C,D).

in the latent space partitioning: spin textures generated at negative fields predominantly occupy the positive μ_1 region, and vice versa. Moreover, the sequential arrangement of SkL(+), the helical state, and SkL(-) mirrors the field sweep procedure precisely. This is likely due to the VAE keeping similar clusters closer together and it learns that the transition between SkL(+) and SkL(-) is facilitated through the helical phase. Such learning might stem from the presence of mixed phases[4], where the system contains both skyrmions and spin spirals, rather than an instantaneous switch from SkL(+) to the helical phase. Without these mixed phases, we speculate that the clusters would be completely disentangled.

Additionally, we note that the SkL(+) occupies a larger region of the latent space compared to the helical/SkL(-) state, implying a greater diversity in the potential manifestations of SkL(+) from the uniform state, for example different lattice orientations. Conversely, the diminished size of the helical and SkL(-) regions indicates fewer pathways for these states to emerge from the SkL(+).

[4]This is due to non-zero defect level used to generate 'clean' systems

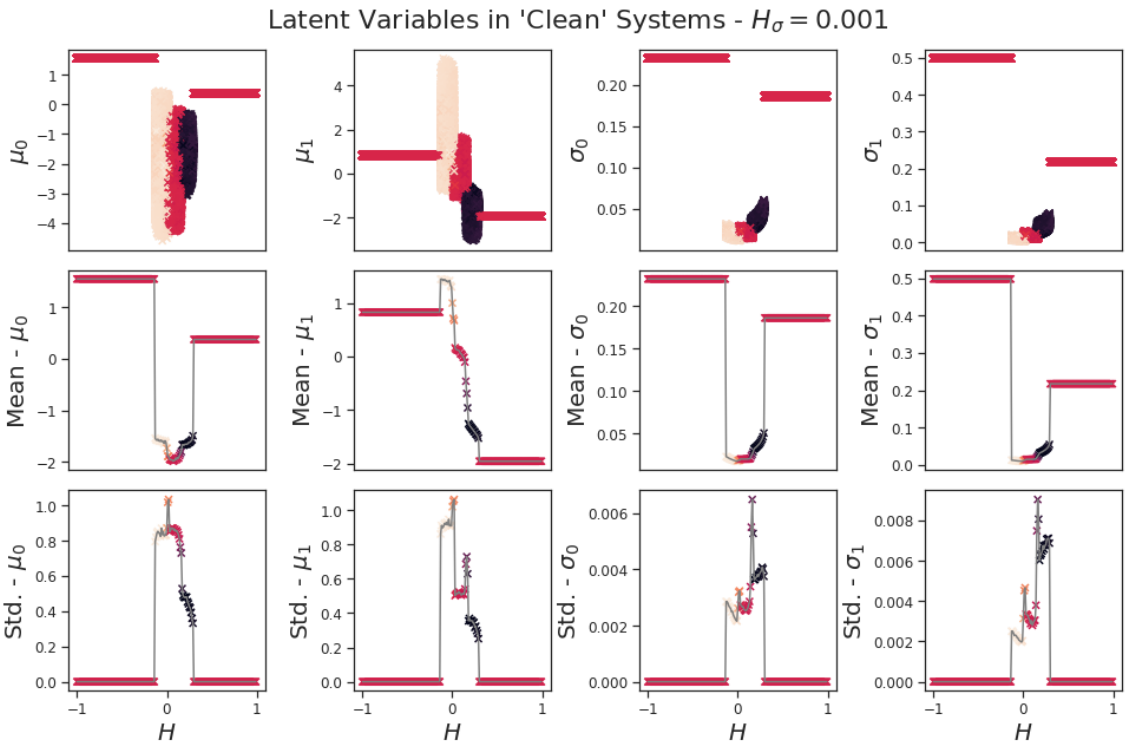

FIGURE 7.9: **Latent variables of 'clean' systems**: The test datasets are transformed into the means (A,C) and standard deviations (B,D) of the latent variables, with the colors representing the external field (A,B), and skyrmion number, (C,D).

The field-dependent behavior of the latent variables (LVs), shown in Figure 7.9 together with their means and standard deviatons, reflects the partitioning observed in the latent space. Specifically, the LVs remain relatively stable for the uniform states (both in $\pm z$), while exhibiting varying but consistent distributions for the skyrmion and helical states as a function of H. The abrupt shifts in LV distribution indicate first-order phase transitions. These LVs are more effective in distinguishing hysteresis phases compared to our macroscopic properties; while N_{sk} failed to differentiate between the two uniform and helical states, and the spin-spin correlation (Chapter 6) proved overly sensitive to local changes such as skyrmion merging, the LVs uniquely distinguishes each of the five phases. Additionally, the local peaks in standard deviation for σ_1 can effectively pinpoint the transition fields at each boundary (Figure 7.10A), aligning closely with the fields identified by N_{sk}, as shown in Table 7.1.

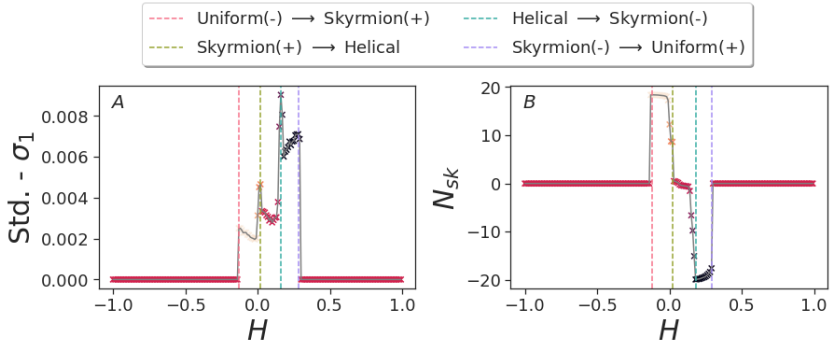

FIGURE 7.10: A) Standard deviation of the VAE latent variable, σ_1, and B) N_{sk} are a function of Zeeman field, H, for the 'clean' system. The vertical dashed lines correspond to the transition fields at the boundary of various phases.

Transition Fields in 'Clean' Systems

Transition	Std. - σ_1	N_{sk}
Uniform(-) → Skyrmion(+)	-0.13	-0.12
Skyrmion(+) → Helical	0.02	0.02
Helical → Skyrmion(-)	0.16	0.18
Skyrmion(-) → Uniform(+)	0.28	0.29

TABLE 7.1: Comparison of transition fields in 'clean' systems between the VAE Std. - σ_1, and the manual computation of N_{sk}.

Classifying phases in 'Dirty' Systems - $H_\sigma = 0.1$

We performed a similar analysis for 'dirty' systems, with the latent space illustrated in Figure 7.11. In this case, only the Unf(-) is distinctly separate from the others, while the cluster representing Unf(+) overlaps with other phase clusters, though still converging to a single point in the latent space. For the intermediate states, the latent datapoints form a continuous surface that, similar to the 'clean' case, reflects the hysteresis protocol through its partitioning. Once again, μ_1 appears to exhibit some correlation with the Zeeman field.

One notable observation is that the VAE better distinguishes textures at negative fields than positive ones, with a broader spread of the contiguous surface at negative fields, and a more gradual change in N_{sk} (Figure 7.11(C, D)) across the cluster. Conversely, the surface at positive fields is much smaller in both μ and σ. The challenge of differentiating positive spin textures is particularly evident in the N_{sk} surface for σ: from Unf(-) cluster, it exhibits a smooth, continuous branch up to the point with

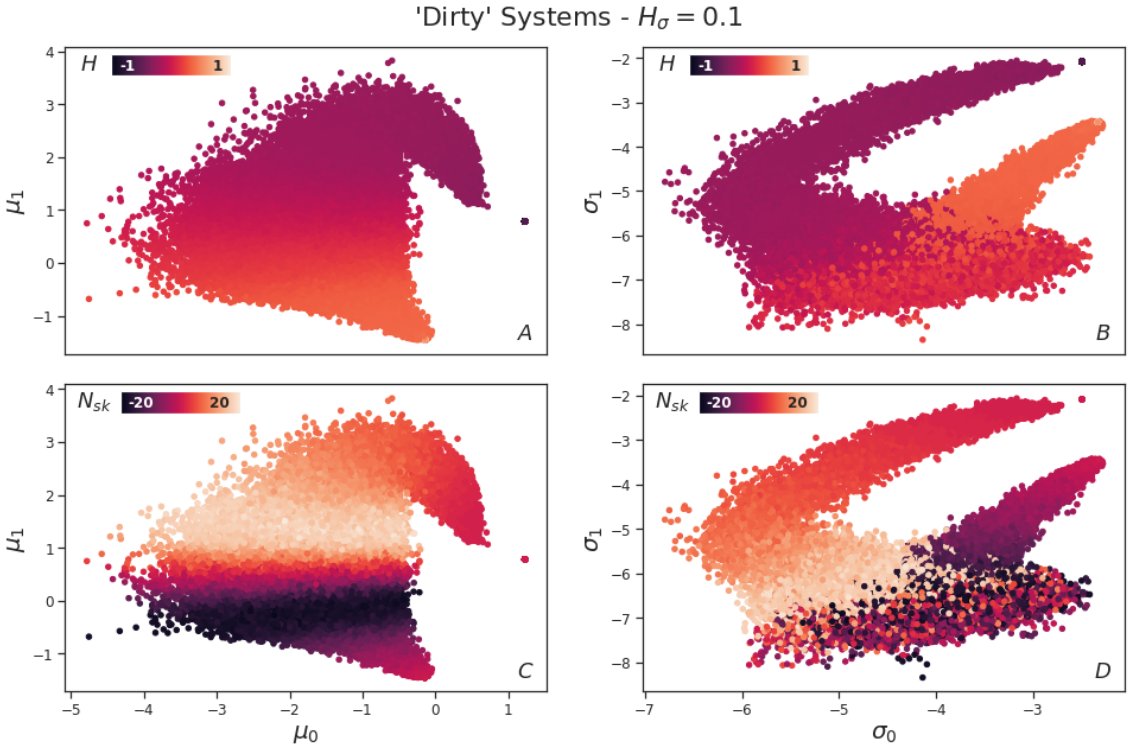

FIGURE 7.11: **Latent space of 'dirty' systems**: The test datasets are transformed into the means (A,C) and standard deviations (B,D) of the latent variables, with the colors representing the external field (A,B), and skyrmion number, (C,D).

the maximum number of skyrmions in $+z$, Sk(+), followed by a wide cluster containing textures with varying skyrmion numbers, and then a much narrower branch with fewer skyrmions. Interestingly, N_{sk} does not appear disordered in μ space, indicating that μ may be learning the skyrmion number (or the external field), while σ seems to distinguish the mixed phase consisting of Sk(+), meandering domains, and Sk(-) as a distinct entity. The broadness of this phase suggests a variety of configurations that the mixed state can adopt, and conversely, the shrinking of this region to a narrower branch at even higher fields, albeit with a non-zero N_{sk}, suggests there are fewer configurations of isolated skyrmions compared to mixed states.

Upon analyzing the distributions of the μ and σ components with respect to H (Figure 7.12), the previously observed discontinuous changes in distributions with N_{sk} are no longer apparent. Specifically, in the standard deviation plots of σ, only two turning points are identifiable, corresponding to phase transitions into and out of the uniform state. This observation aligns with our earlier hypothesis in

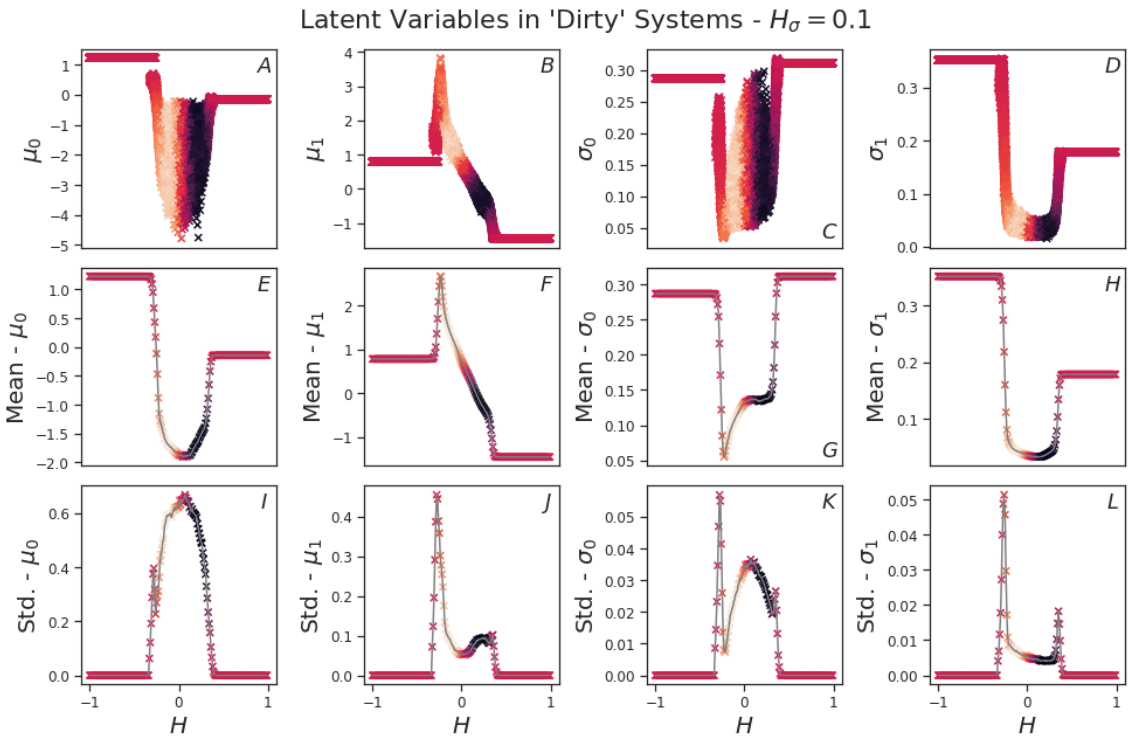

FIGURE 7.12: **Latent variables of 'dirty' systems**: The test datasets are transformed into the means (A,C) and standard deviations (B,D) of the latent variables, with the colors representing the external field (A,B), and skyrmion number, (C,D).

Chapter 6 about the smearing of the first-order transitions between the intermediate phases[254–256]. Similar to the 'clean' case, we can pinpoint the transition field boundaries around the uniform state and compare them against the behavior of N_{sk} (Figure 7.13 and Table 7.2). The high level of agreement between the two suggests that the VAE effectively clustered these phases with an accuracy that was comparable to our manual methods, which were based on our understanding of the system.

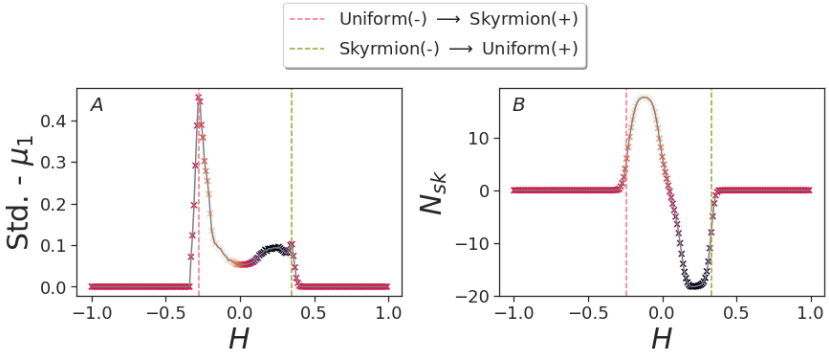

FIGURE 7.13: A) Standard deviation of the VAE latent variable, σ_1, and B) N_{sk} are a function of Zeeman field, H, for the 'dirty' system. The vertical dashed lines correspond to the transition fields at the boundary of various phases.

Transition Fields in 'Dirty' Systems

Transition	Std. - σ_1	N_{sk}
Uniform(-) → Skyrmion(+)	-0.26	-0.24
Skyrmion(-) → Uniform(+)	0.35	0.33

TABLE 7.2: Comparison of transition fields in 'dirty' systems between the VAE Std. - σ_1, and the manual computation of N_{sk}.

7.5 Classifying Ordered vs. Disordered SkL

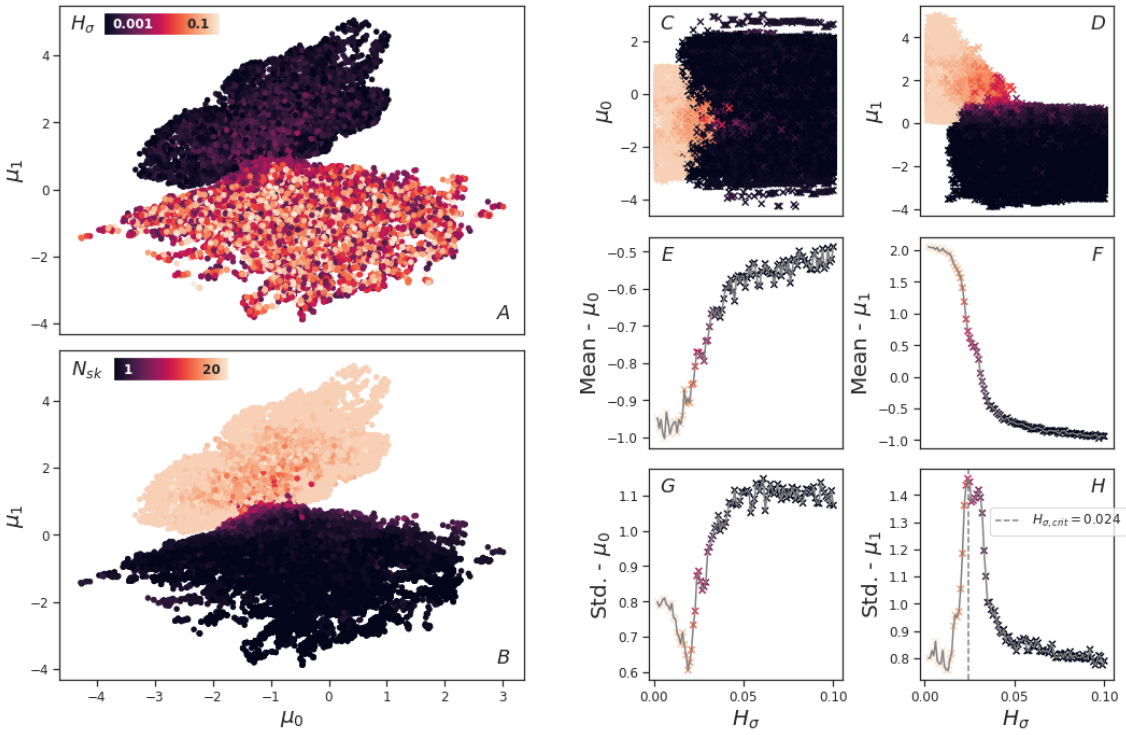

FIGURE 7.14: **Latent space of systems along skyrmion nucleation boundary**: The test datasets are transformed into the means (A,B) of the latent variables, with the colors representing the defect level (A), and skyrmion number (B). The individual latent variables are plotted as function of defect level for μ_0 (C), μ_1, along with their respective means, (E-F) and standard deviations, (G-H). The color of these diagrams represent N_{sk}.

If the VAE is relatively accurate at classifying phases within both 'clean' and 'dirty' systems, it should be proficient in distinguishing spin textures along the skyrmion nucleation boundary as a function of defect level. A comprehensive analysis of these spin textures can be found in Chapter 6. To summarise, the initial switching configuration either comprises a full regular skyrmion lattice or a few individual

skyrmions, depending on the defect level. Our dataset contains all these spin configurations, with 1000 realizations at each defect level in $H_\sigma \in [0.001, 0.1]$, which was once again shuffled and divided into training and test sets with a ratio of 80% to 20%. The resulting latent space for μ is illustrated in Figure 7.14(A, B).

Here, we observe two distinct clusters: one exhibiting a high number of skyrmions and the other featuring only a few, with a slight overlap between them. Configurations with $N_{sk} \geq 20$, predominantly originating from very low defect levels, are positioned along the outer periphery of the upper cluster. In contrast, the inner data points contain fewer skyrmions, likely representing a partially nucleated skyrmion lattice, a phenomenon commonly observed around relatively low defect levels, $H_\sigma \approx 0.01$. Notably, the bottom cluster of H_σ in (A) reveals that while certain realizations at very low defect levels yield a few skyrmions at the nucleation boundary, the top cluster shows that there are no instances of generating a complete array of skyrmions from a high defect level. This supports our previous observations, that higher defect levels always leads to a more gradual proliferation of skyrmions, whereas certain instances at low defect levels can yield a few skyrmions at the switching point, potentially followed by a transition to a fully ordered lattice at subsequent field points. The data points at the intersection of the two clusters occur at approximately the critical defect level, which can be determined from the defect-dependent behaviour of the latent variables. Specifically, we observe an increase in the spread of μ_1 (H), with a peak corresponding to $H_{\sigma,crit} = 0.024$.

In Chapter 6, we established that the OSkL-DSkL phase transition can be identified solely through the spin-spin correlation between the configuration at the nucleation boundary and a uniform state. Since this is a linear transformation, PCA should theoretically suffice in differentiating these phases. Unfortunately, due to the large size of the training dataset (1.2 TB), using the PCA model from `scikit-learn` is not feasible, as it requires loading the full dataset into memory. Instead, we repurpose the VAE model, removing any sources of non-linearity in the function, and the sampling layer such that the model directly outputs z. The loss function now only contains the reconstruction error between the true and predicted configurations. This autoencoder performs a dimensionality reduction that is similar to PCA, but does not enforce the constraint that the latent variables should be ordered by

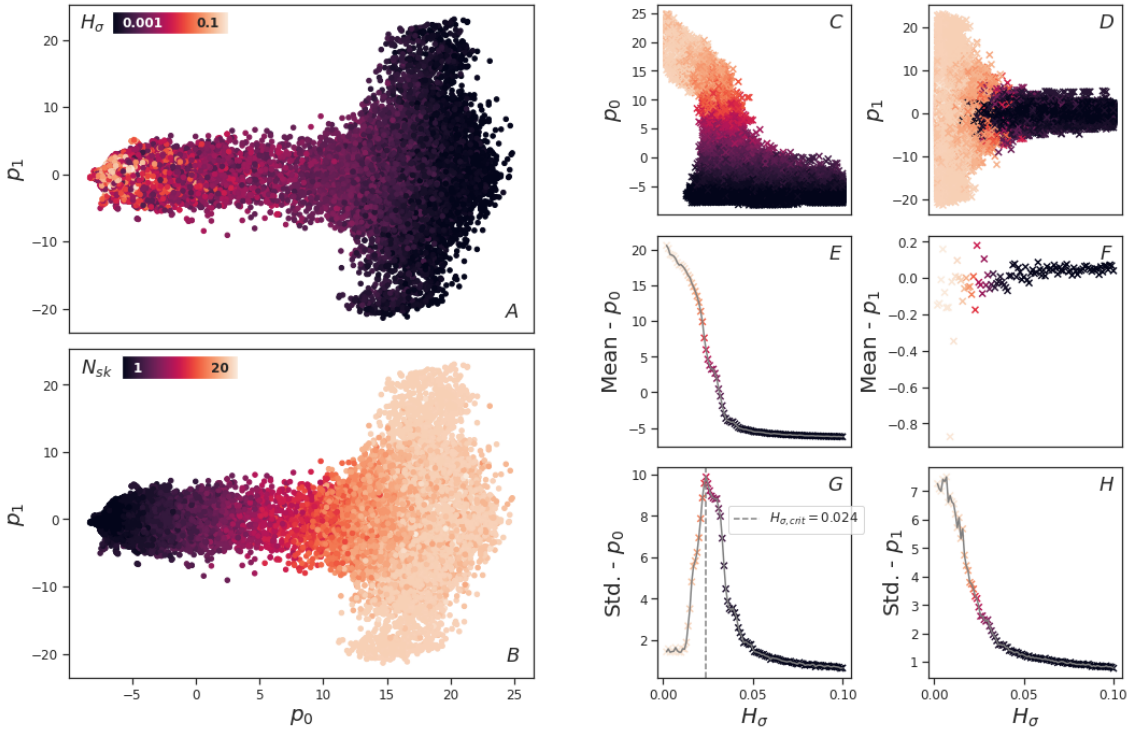

FIGURE 7.15: **PCA subspace of systems along skyrmion nucleation boundary**: The test datasets are transformed into the subspace spanned by the first two principal components (A,B), with the colors representing the defect level (A), and skyrmion number (B). The individual principal components are plotted as function of defect level for μ_0 (C), μ_1, along with their respective means, (E-F) and standard deviations, (G-H). The color of these diagrams represent N_{sk}.

their variance levels. To resolve this, we feed these reduced variables into a PCA with two principal components, leading to a rotation of the latent space to align the first component with the axis of maximum variance.

The resulting latent space, displayed in Figure 7.15, shows that the first principal component (p_0) is sufficient in distinguishing between the OSkL and DSkL phases. Similar to the VAE latent variables, we observe a significant spread of the distribution of p_0 as a function of H_σ (C), where the peak of the standard deviation of p_0 identifies the same critical defect level as the VAE, $H_{\sigma,crit} = 0.024$. This value is relatively close to the true critical defect level, $H_{\sigma,crit} = 0.02056$, determined via a finite-size analysis in Chapter 6. Consequently, both VAE and PCA demonstrate the capacity to distinguish the OSkL from the DSkL phase solely from the configuration at the skyrmion nucleation boundary. However, further analysis that reveals that this ability does not extend to other paths in the phase space along varying defect levels.

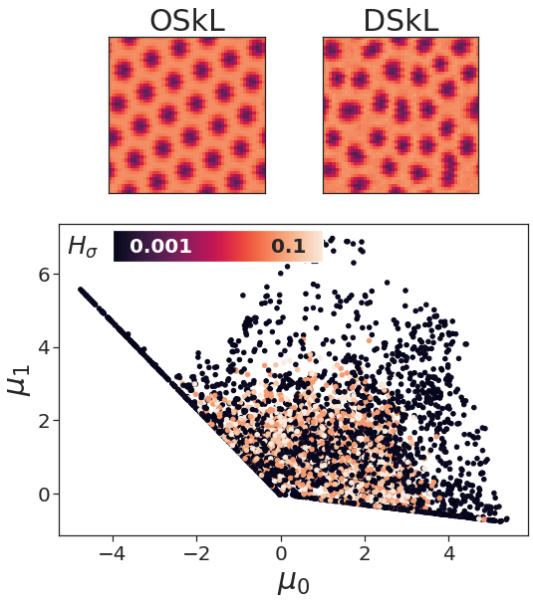

FIGURE 7.16: Latent space of the means for OSkL and DSkL, with the color representing the defect level.

We extend our study to identifying the $H_{\sigma,crit}$ along any phase trajectory at a constant field, with the key focus being on distinguishing an OSkL from a DSkL, using configurations with a similar number of skyrmions. To achieve this, we begin with a simple dataset comprising of only OSkL at very low defect levels, $H_\sigma \in \{0.001, 0.002, 0.003\}$, and DSkL at high defect levels, $H_\sigma \in \{0.08, 0.09, 0.1\}$ at $H = 0.2$. An example configuration of each phase is displayed in Figure 7.16. Surprisingly, the VAE fails to differentiate these phases, as indicated by the overlapping clusters in the latent space of Figure 7.16. Our use of CNNs, which have limited receptive fields and therefore struggle to capture long-range order, may have contributed to this. This limitation resembles the Pathfinder problem in which CNNs can learn local features, such as the outline of an object, but encounter difficulty in learning distant connections, such as whether two points in an image are linked by a path[257]. Some studies have addressed this issue by implementing network modifications, such as horizontal gated recurrent units[257] or skip connections[258]. However, we defer the implementation of such changes to future work.

7.6 Classification of all Phases in Materials with Defects

Our analysis has revealed at least one limitation of the VAE—its inability to distinguish between OSkL-DSkL phases from configurations of ordered and disordered arrangements of skyrmions. To uncover any additional issues, we systematically explore all phases present in our materials and compare the clustering performance of the VAE with our manual approach using N_{sk} and FFT images. For this study, we compile a dataset containing 1000 realizations of all potential phases listed below:

- **Unf(-)**: Uniform state with spins pointing in $-z$. Taken at $H_\sigma = 0.001$, $H = -0.9$.
- **Unf(+)**: Uniform state with spins pointing in $+z$. Taken at $H_\sigma = 0.001$, $H = 0.9$.
- **SkL(+)**: OSkL where the cores of the skyrmions point in $+z$. Taken at $H_\sigma = 0.001$, $H = -0.1$.
- **SkL(-)**: OSkL where the cores of the skyrmions point in $-z$. Taken at $H_\sigma = 0.001$, $H = 0.2$.
- **DSkL(+)**: DSkL where the cores of the skyrmions point in $+z$. Taken at $H_\sigma = 0.1$, $H = -0.1$.
- **DSkL(-)**: DSkL where the cores of the skyrmions point in $-z$. Taken at $H_\sigma = 0.1$, $H = 0.2$.
- **Sk(+)**: Isolated skyrmions where the cores point in $+z$. Taken at $H_\sigma = 0.1$, $H = -0.25$.
- **Sk(-)**: Isolated skyrmions where the cores point in $-z$. Taken at $H_\sigma = 0.1$, $H = 0.32$.
- **Mixed**: A mixed state of meandering domains an skyrmions, Sk(-). Taken at $H_\sigma = 0.1$, $H = 0.13$.
- **Helical**: Helical state. Taken at $H_\sigma = 0.001$, $H = 0.1$.
- **Meandering**: Meandering domains, which nucleate from Unf(-) at low temperatures. Taken at $H_\sigma = 0.001$, $H = 0.01$, $T = 0.5$.

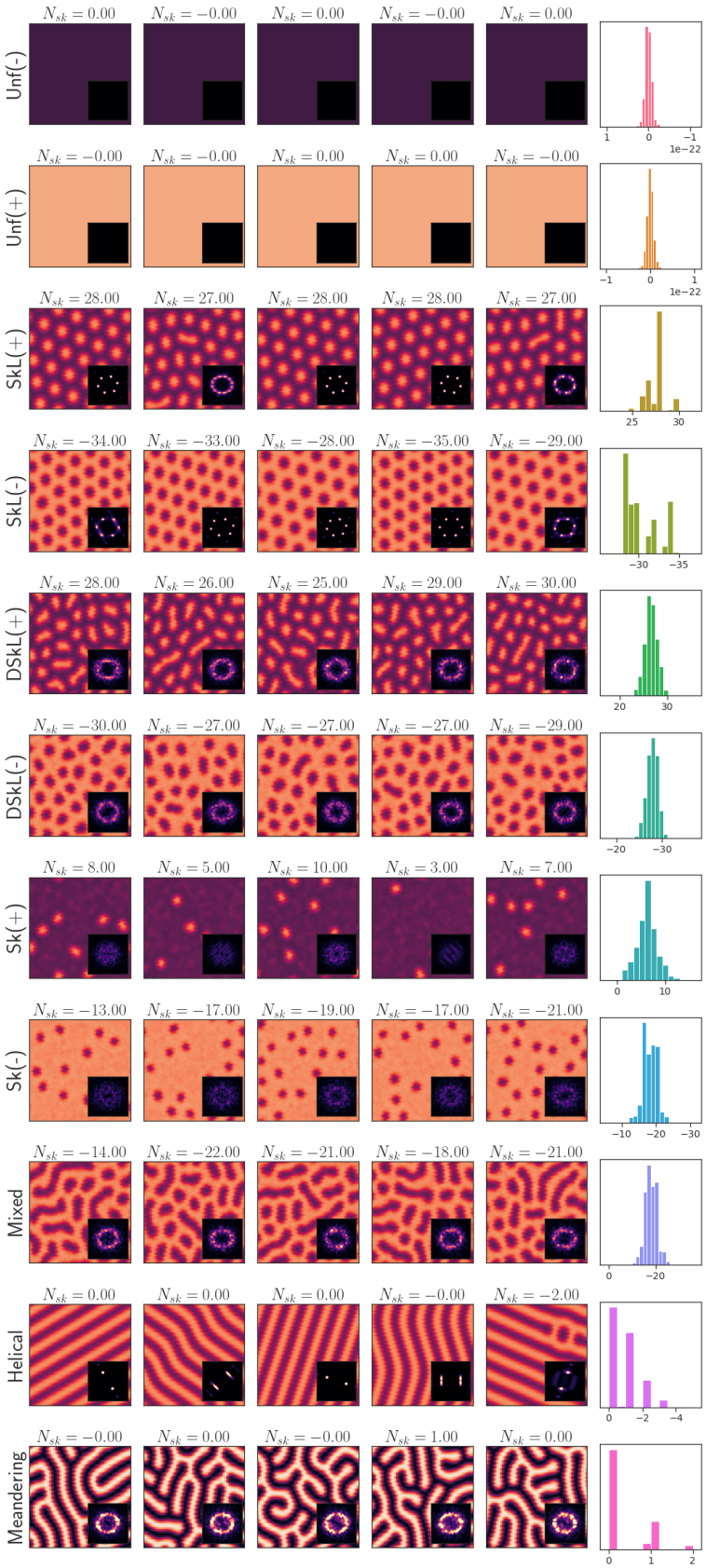

FIGURE 7.17: Example realizations of the spin configurations for all possible phases in our system. The inset shows the reciprocal space from FFT and the final column contains the distribution of N_{sk} over all 1000 realizations.

We took all configurations at $T = 1.0$ unless specified otherwise. Example spin textures of these phases are shown in Figure 7.17, together with their FFT reciprocal space and N_{sk}. It is worth noting that due to the small level of defect in 'clean' systems, certain realisations might not be representative of the phase: for instance, the second realization of SkL(+) is not a fully ordered SkL, as evidenced by the smeared FFT peaks. Additionally, the helical state might include Sk(-) at locations where the spin spirals separate, e.g. the fourth realisation for Helical in Figure 7.17.

The latent space representation of this diverse set of phases, shown in Figure 7.18, highlights the VAE's ability to differentiate all phases with distinct clusters, with the exception of SkL(+)-DSkL(+) and SkL(-)-DSkL(-). Although the VAE cannot capture the LRO in skyrmion lattices, it can still distinguish between the spin spirals in the Helical state and the meandering domains. The smaller separation of LRO information in the Helical state compared to the SkL may cause this difference, with the latter's LRO information likely extending beyond the CNN's receptive field[257].

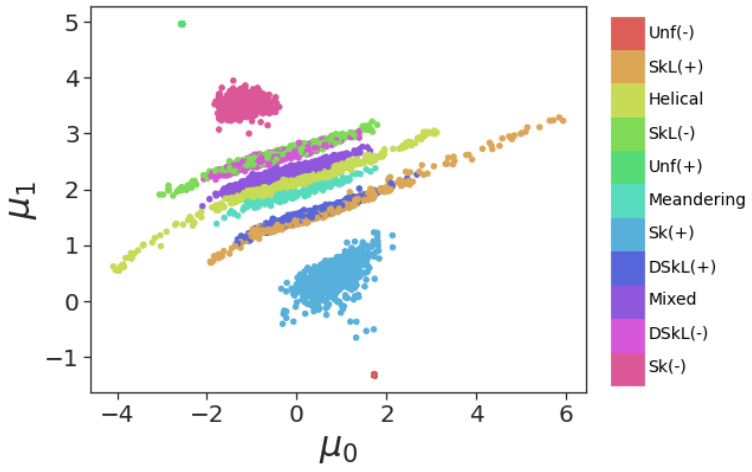

FIGURE 7.18: **Latent space of all phases**: The test dataset is transformed into the means of the latent variables, where the colors represent the phase type.

Furthermore, we observe that the VAE learns to distinguish individual components (skyrmions and meandering domains) separately, as evidenced by the Mixed state forming its own distinct cluster separate from Sk(-), DSkL(-), and the Meandering state. Had the VAE solely relied on N_{sk}, the Mixed state would cluster with Sk(-); conversely, if it solely detected the presence of meandering domains, it would cluster with the Meandering state. Thus, the VAE variables potentially serve as better

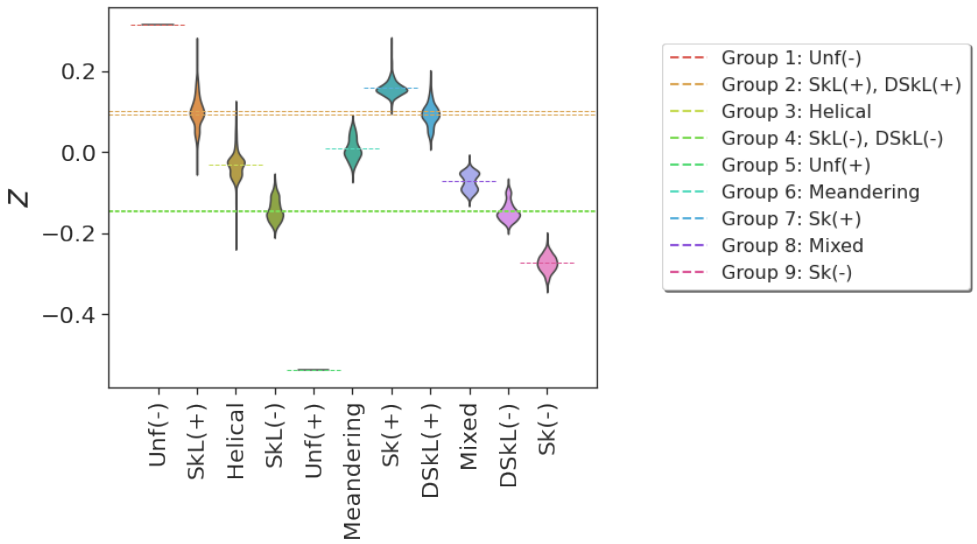

FIGURE 7.19: **1D VAE of latent variables**: Distribution plots of the VAE reduced variable, z, from $\mu_{0/1}, \sigma_{0/1}$, where the vertical dashed lines correspond to the mean at each phase. If the dashed line extends across the whole diagram, then multiple real phases are grouped together.

phase indicators than N_{sk}.

Additionally, while the VAE path in μ_1 no longer reflects a real phase path, it still encodes the Zeeman field progression, transitioning from the Unf(-) state at very low μ_1 to Unf(+) at the highest μ_1. Once again the broader clusters, observed at lower fields, suggest a greater variety in configurations in the initial nucleation of Sk(+) and DSkL(+)-SkL(+) compared to later Sk(-) and DSkL(-)-SkL(-) phases.

We condense the learned information further from $\mu_{0/1}, \sigma_{0/1}$ into a single variable, z, using a VAE composed of dense layers with $(32, 16)$ nodes in the encoder. The distributions of z for each phase are illustrated with violin plots in Figure 7.19. We categorise similar VAE phases based on whether the mean of one phase, $\langle p_i \rangle$, falls within 0.1 standard deviations of another phase, $\langle p_j \rangle$, resulting in only two phase groups with multiple phases, SkL(+)-DSkL(+) and SkL(-)-DSkL(-). These are highlighted by the full vertical dashed lines in Figure 7.19. All other groups consist of only a single phase, yielding 9 distinct phases detected by the VAE.

We evaluate the classification performance of the VAE against our manually computed parameters, N_{sk} and FFT. The FFT analysis yields a 50×50 image representing the reciprocal space of the spin texture, where peaks correspond to the

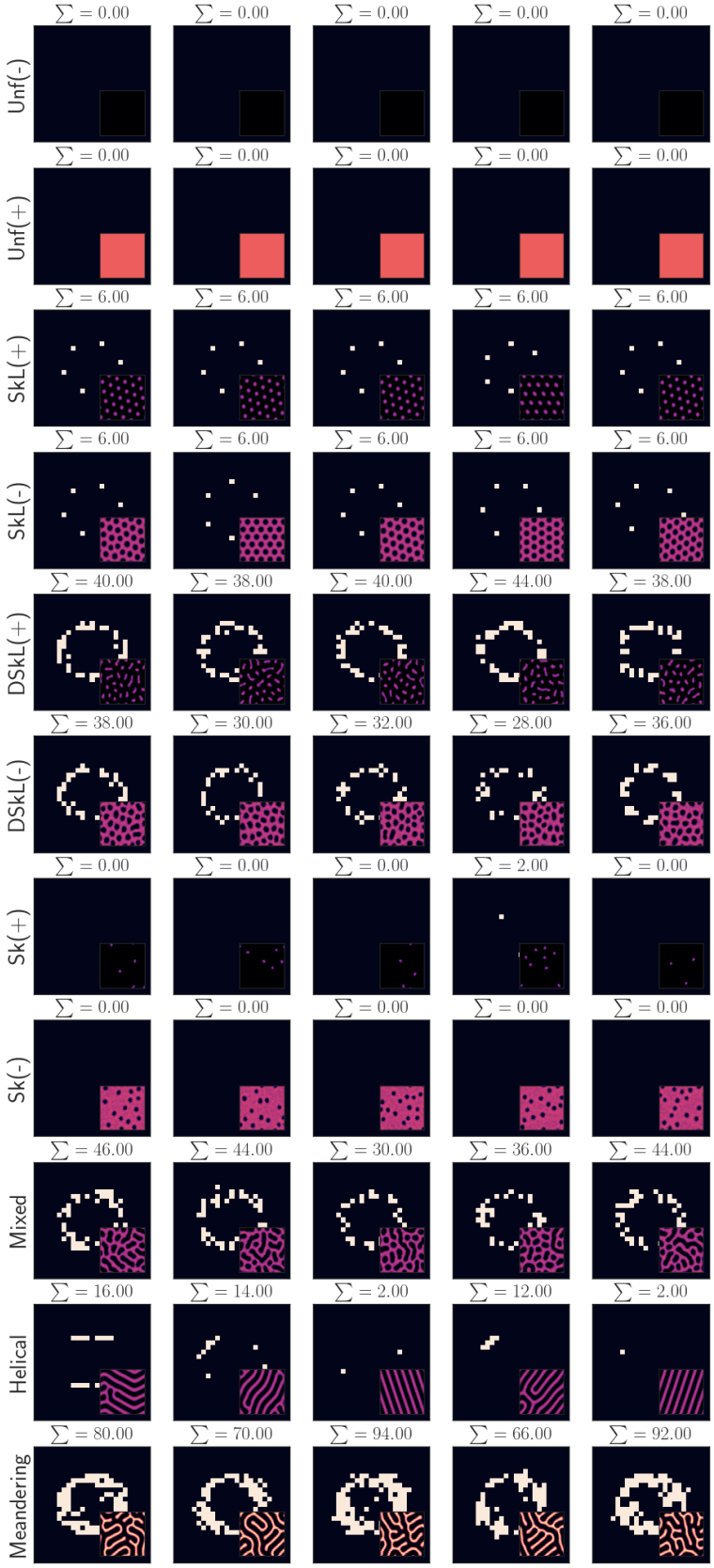

FIGURE 7.20: **FFT Analysis**: The peaks of the FFT reciprocal space, where pixels above the threshold ($I_{\text{thresh}} = 0.00006 I_{\text{max}}$) are set to 1, otherwise they are set to 0. The inset displays the real-space spin configuration, and the title contains the $\sum \text{FFT}_{\text{peak}}$.

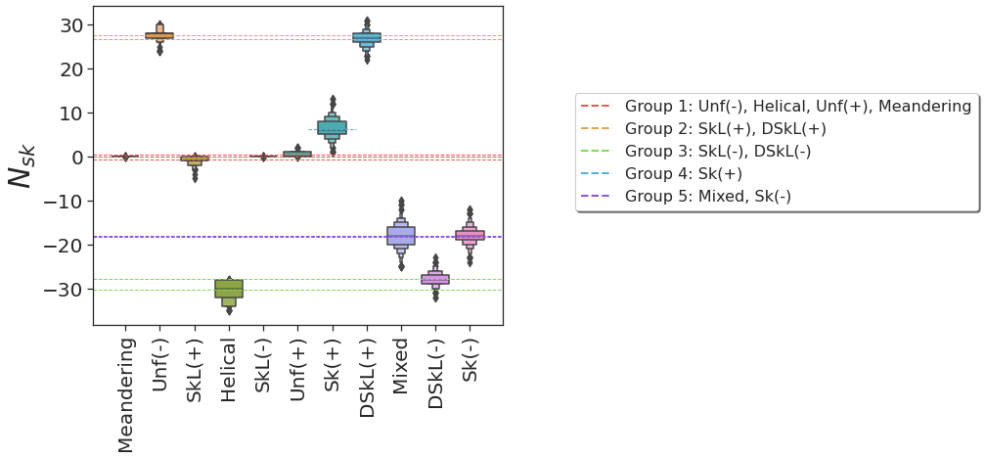

FIGURE 7.21: N_{sk} **of all phases**: Distribution plots of the PCA reduced variable, p, from $\mu_{0/1}, \sigma_{0/1}$, where the vertical dashed lines correspond to the mean at each phase. If the dashed line extends across the whole diagram, then multiple real phases are grouped together.

spatial frequencies of repeated textures such as skyrmions or spin spirals. To facilitate comparative analysis, we simplify this to a single variable by setting all pixels above a certain threshold intensity ($I_{\text{thresh}} = 0.00006 I_{\text{max}}$) to 1, and the rest to zero, and then summing over the matrix (Figure 7.20). In highly ordered SkL phases, $\sum \text{FFT}_{\text{peak}} = 6$, reflecting the six peaks in the image, a clean helical phase should yield $\sum \text{FFT}_{\text{peak}} = 2$, while progressively dirty systems (DSkL($\pm$), Mixed, Meandering) exhibit a ring structure with much larger areas. Conversely, the uniform and isolated skyrmion states should have the smallest pixel areas, as they lack LRO spin structures. It is worth noting that due to the small level of defects in 'clean' systems, there may be some smearing of the sharp peaks in SkL($\pm$) and Helical states, leading to larger than expected areas.

Categorising the phases using the same method applied to the VAE p parameter, N_{sk} results in 5 distinct groups, where 10 of the 11 phases are clustered with other phases (Figure 7.21). For instance, the first group, characterised by $N_{sk} \approx 0$, contains the 4 non-topological phases (Unf($\pm$), Helical, and Meandering), whereas the VAE distinguishes these as distinct phases, and thereby outperforms N_{sk}. Additionally, N_{sk} struggles to differentiate the Mixed state from Sk(-), as the average N_{sk} values for these phases are almost identical. This difficulty arises because the ends of the meandering domains in the Mixed state are considered half-skyrmions, which shrink to isolated skyrmions (Sk(-)) at higher fields, resulting in no change in N_{sk}. In

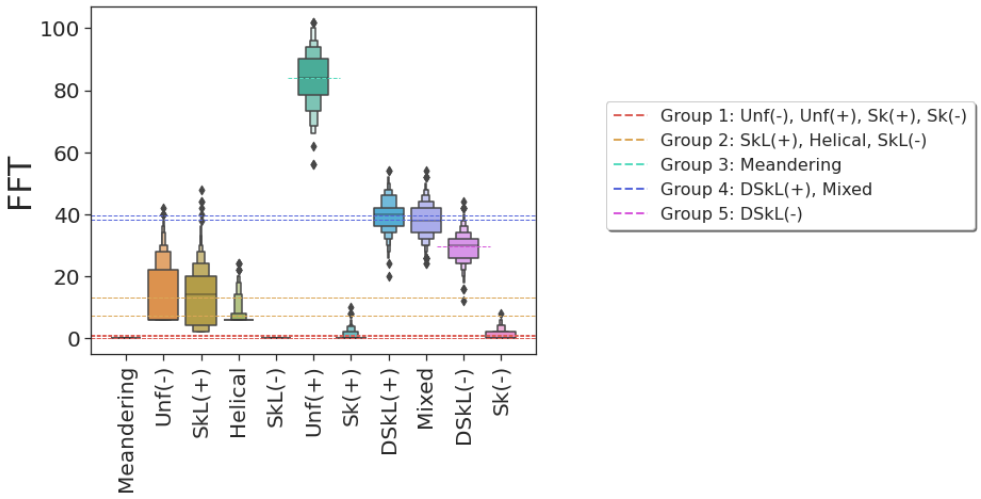

FIGURE 7.22: **FFT of all phases**: Distribution plots of the PCA reduced variable, p, from $\mu_{0/1}, \sigma_{0/1}$, where the vertical dashed lines correspond to the mean at each phase. If the dashed line extends across the whole diagram, then multiple real phases are grouped together.

contrast, the VAE effectively distinguishes these phases, with a notable difference in p values for the two clusters. Coupled with these difficulties, N_{sk} also faces the same challenge as p in failing to differentiate SkL($\pm$) from DSkL($\pm$)[5], thus offering no advantage over the VAE parameter.

When considering the FFT results, this metric categorises phases based on the level of LRO detected in the real-space configuration (Figure 7.22), also resulting in 5 distinct group. The first group comprises systems with no LRO, including Unf($\pm$) and Sk($\pm$), with an average peak area of 0. However, our VAE outperforms this metric by demonstrating a smaller difference in p between Sk(+) and Unf(-), but a significant difference between Unf(+) and Sk(-), while still maintaining separate clusters in the latent space (7.19). The smaller difference for Unf(-)-Sk(+) arises from the greater number of ways Sk(+) can emerge from Unf(-), resulting in a wider area spanned by Sk(+) in the latent space, which is therefore closer to Unf(-). Conversely, because Sk(-) forms a denser cluster, there is a sharper distinction between the Sk(-) and Unf(+) in the latent space.

The three phases characterised by LRO, SkL($\pm$) and Helical, form the second phase group in the FFT analysis. These phases exhibit a broad distribution in peak area due to the smearing effect caused by a small amount of defect present in 'clean' systems, leading to a large standard deviation that captures all three state into a

[5]The standard deviation of the SkL(-) phase was relatively high, leading to a group with DSkL(-)

single group. DSkL(+) and Mixed form another group distinguished by their distinctive ring pattern in FFT. Although DSkL(-) exhibited a similar ring pattern, it was likely not grouped with these phase due to slightly more order in the structure at this particular field ($H = 0.2$), attributed to the smaller skyrmions in the system. DSkL(+) featured more elongated skyrmion structures, resulting in a peak are closer to that of the Mixed state.

In summary, although the FFT and N_{sk} measures yield 5 distinct phase groups, the VAE generates 9, showing its superiority in phase classification compared to our manual methods, even in the absence of contextual information, such as the interactions in the Hamiltonian. However, our analysis shows that it is not foolproof, as it encounters difficulty in distinguishing between the phases where LRO information is distant, for example SkL($\pm$)-DSkL($\pm$). Consequently this leads to difficulties in identifying the $H_{\sigma,crit}$ for all phase paths except at the skyrmion nucleation boundary.

7.7 Conclusions

In conclusion, our systematic study of VAE reveals its capability of generating properties from spin textures that more uniquely characterise its phases compared to conventional properties. However, it still struggles to classify phases that exhibit LRO over a wider distance, such as in skyrmion lattices. This limitation does not imply the absence of a phase transition between OSkL and DSkL but rather indicates that the CNNs in our VAE struggle to encode spatially distant information. Similar limitations have been observed in other studies, suggesting that implementing necessary network modifications could enable the VAE to successfully differentiate all phases in materials with defects. While several studies have found success in using the VAE for phase classification[259–264], our findings emphasize the importance of careful model design, as an off-the-shelf approach may not ensure full differentiation of all phases.

Chapter 8

Supervised Learning of Defect Fields

8.1 Motivation

Standardising the manufacturing process of sputtered thin films is imperative for the development of skyrmion-based devices, as defects can significantly disrupt the formation of densely packed OSkL. Consequently, the implementation of online defect detection is essential. While certain studies have leveraged a convolutional neural network (CNN) to predict global Hamiltonian parameters[265–267], our method aims to identify the complete localised defect field across the lattice. This enables a more comprehensive analysis of the defect distribution and its influence on magnetic behavior.

We use a deep learning architecture, based on U-Net[1], that learns the relationship between the spin texture and the localised defects, where defects are modelled as random pinning fields[2]. Expanding upon this framework, our objective is to identify all system parameters—whether local (L) or global (G)—that can accurately reproduce a specific texture. For instance, in a real polycrystalline material, site-dependent uniaxial anisotropy, $\boldsymbol{J}_k \in \mathbb{R}^{N_x \times N_y}$, exchange strength, $\boldsymbol{J}_{ex} \in \mathbb{R}^{N_x \times N_y}$, and a global DMI strength, J_{dm}, may be present. Consequently, we seek to learn the

[1] See Chapter 3, Section 3.4 for more information on the U-Net.
[2] See Chapter 2 for more information on defect modelling.

mapping f defined as:

$$f(\boldsymbol{X}) = \begin{bmatrix} \boldsymbol{J}_k \\ \boldsymbol{J}_{ex} \\ \boldsymbol{J}_{dm} \end{bmatrix}, \tag{8.1}$$

where $\boldsymbol{X}$ is the spin texture. Notably, if in future, the mean field model can be calibrated to experimental data, then this neural network can be trained using experimental real-space spin configurations, offering a means of extracting material parameters without the need for trial-and-error fitting of the Hamiltonian.

8.2 Learning the RP Field (1L)

8.2.1 Model and Dataset

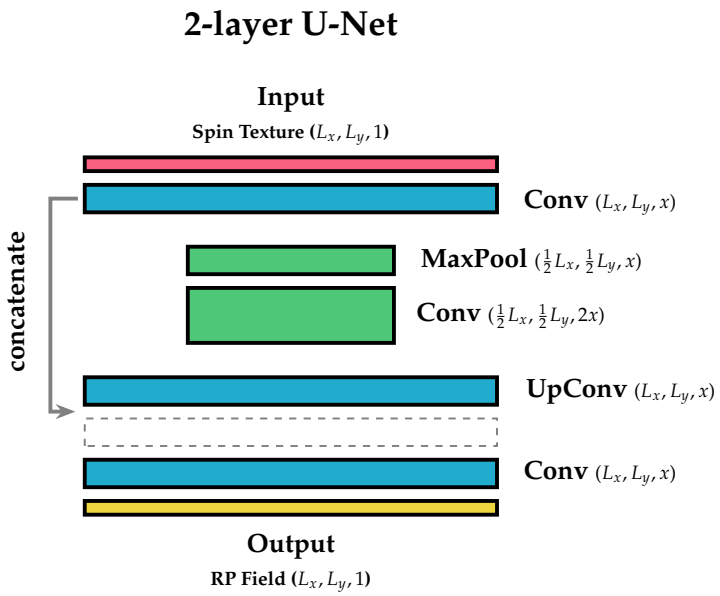

FIGURE 8.1: 2-layer U-Net architecture.

We initially focus on learning a single localized parameter (1L): the defect field, $H_{z,i}$. To accomplish this, we repurpose the U-Net model, originally designed for image segmentation tasks[3], as a general mapping function between a 2D input (the true defect field) and a 2D output (the predicted defect field). Our model is a simplified version of the original U-Net consisting of 4 units in the contracting path, each containing a 3×3 convolutional layer followed by a 2×2 max pooling layer, which

[3]For more details, refer to Chapter 3, Section 3.4.

reduces the dimensions of the input by half in the x and y directions. The expanding path reverses this procedure, including the skip connections to the appropriately sized matrices from the encoder path. The final layer is a 1×1 convolutional layer leading to a single-channel output with a sigmoid activation function that predicts the RP fields. Although Figure 8.1 illustrates a simple two-layer architecture, our model comprises four blocks with convolution layers containing [32, 64, 128, 256] nodes, and uses a `LeakyReLu` activation function[4], over the standard `ReLU`, to prevent dying nodes[5]. The loss function, used for training, is the sum of the mean squared error (MSE) of the RP field at each lattice point,

$$\text{loss} = \sum_{i,j} (\boldsymbol{Y}_{ij} - \hat{\boldsymbol{Y}}_{ij})^2, \tag{8.3}$$

where $\boldsymbol{Y}_{ij}$ and $\hat{\boldsymbol{Y}}_{ij}$ are the true and predicted RP fields respectively, and $\boldsymbol{Y}_{ij}, \hat{\boldsymbol{Y}}_{ij} \in \mathbb{R}^{48\times 48}$. We crop the original 50×50 dimensions to 48×48 in order to to get well-behaved sizes during downsampling and upsampling, and only consider the z-component of the spin texture to reduce the complexity of the problem. Additionally, both the input and output are normalised to values within the range $[0, 1]$.

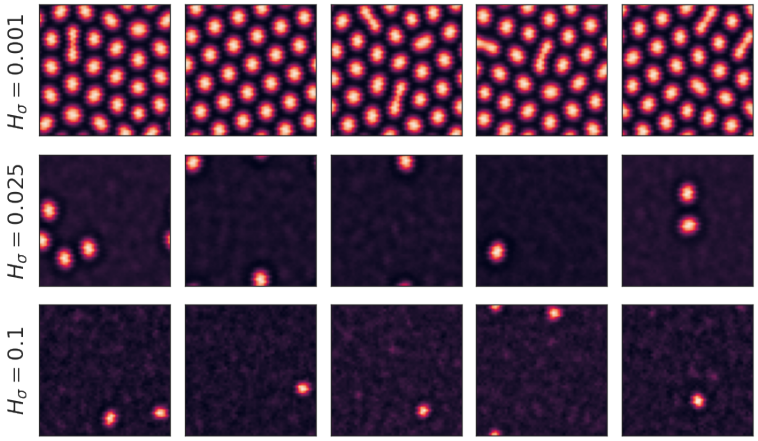

FIGURE 8.2: Five example spin configurations from our training set at $H_\sigma \in \{0.001, 0.025, 0.1\}$ along the skyrmion nucleation boundary.

[4]The `LeakyReLu` activation function can be mathematically defined as,

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases} \tag{8.2}$$

where: α is a small positive constant, typically 0.01.

[5]Dying nodes occur due to the vanishing gradient problem, where the gradients are small and therefore when these are back-propagated to the earliest layers, these weights may not be updated.

Figure 8.2 shows some examples of skyrmion spin textures at $H_\sigma \in \{0.001, 0.025, 0.1\}$. The dataset only considers configurations at the skyrmion nucleation boundary, where low-H_σ yields a full hexagonal lattice of skyrmions, high-H_σ results in only a single or a few skyrmions, and intermediate-H_σ have realisations exhibiting a full lattice or only a few skyrmions (see Chapter 6). The input dataset, containing 36,000 spin textures captured at 18 σ points with $\sigma \in [0.001, 0.1]$ and 2000 realizations for each, was shuffled and partitioned with 50% allocated to the training dataset, and 25% each to the validation and test datasets. The validation dataset was used for architecture and hyperparameters optimisation and was excluded from the training process. In particular, the number of layers, number of nodes in each layer, and model parameters which minimised the validation loss was chosen as the optimal network.

8.2.2 Results

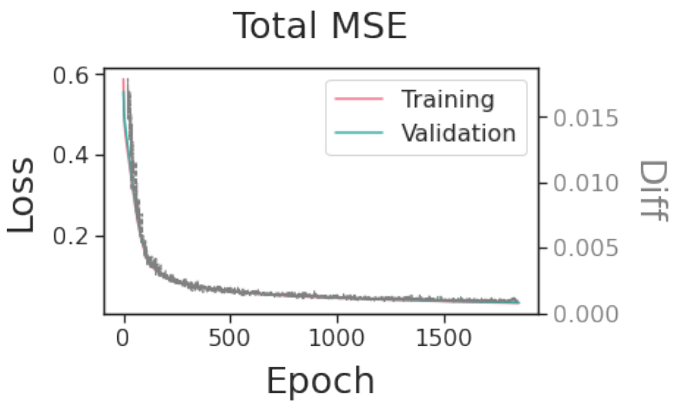

FIGURE 8.3: Convergence of the loss function over the 1855 epochs for the training and validation dataset. The grey dashed line shows the difference between the validation and training loss.

The model was trained using the stochastic gradient descent optimiser (see Chapter 3) with a learning rate of 0.1 over 4000 epochs, and an `EarlyStopping` checkpoint which terminates the training when the total change in the validation loss is below 0.00001. A `ModelCheckpoint` saves the network weights when the total loss of the validation dataset is at its lowest. All models in this chapter use this training protocol and checkpoints unless specified otherwise. The model converges after 1855

epochs, with similar performance on both the training and validation set, as indicated by the loss function in Figure 8.3. The `EarlyStopping` checkpoint prevents overfitting, yielding an optimised model with a total MSE of 0.004180 on the test set.

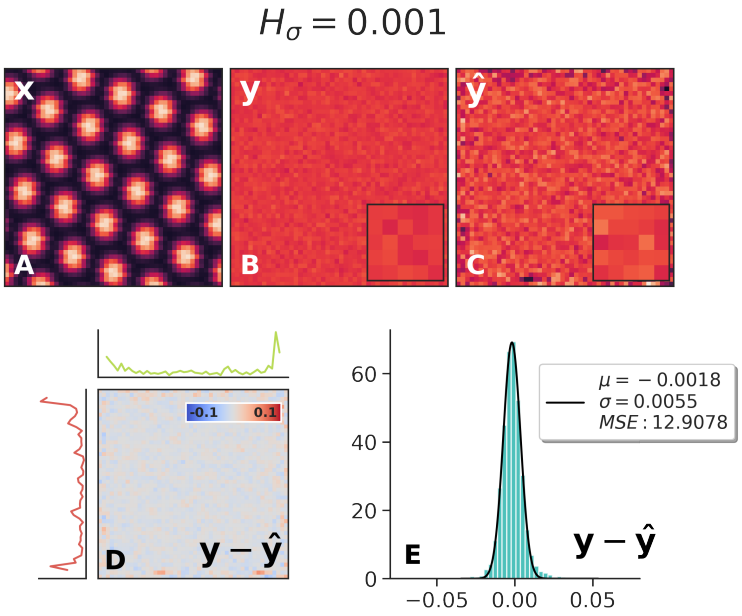

FIGURE 8.4: An example prediction of the RP field on unseen spin texture in a 'clean' system, $H_\sigma = 0.001$. A) The input spin configuration, x, B) the true RP field, y, C) the RP field predicted by the model, $\hat{y}$, D) the pixel-wise difference between the true and predicted fields, $y - \hat{y}$, and E) a distribution of this difference. The color scale is identical for the true and predicted RP fields (B,C), but due to the small sizes of the defects, the difference between the two is magnified.

Figures 8.4-8.5 shows an example prediction of the RP field on an spin texture from the test dataset for a 'clean' (8.4) and 'dirty' (8.5) system. Despite being an unseen input, the model accurately predicts the RP field at the pixel level, with minimal differences between the true (y, B) and predicted field ($\hat{y}$, C) at most lattice sites (D). However, the mean square error ($\sum_i (y_{ij} - \hat{y}_i)^2$) along each axis (top and left of D), reveals discrepancies at the system boundaries. This could stem from the splicing of the original input and RP field from 50×50 to 48×48 matrices, required by the U-Net, or due to the zero-padding implemented by TensorFlow's convolutional layers, which may not preserve periodic boundary conditions.

Furthermore, the $y - \hat{y}$ plot of the 'dirty' system (8.5D) shows an imprint of the original skyrmions, indicating that the U-Net struggles to learn the defect field from these structures. This is to be expected, as we hypothesise that defects exert a more localised influence on uniform spins, embedding themselves directly into the

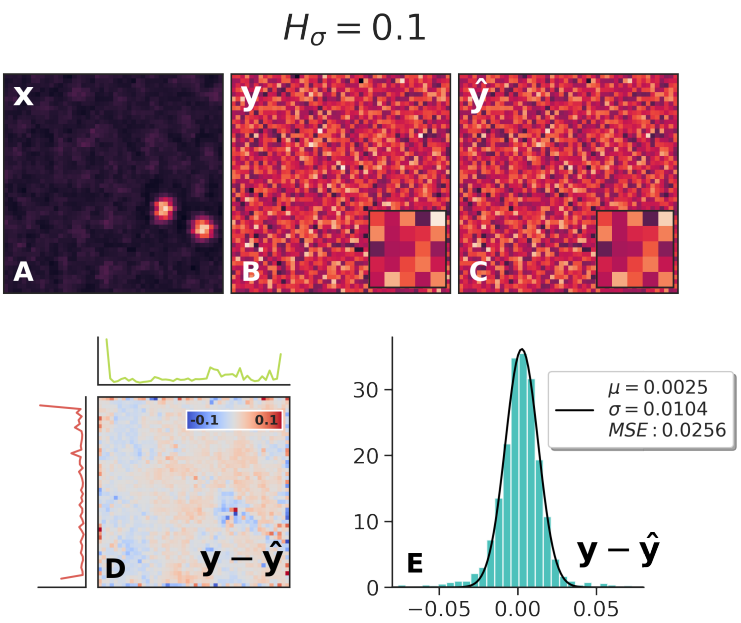

FIGURE 8.5: An example prediction of the RP field on unseen spin texture in a 'dirty' system, $H_\sigma = 0.1$. A) The input spin configuration, x, B) the true RP field, y, C) the RP field predicted by the model, $\hat{y}$, D) the pixel-wise difference between the true and predicted fields, $y - \hat{y}$, and E) a distribution of this difference.

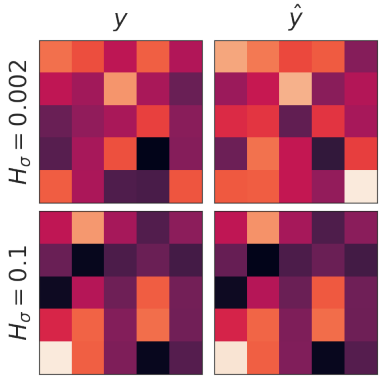

FIGURE 8.6: Comparison of random patches of the RF field between the true values (first column) and the predicted values (second column) for the 'clean' ($H_\sigma = 0.002$) and 'dirty' ($H_\sigma = 0.1$) cases.

z-component of the spins, whereas they are more likely to alter the shape, stability, or location of topological particles, rather than its internal structure. Nevertheless, the error distribution (E), fitted with a Gaussian ($Ae^{\frac{-(x-\mu)^2}{2\sigma^2}}$), reveals that the average pixel difference (μ) is approximately only 6.72% of the original RP field in the 'dirty' system. For comparison, smaller patches of the RP field are plotted in Figure 8.6 for different defect levels, showing a high degree of similarity between the true and predicted values, especially in the 'dirty' case.

We evaluate the U-Net's performance across different defect levels by analysing

the error distributions, illustrated in Figure 8.7. To ensure comparability, the MSE score, giving the summed pixel by pixel squared difference, has been normalised by squared sum of the true RP field, facilitating a comparison against a baseline prediction of 0. As defect levels increase, the MSE decreases significantly, highlighting the U-Net's improved ability to predict the defect field in such systems. This is likely attributed to a combination of higher-defect configurations containing mostly uniform regions, which are more sensitive to defects, and a stronger defect signal. These findings suggest that the training dataset should include more spin configurations at low defect levels to enhance prediction accuracy in these systems.

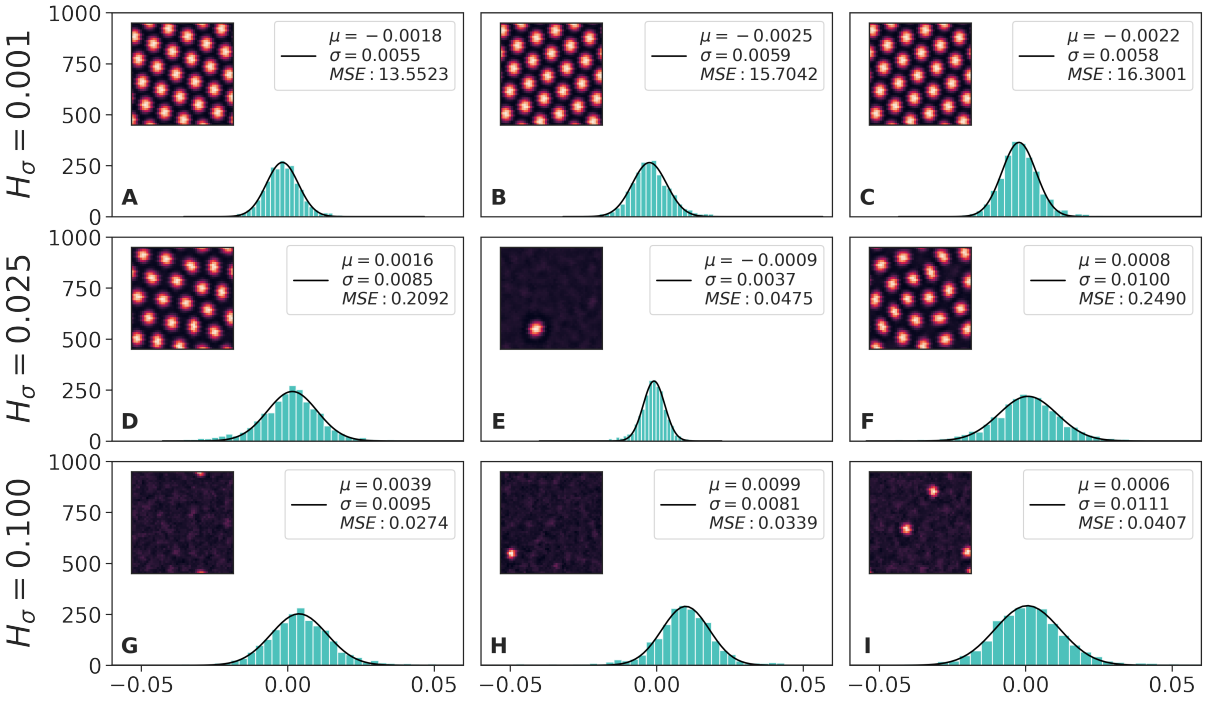

FIGURE 8.7: The distribution of difference, $y - \hat{y}$, for random examples from the test dataset at different H_σ levels. The inset shows the input spin texture and the legend gives the (μ, σ) values of the fitted Gaussian ($Ae^{\frac{-(x-\mu)^2}{2\sigma^2}}$), together with the scaled MSE.

At the intermediate defect level ($H_\sigma = 0.025$), we observe differing spreads in the error distribution depending on the input spin texture: a skyrmion lattice results in a wider distribution, while single skyrmions lead to a narrower distribution. As a function of defect level, both the spread, captured by the fitted σ, and MSE demonstrate a split in their distributions (see the inset of Figure 8.8A and Figure 8.8B), with a relatively unimodal distribution in 'clean' and 'dirty' systems (8.8(C, E,

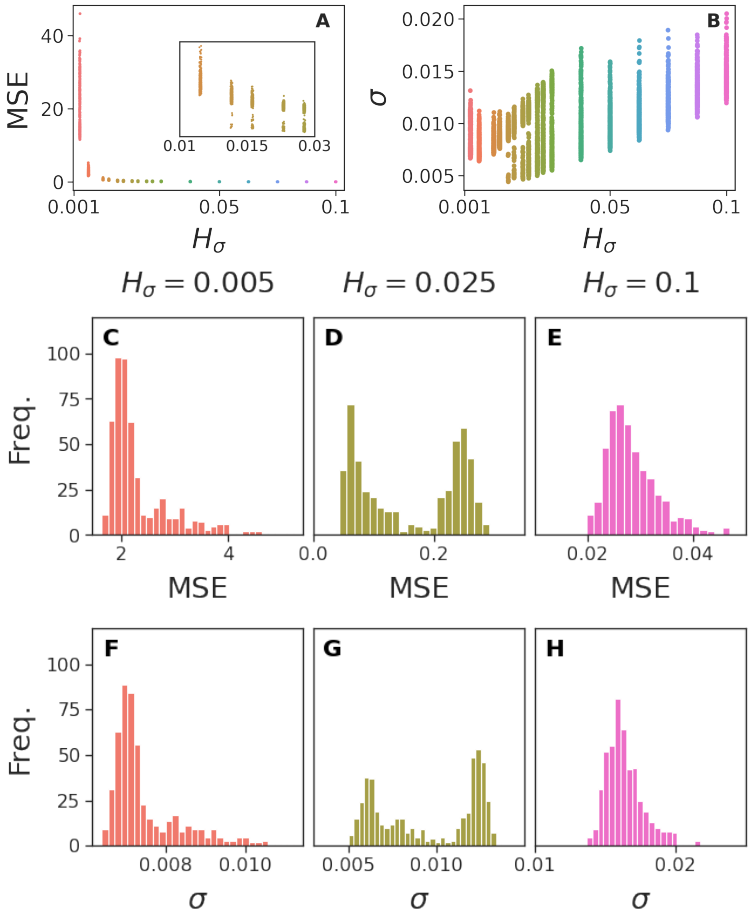

FIGURE 8.8: A) The normalised MSE and B) the fitted σ as a function of H_σ for all inputs in the test set, illustrating the split from a unimodal distribution to a bimodal one at intermediate H_σ. Histograms of the (C-E) MSE, and (F-G) σ showing a unimodal distribution at low and high H_σ (C,E,F,H), and a bimodal one at $H_\sigma = 0.025$ (D,G).

F, H)), that becomes bimodal at intervening defect levels (8.8(D, G)), close to the critical OSkL-DSkL transition boundary. This is the most interesting outcome from the U-Net analysis, where σ and MSE mirrors the behaviour of the spin-spin correlation observed in Chapter 6, lending further credibility to the existence of a phase transition dependent on defect level ($H_{\sigma,crit}$), particularly considering previous studies have observed a similar degradation in parameter estimation accuracy near transition boundaries[268].

While a simple U-Net successfully predicts the localised defect field for spin textures generated by a specific set of Hamiltonian parameters (J_{ex}, J_{dm}), real materials may possess intrinsic interaction strengths that differ significantly from those simulated in our study. Unfortunately, when evaluating the same model on spin textures generated under different Hamiltonian parameters ($J_{ex} \in \{0.4, 0.45, 0.55, 0.6\}$), we

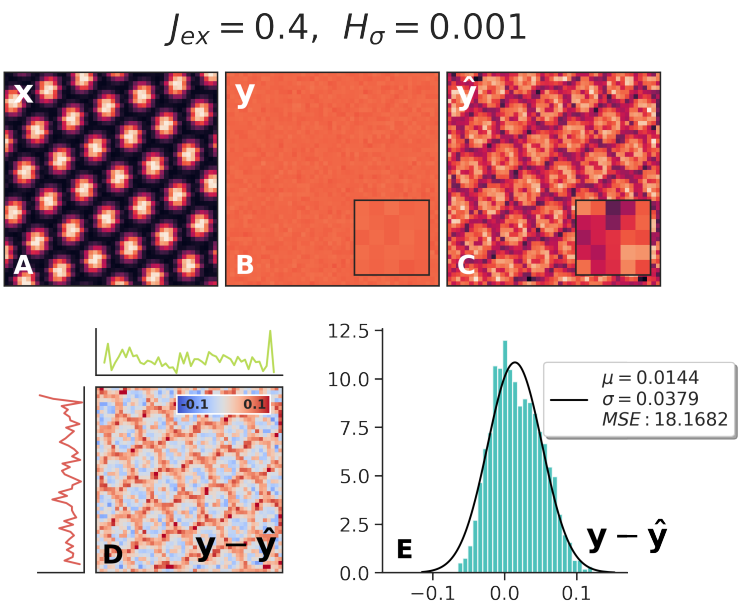

FIGURE 8.9: A) The full error distribution, $y - \hat{y}$, and B) U-Net MSE evaluated at different J_{ex}. Example predictions at $H_\sigma = 0.001$ for $J_{ex} = 0.4$ (C-G) and $J_{ex} = 0.6$ (H-L), displaying the imprinted skyrmion lattice in the predicted RP field.

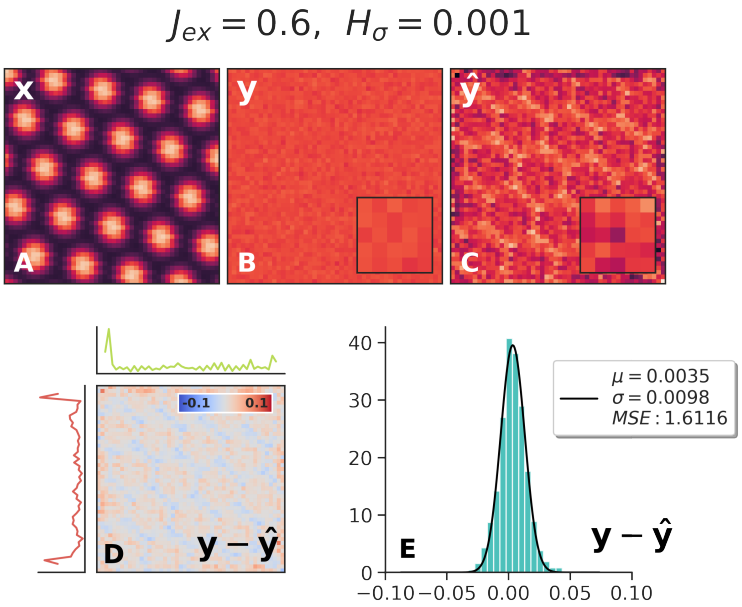

FIGURE 8.10: A) The full error distribution, $y - \hat{y}$, and B) U-Net MSE evaluated at different J_{ex}. Example predictions at $H_\sigma = 0.001$ for $J_{ex} = 0.6$ (C-G) and $J_{ex} = 0.6$ (H-L), displaying the imprinted skyrmion lattice in the predicted RP field.

observe a significant degradation in accuracy. For instance, example predictions for 'clean' systems at $J_{ex} = 0.4$ (Figure 8.9) and $J_{ex} = 0.6$ (Figure 8.10) show sharp imprints of the skyrmion lattice in the predicted RP field ($\hat{y}$), particularly pronounced at lower J_{ex}. We hypothesise that this error arises from the variation in skyrmion

sizes (proportional to J_{ex}/J_{dm}), making it challenging for a model trained on fixed-size skyrmions to accurately predict RP fields for skyrmions of different sizes. Future research may explore the implementation of scale and rotation invariant filters, extending the model's capability to predict defect maps across a wider range of parameters. However, in our study, we take a different approach by modifying the architecture to enable the prediction of both the local RP field and the global J_{ex}.

8.3 Learning the RP Field and J_{ex} (1L, 1G)

8.3.1 Model and Dataset

The following section explores an extension of the U-Net architecture to concurrently learn both a local parameter (1L), the RP field, and a global parameter (1G), the exchange strength, through the integration of a secondary branch originating from the bottleneck (Figure 8.11). Following the final convolution layer, this additional branch flattens the layer and subsequently reduces the number of nodes through k Dense layers, each comprising n_k nodes, before finally generating a prediction for J_{ex}. Meanwhile, the first branch executes the same up-convolution procedure as previously to predict the RP field. The updated architecture is illustrated in Figure 8.11.

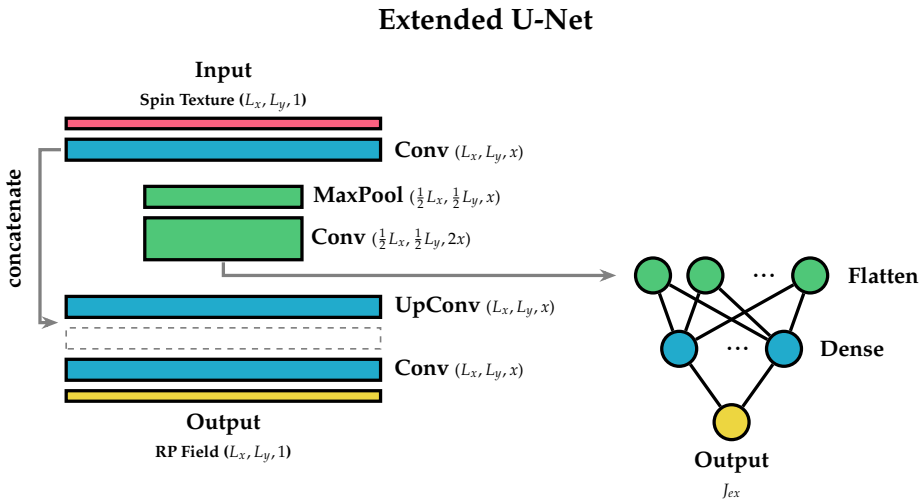

FIGURE 8.11: The extended U-Net architecture, with two branches. The first branch is the original 2-layer U-Net, predicting the RP field, while the second branch extends from the bottleneck by flattening the nodes in the convolution layer and then applying successive Dense layers, before outputing a single value for J_{ex}.

To train the model, we construct a loss function by combining the MSE of the RP field with the MSE of J_{ex}, as follows:

$$\text{loss} = \gamma \cdot \sum_{i,j} (\boldsymbol{Y}_{ij} - \hat{\boldsymbol{Y}}_{ij})^2 + (\boldsymbol{Z} - \hat{\boldsymbol{Z}})^2, \tag{8.4}$$

where $\boldsymbol{Y}_{ij}$ and $\hat{\boldsymbol{Y}}_{ij}$ are the true and predicted RP fields as before, and $\boldsymbol{Z}$, $\hat{\boldsymbol{Z}}$ are the true and predicted exchange strength, J_{ex}. The parameter γ can be adjusted to scale the loss from the RP field relative to the loss in J_{ex}. Inputs and outputs are again normalised, whilst maintaining the relative weighting between each output parameter. More specifically, parameter $p \in \{J_{ex}, H_{z,i}\}$ is scaled across all realisations as follows,

$$\begin{aligned} p_s &= \frac{p - p_{\min}}{p_{\max} - p_{\min}}, \\ p_{\min} &= \min\Big(\min\left(J_{ex}\right), \min\left(H_{z,i}\right)\Big), \\ p_{\max} &= \max\Big(\max\left(J_{ex}\right), \max\left(H_{z,i}\right)\Big). \end{aligned} \tag{8.5}$$

We train the model on spin configurations at the skyrmion nucleation boundary across 14 distinct H_σ values within the range $[0.001, 0.1]$ along with 5 J_{ex} values in [0.4, 0.6], each comprising 2000 realisations. Example realisations of the dataset at these various points are illustrated in Figure 8.12. The dataset contains 140,000 configurations, which are then shuffled and divided into training, validation and test sets with sizes of [70,000, 35,000, 35,000], respectively.

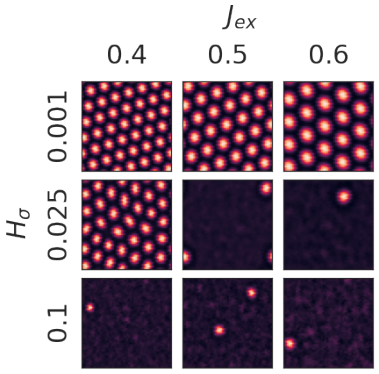

FIGURE 8.12: Example spin configurations from our training set at $J_{ex} \in \{0.4, 0.5, 0.6\}$ and $H_\sigma \in \{0.001, 0.025, 0.1\}$.

We optimise the hyperparameters (γ) and architecture (number of layers and nodes in the U-Net and Dense branches) by identifying settings with the lowest

validation loss, and then evaluate our optimal model's performance on the unseen test set. The optimised model is characterised by the following parameters: $\gamma = 2$, U-Net architecture = [16, 32, 64, 128, 256], Dense architecture = [256, 128, 64, 32, 16], resulting in a test loss of 0.0206645. Details of the tuning results are given in the Appendix, Section A.3.

8.3.2 Results

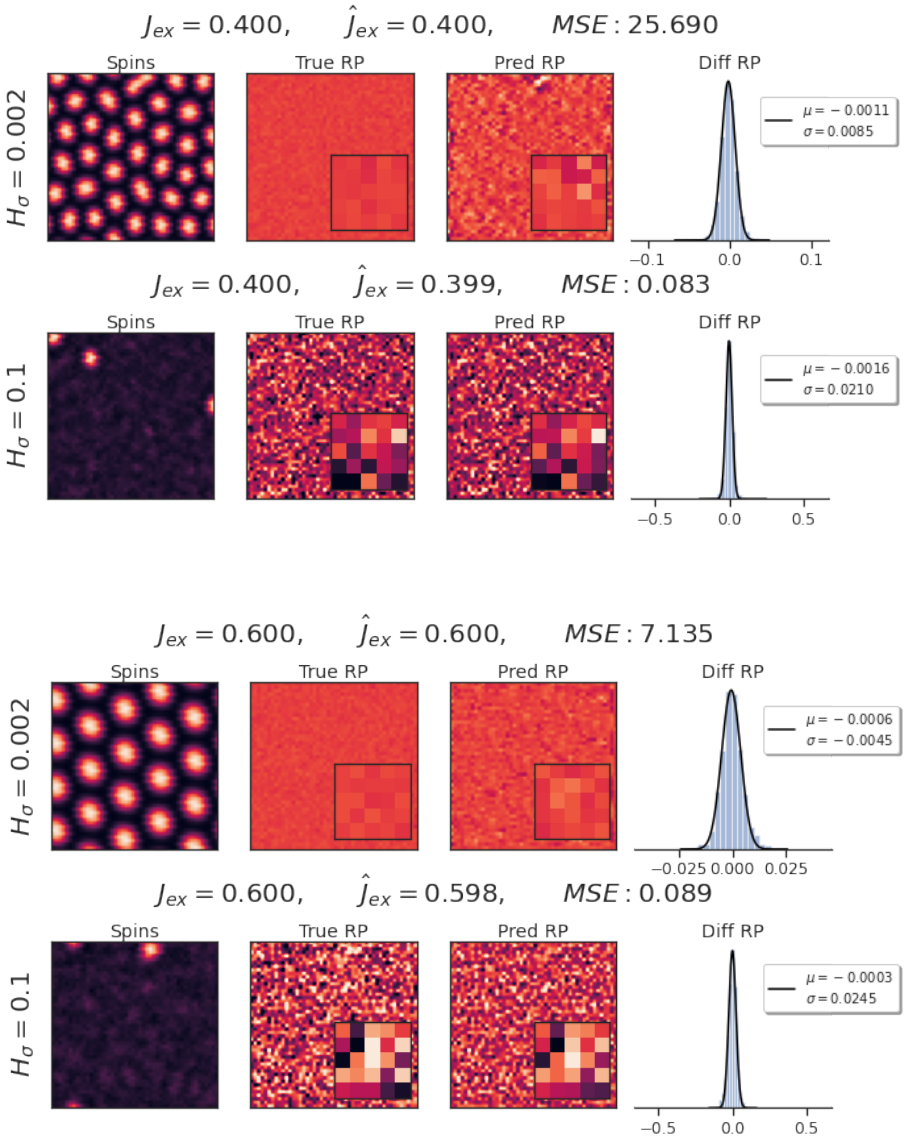

FIGURE 8.12: Each row shows the input and output for a pair H_σ-J, where $H_\sigma \in \{0.002, 0.1\}$ and $J_{ex} \in \{0.4, 0.6\}$. The first column shows the input spin configuration, second and third columns are the true and predicted RP field, fourth column is a histogram of the pixel-by-pixel difference between the true and predicted RP field, which has been fitted to a Gaussian distribution (black line). The true and predicted J_{ex} are also given, along with the normalised MSE.

Figure 8.12 showcases example predictions for 'clean' and 'dirty' systems at $J_{ex} \in \{0.4, 0.6\}$ from the test set using the optimised model. Due to the inclusion of these parameters in the training set, the model has learned to predict both the RP field and J_{ex} for different skyrmion sizes and we now no longer see the skyrmion imprint the predicted RP field. Once again, we see a high prediction accuracy for not only the localised RP field but also the exchange strength, however, as noted previously, this accuracy (normalised MSE) decreases for lower defect and J_{ex} levels.

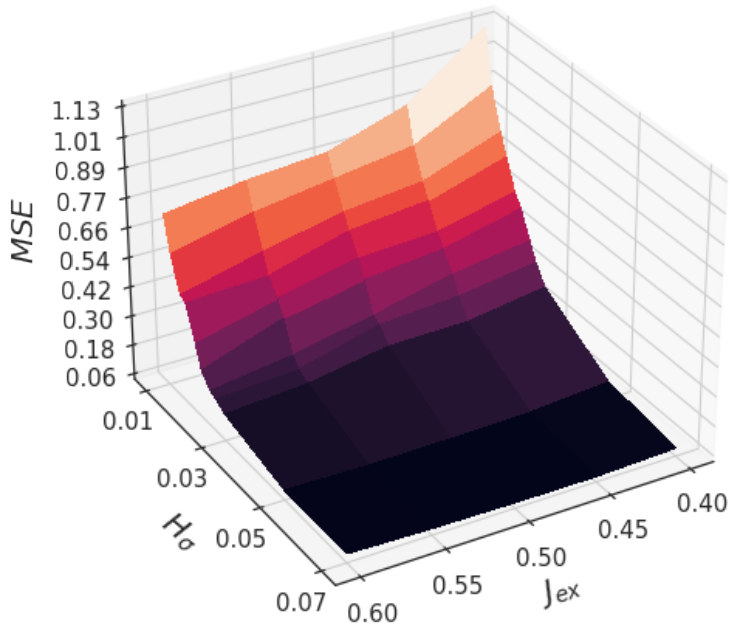

FIGURE 8.13: A surface plot of the averaged MSE as a function of H_σ and J_{ex}.

The behaviour of the normalised MSE, averaged over all datapoints and stratified by H_σ and J_{ex} (Figure 8.13), reveals a clear deterioration in performance with decreasing values of both parameters. As J_{ex} decreases, the skyrmions shrink in size allowing for more skyrmions to exist within the lattice. Consequently, smaller J_{ex} in 'cleaner' systems leads to an increased number of skyrmions, which are less sensitive to the defect field than uniform regions, thereby contributing to the increase in MSE. Nevertheless, even at the highest MSE levels, the average difference in the defect field at all lattice points is close to zero, indicating a relatively high prediction accuracy (Figure 8.12).

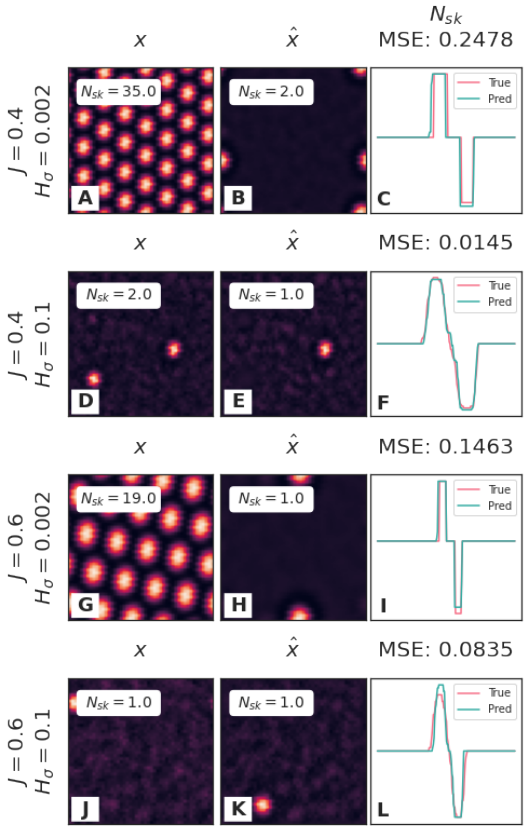

FIGURE 8.14: A comparison of the true switching configuration with the results using the predicted U-Net parameters, together with the N_{sk} hysteresis behaviour for various pairs of (J_{ex}, H_σ).

To further validate our findings, we feed the predicted RP field and J_{ex} as inputs for the mean field algorithm, and compare the resulting spin configuration and hysteresis behaviour with the original data. Figure 8.14 compares the true switching spin configuration with that generated from the predicted U-Net parameters, together with N_{sk} behaviour along the field sweep for different J_{ex} and H_σ. Although the predicted RP field fails to exactly reproduce the correct spin texture at the switching point, particularly in 'clean' systems, it generally mirrors the hysteresis behaviour seen in the true RP field case. Specifically, 'clean' systems, although producing only a few skyrmions at the nucleation boundary, eventually transitions to the full OSkL state at the subsequent field step with a constant N_{sk}. 'Dirty' systems, simulated by the predicted U-Net parameters, also follow the expected behavior from true parameters, exhibiting a smooth N_{sk} behaviour along the field sweep, with the nucleation boundary occurring at a lower field.

The discrepancies in the switching configuration likely stem from minute differences in the RP fields, where locations with slightly higher predicted defect values may nucleate individual skyrmions at earlier field steps, as opposed to the full lattice for 'clean' systems. Furthermore, we note that the skyrmions in the original and predicted configurations appear to be of similar sizes, and, according to N_{sk}, approximately a similar number of skyrmions nucleate in the full OSkL. This confirms that the predicted J_{ex} leads to topological particles similar to those observed with the true J_{ex}.

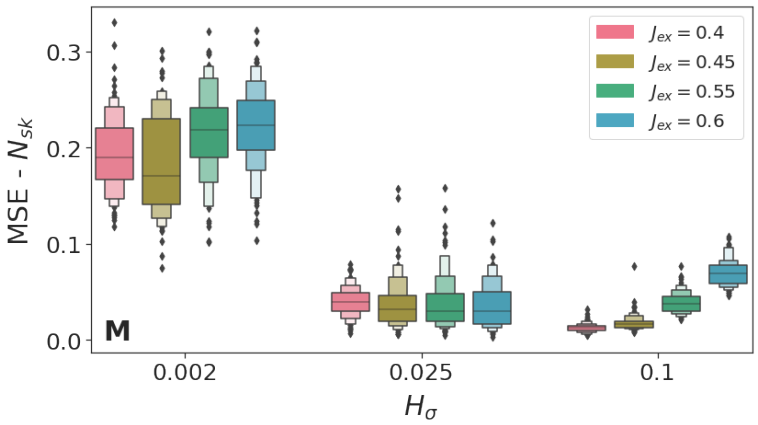

FIGURE 8.15: Boxplots of the scaled MSE on N_{sk} stratified by J_{ex} and H_σ.

To quantitatively compare the hysteresis behaviour in N_{sk} between the true parameters and U-Net predicted parameters, we compute the MSE of the N_{sk}, normalising by the area under the N_{sk} curve to enable comparison against a baseline prediction of $N_{sk} = 0$. Similar to the U-Net MSE (Figure 8.13), this measure also reveals a decline in performance at lower defect levels, but contrastingly also shows a decline in performance with increasing J_{ex}. This aligns with our understanding from Chapter 6, where larger skyrmions (higher J_{ex}) were found to be more sensitive to the underlying defect field, and therefore even a small percentage change can pin a skyrmion to a different location, leading to a change in the arrangement of skyrmions, and therefore a change in N_{sk}.

Nevertheless, assessing performance using MSE - N_{sk} provides a more reliable indicator than the U-Net MSE, as it evaluates the actual magnetic behaviour desired for simulation. Notably, this metric remains relatively low across all J_{ex} and H_σ, indicating that the U-Net demonstrates high accuracy in predicting both global and

local Hamiltonian parameters capable of reproducing mean field behaviour.

8.4 Learning parameters of polycrystalline materials (2L, 1G)

8.4.1 Model and Dataset

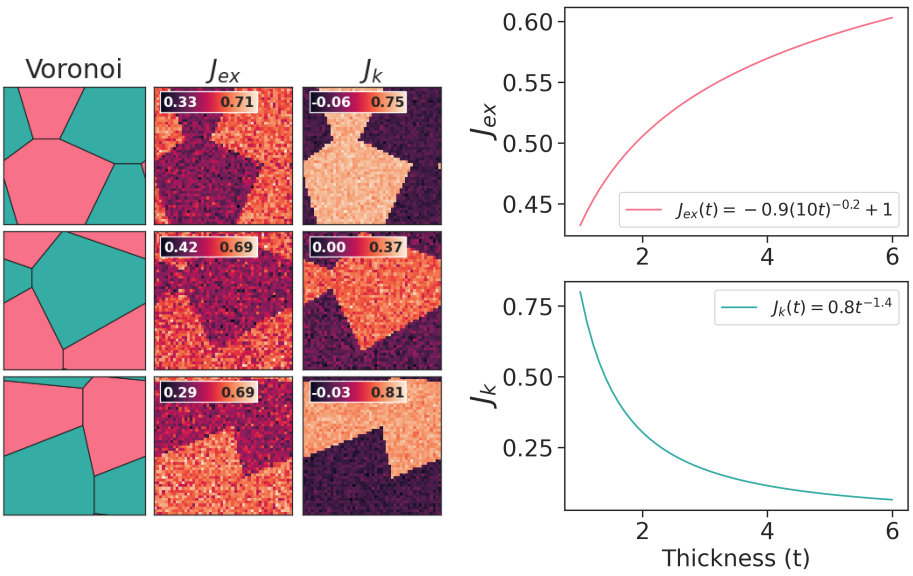

FIGURE 8.16: A) Examples of the Voronoi tessellation of the lattice thickness (first column), with the generated random-bond J_{ex} (second column) and the random anisotropy, J_k (third column). B) Thickness dependence of the average J_{ex} and J_k.

We increase the complexity of the extended U-Net by introducing a second localised parameter, (2L, 1G). To mimic real materials more closely, we replace the random pinning model with a random anisotropy model, accompanied by bond defects in the exchange strength. These constitute our two local parameters (J_k and J_{ex}), along with one global parameter, the DMI strength (J_{dm}).

Experimentally, thin film polycrystals contain regions of varying thickness with different anisotropy and isotropic exchange strengths. Many studies find that J_k tends to increase with decreasing film thickness due to heightened surface effects[269, 270], while conversely, exchange strength decreases owing to the reduced number of interactions along the sample thickness[271]. To simulate these polycrystals, we tessellate the lattice to vary its thickness throughout the structure. Using the `freud` package[272], we generate a Voronoi diagram with periodic boundary conditions, containing the regions for two different thicknesses (see Figure 8.16). Each cell

within this diagram is assigned a thickness value randomly sampled from the range $[1, 6]$. From these thickness values, we derive the average J_{ex} and J_k parameters for their respective regions, modeled with a power-law behavior illustrated in Figure 8.16B. While this behavior is not directly calibrated from empirical data, it is based on our experimental understanding of such systems, enabling us to evaluate the capabilities of the extended U-Net. Additionally, we introduce defects into the system through a pinning field on J_{ex} and J_k, drawn from $\mathcal{N}(0, \sigma)$, using a randomly generated defect level, $\sigma \in [0.001, 0.05]$. Figure 8.16A illustrates the Voronoi tessellation alongside the local J_{ex} and J_k fields.

Our dataset contains the switching spin configuration for 8000 realisations of the local parameters across 5 DMI strengths in the range $J_{dm} \in [0.4, 0.6]$, resulting in 40000 configurations. This dataset was divided into training, validation, and test sets, with proportions of 50%, 25%, and 25%, respectively. The outputs were min-max normalized, retaining their relative scales, following the approach in (1L, 1G) with the following minimum and maximum values,

$$\begin{aligned} p_{\min} &= \min\Big(\min(J_{ex}), \min(J_k), \min(J_{dm})\Big) \\ p_{\max} &= \max\Big(\max(J_{ex}), \max(J_k), \max(J_{dm})\Big) \end{aligned} \tag{8.6}$$

We train the same optimised architecture and hyperparameters as in (1L, 1G), resulting in a test set MSE of 0.02341.

8.4.2 Results

An example prediction from the test set is depicted in Figure 8.17. The figure shows the input spin texture as the first image, followed by the true and predicted exchange strength (J_{ex} and $\hat{J}_{ex}$) in the top row, and the true and predicted anisotropy (J_k and $\hat{J}_k$) on the bottom row. The true and predicted values of the global DMI strength (J_{dm} and $\hat{J}_{dm}$) are given in the title, along with the total MSE normalised against a baseline of 0.5.

In addition to examining site-specific values of the local parameters, we consider the distributions of J_{ex} and J_k (final column of 8.17) in our analysis. This is essential since precise variations in these parameters cannot be experimentally determined,

and instead only their averages can be roughly approximated across different thickness regions using the sample magnetisation[273]. The distributions of J_{ex} and J_k offer insights into several aspects of the sample, including the number of different thicknesses (modes), the average interaction strengths in these regions, and the defect level in the system (spread about each mode). We determine the average interaction strengths by fitting the true and predicted distributions to a mixture of two Gaussians, modelled as follows,

$$\rho(x, A, \mu, \sigma) = Ae^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \tag{8.7}$$

$$\text{Model}(x) = \rho(x, A_1, \mu_1, \sigma_1) + \rho(x, A_2, \mu_2, \sigma_2), \tag{8.8}$$

where A_i, μ_i, and σ_i represent the amplitude, mean, and standard deviations of distribution i.

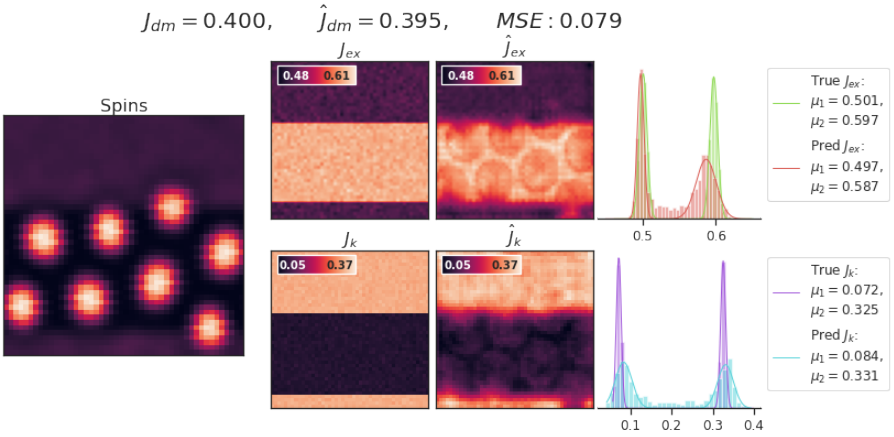

FIGURE 8.17: Two example predictions from the test set. The first image shows the input sp in configuration, the next two images are the true and predicted J_{ex} (top row) and J_k (bottom row). The final column is a histogram of the pixel-by-pixel difference between the true and predicted J_{ex} (top row) and J_k (bottom row).

The most notable outcome of the U-Net predictions is the presence of skyrmion imprints, particularly evident in $\hat{J}_{ex}$ and somewhat visible in $\hat{J}_k$. Similar imprints were only observed in other models when tested with unseen skyrmion sizes, such as the (1L) model trained with $J_{ex} = 0.5$ but evaluated at different J_{ex} (see Figures (Figure 8.9 and 8.10)). We suspect a similar issue here, where the relatively smaller training dataset in comparison to the number of free Hamiltonian parameters may result in uncommon skyrmion sizes caused by a specific set of parameters and defect

level. Consequently, these configurations are not adequately learned by the model, leading to lower prediction accuracy. This decline in performance is further evident in the histograms for the true and predicted J_{ex}, where the inaccuracies result in a wider distribution in $\hat{J}_{ex}$ about the specific mode corresponding to the skyrmion imprints, and an underestimation of this average exchange strength.

To further explore the degradation of model performance with number of free Hamiltonian parameters, a comparative study was conducted on all U-Net variations using a randomized sample of 18,000 examples for training and 9,000 samples each for validation and testing. To ensure comparability, the model architectures are adjusted to roughly contain a similar number of model weights[6], and were trained for 1000 epochs. Finally, the test set MSE is normalised by the number of weights and number of output parameters. The results, summarised in Table 8.1, clearly show a decrease in model performance as the number of Hamiltonian parameters increases, implying that a progressively larger dataset is needed to maintain similar performance levels with increasing model complexity.

U-Net Model Comparison

Model [# Input Parameters]	# Model Weights	U-Net MSE	Normalised MSE
(1L)[2304]	1, 592, 451	0.0157523	4.6957×10^{-12}
(1L, 1G)[2305]	1, 538, 373	0.0352776	9.9487×10^{-12}
(2L, 1G)[4609]	1, 456, 001	0.2229303	3.1242×10^{-11}

TABLE 8.1: Performance comparison between the different U-Net variants with an increasing number of input Hamiltonian parameters.

Nevertheless, despite the smaller dataset, the U-Net retains its capacity to accurately identify both the localised and global parameters, inclusive of their relative scale, even within a material composed of varying thicknesses. In the next section, we benchmark this model against simpler networks to identify the specific components that contribute to parameter estimation.

[6]Exact matching of the number of weights was not feasible due to the requirement of CNN filter sizes to be in powers of 2 for optimised training on a GPU.

8.5 Benchmark Models

8.5.1 Comparing U-Net to other networks

We evaluate the performance of our extended U-Net in comparison to other available models through several benchmark tests. Initially, we examine the simplest scenario involving a single localised parameter (1L) and compare our standard U-Net with alternative models, including a neural network solely comprising `Dense` layers (DenseNet), a fully convolutional network (FCN) lacking the U-Net's symmetric architecture and skip connections[7], and a U-Net variant without skip connections (U-Net-NoSkip).

Given the complexity of these alternative architectures, they may struggle to progress beyond predicting a constant normalised value of 0.5 for all RP field values, due to the vanishing gradient problem. However, since 0.5 is a relatively accurate (normalised) prediction for systems with low defect levels, this issue is not reflected in the MSE. To address this, we also compute the ratio of MSE of the true RP fields relative to 0.5, against the MSE of the predicted RP fields relative to 0.5.

$$\text{Score} = \frac{\sum_{ij}(\boldsymbol{Y}_{ij} - 0.5)^2}{\sum_{ij}(\hat{\boldsymbol{Y}}_{ij} - 0.5)^2} \tag{8.9}$$

This score tends to 1 when the difference between the true and predicted RP fields is zero, but results in a larger value if predictions, $\hat{\boldsymbol{Y}}_{ij}$, are close to 0.5.

The results of the benchmark tests are given in Tables 8.2-8.4 for both MSE and the Score, with the final row showing the results of the U-Net used in Section 8.2. In FCN, the value enclosed in brackets denotes the stride of the convolutional layer, and for the U-Net-NoSkip, a 2×2 `MaxPool` follows every convolutional layer except the last.

The score indicator demonstrates that FCN and U-Net-NoSkip significantly outperform DenseNet, indicating the necessity of convolutional layers for effectively learning localised parameters. The DenseNet suffers from the vanishing gradient problem, where the model's depth and complexity cause gradients propagated to earlier layers to become exceedingly small, resulting in minimal adjustments to these

[7] For more information, refer to Chapter 3, Section 3.4.

DenseNet

Architecture	MSE	Score
(512, 256, 128, 64, 32)	0.333539	10179.3993
(128, 64, 128)	0.333983	4788.7693
(128, 64)	0.333952	2727.7160
(64, 32)	0.333646	5323.1361
U-Net	**0.004180**	**1.11794**

TABLE 8.2: Performance comparison for various DenseNet architectures using MSE and Score, against the standard U-Net network.

FCN

Architecture	MSE	Score
(32[1], 64[1], 128[2], 256[2])	0.313781	14.0541
(32[1], 64[1], 128[2])	0.256848	4.332644
(16[1], 32[2], 64[2])	0.318490	22.0412
(16[1], 32[1], 64[2])	0.259481	4.5367
(16[2], 32[2])	0.330277	355.5465
U-Net	**0.004180**	**1.11794**

TABLE 8.3: Performance comparison for various FCN architectures using MSE and Score, against the standard U-Net network.

U-Net-NoSkip

Architecture	MSE	Score
(64, 128, 256, 512, 1024)	0.330219	350.7250
(32, 64, 128, 256, 512)	0.330091	367.3465
(16, 32, 64, 128, 256)	0.329793	364.0035
(32, 64, 128, 256)	0.329848	389.0372
(32, 64, 128)	0.330202	364.1737
(16, 32, 64)	0.330218	312.7810
(16, 32)	0.187955	3.8886
(32, 64)	0.136675	2.8746
U-Net	**0.004180**	**1.11794**

TABLE 8.4: Performance comparison for various U-Net architectures with no skip connections using MSE and Score, against the standard U-Net network.

layers' weights. Additionally, the early stopping criterion terminates training when weight changes are minimal, effectively locking weights to their initial random values and leading to RP field predictions of 0.5. In contrast, convolutional layers reduce the number of trainable parameters, thereby mitigating the vanishing gradient issue.

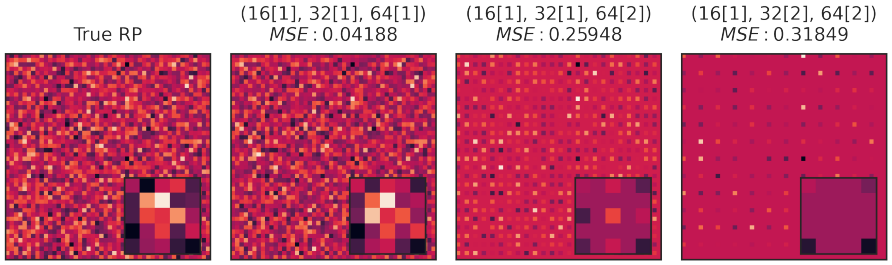

FIGURE 8.18: **FCN example predictions**: A comparison between the true RP field (A), and the predictions FCN architectures with an increasing level of downsampling: (16[1], 32[1], 64[1]), (16[1], 32[1], 64[2]), (16[1], 32[2], 64[2]).

While convolutional layers are essential for learning localised parameters, the consecutive downsampling via strided convolution followed by up-convolution in the final layer of the FCN often introduces checkerboard artifacts in the predicted RP field (refer to Figure 8.18). This figure displays an example random pinning field alongside predictions from three different FCN networks with increasing levels of downsampling: (16[1], 32[1], 64[1]), (16[1], 32[1], 64[2]), and (16[1], 32[2], 64[2]). The network without downsampling, (16[1], 32[1], 64[1]), employs a final convolutional layer with just one filter instead of an up-convolution. This setup preserves spatial information, yielding high prediction accuracy across all lattice sites. However, the absence of downsampling results in a higher number of trainable weights, which slows down convergence.

Introducing downsampling with a stride-2 convolutional layer accelerates training, but it leads to checkerboard artifacts in the predicted RP field, with accurate predictions occurring only at every other lattice site. These artifacts stem from the striding of the up-convolution, causing uneven overlap of the kernel over lattice sites, resulting in better predictions at frequently visited sites and predictions closer to the average (0.5) at less frequently visited sites[274]. With two convolutional layers using stride 2 (final column), the final up-convolution uses a stride of 4, resulting in sparser overlapped lattice sites and more inaccuracies. To summarise, the inclusion of convolutional layers in the FCN somewhat resolves the vanishing gradient problem of DenseNets, leading to accurate predictions at certain lattice sites and, consequently yields a higher Score (Table 8.3). But the loss of information from downsampling prevents pixel-perfect predictions of the localised parameter.

The checkerboard artifacts of FCN can be resolved through the symmetric downsampling and upsampling of the U-Net architecture. Figure 8.19 illustrates an example prediction from two different U-Net architectures ((16, 32) and (16, 32, 64)), where the left-most images depict the input spin texture and the target (true) RP field, followed by the predictions on the right. Here, the figure shows the predicted RP field from (16, 32) (top row), and (16, 32, 64) (bottom row), in a network without (first column) and with (second column) skip connections.

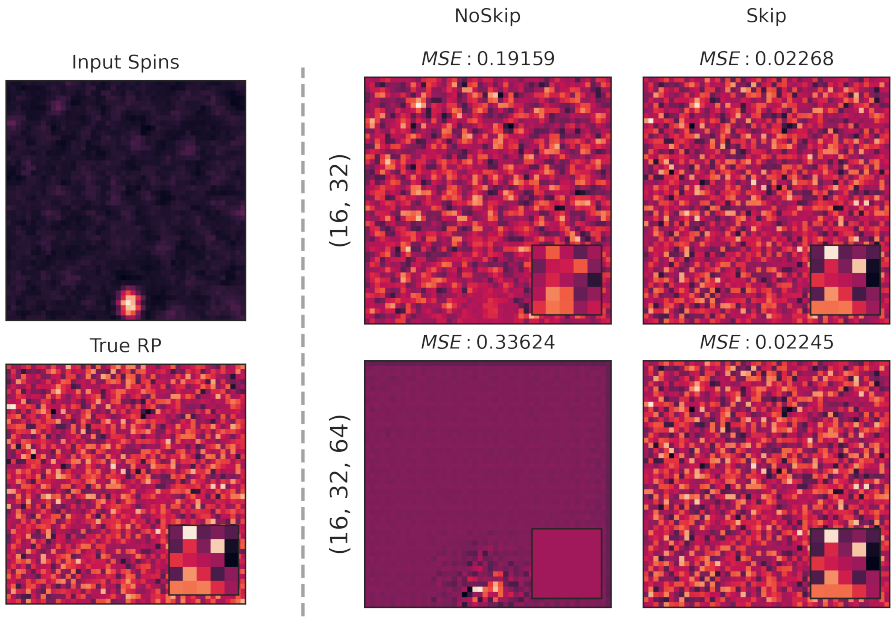

FIGURE 8.19: **U-Net example predictions**: The left-hand images shows the input spin configuration(top) and true RP field (bottom), whereas the right-hand side displays the predictions from networks of size (16, 32) in the first row and (16, 32, 64) in the second row, without (first column) and with (second column) skip connections.

A relatively small architecture of (16, 32), featuring a single `MaxPool` and no skip connections, exhibits enhanced performance compared to the FCN, with a test set MSE of 0.1910996. Visually, we no longer observe the presence of checkerboard artifacts, likely because upsampling occurs in successive steps with smaller stride as opposed to the single step in FCN, with only one learnable filter. This leads to an increased filter overlap over all lattice sites, resulting in more accurate predictions. However, the absence of skip connections leads to a deterioration in prediction quality, due to information loss from the `MaxPool`. The filters of the convolutional layers learn specific input features, such as circular structures, but the downsampling obscures the exact feature location, resulting in minor inaccuracies in the predicted

field.

With an increase in network complexity to (16, 32, 64), the model fails to predict the local RP field, further exhibiting a sharp imprint of the original skyrmion from the input. This decline in model performance can be attributed to the two `MaxPool` layers, which increases the level of information loss. Skip connections are, therefore, essential for learning the localised parameters of a complex model, owing to their ability to transfer detailed information learned in the encoder path to the upsampling layer, bypassing the `MaxPool`.

The benchmark tests performed thus far underscores the importance of convolutional layers, a symmetric network, and max-pooling layers for efficiently learning a localised parameter. In the next section, we compare the performance of the extended U-Net, which predicts both local and global parameters, with a fully U-Net architecture excluding the extended branch.

8.5.2 Comparing a full U-Net to the extended U-Net

Performance Comparison between Full and Extended U-Net

Model	# Output Parameters	Norm. MSE	Norm. MSE - J_{ex}	Norm. MSE - J_k	Norm. MSE - J_{dm}
Extended	4609	0.112897	0.0044719	0.1074786	0.0009469
Full	6912	0.108245	0.0034053	0.1024386	0.0024015

TABLE 8.5: A comparison of the normalised test set MSE between the full and extended U-Net, across all parameters (**Norm. MSE**), and its individual components, (**Norm. MSE -** J_{ex}/J_k/J_{dm}).

The final benchmark test compares a full U-Net, which incorporates the global parameter as an additional channel resulting in site-dependent predictions, with the extended U-Net that provides a single prediction for the parameter. Both models are trained using the dataset from (2L, 1G), and their prediction accuracy for each parameter is evaluated using MSE normalised against a baseline of 0.5 and by the number of model parameters. Given its increased complexity, the full U-Net takes longer to converge, requiring 1627 epochs compared to 438 epochs for the extended U-Net. However, despite the extended U-Net's faster training, it demonstrates a similar performance to the full U-Net when evaluated using the total MSE across all parameters of the test set (Table 8.5): the extended U-Net has an MSE of 0.112897, while

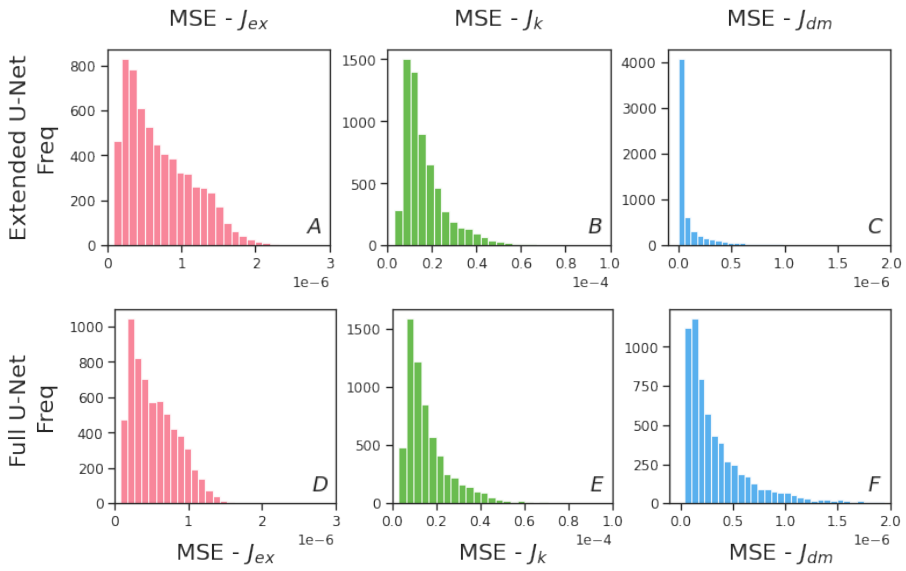

FIGURE 8.20: The distributions of normalised MSE for each parameter (J_{ex}, J_k, J_{dm}) in the extended U-Net (A-C) and the full U-Net (D-F).

the full U-Net has an MSE of 0.108245. Assessing the individual components of the normalised MSE for each output parameter (Figure 8.20 and Table 8.5), we observe a decline in accuracy of the J_{dm} parameter using the full U-Net, offset by a slightly higher accuracy in predictions for the localised parameters, J_{ex} and J_k. In summary, although both architectures are adept at parameter estimation, the extended variant achieves better accuracy for global parameters, with faster convergence.

8.6 Conclusions

In conclusion, we have demonstrated that the U-Net network can be successfully repurposed for local Hamiltonian parameter estimation, with a high accuracy at all lattice sites, and with an additional extension to the model, global parameters can also be simultaneously learned. However, we encountered a number of challenges in using fully convolutional networks, including a minor deterioration in prediction accuracy when estimating parameters from skyrmion structures, especially for unseen skyrmion sizes, and the requirement for a larger dataset as the number of learnable Hamiltonian parameters increases. Despite these challenges, local and global parameters can be accurately predicted even in materials of varying thicknesses, only achievable using our variant of the U-Net architecture. Specifically, the success of

the network can be attributed to the use of convolutional and max-pool layers, essential for improved convergence and overcoming the vanishing gradient problem, along with the symmetric nature of the architecture and skip connections, resulting in pixel-perfect predictions of local parameters. In future, we hope that such a model can be used to predict the Hamiltonian parameters directly from experimental LTEM data, after such systems are initially manually calibrated to the mean field Hamiltonian. Once calibrated, these neural networks can be used to pinpoint exact defect locations within a sample.

Chapter 9

Conclusions and Future Research

Within the scope of skyrmion research, experiments have shown material defects to have adverse and unpredictable effects, both on skyrmion dynamics and their equilibrium properties. We explored the ground states of magnetic systems at thermal equilibrium, using mean field approximation over a broad range of external fields, temperatures, and varying levels of defect. We demonstrated that the mean-field model is capable of reproducing experimental data of Fe_2GeTe_2 two-dimensional van der Waals magnet, which adds the credibility to the overall results reported in this thesis.

Instead of studying the role of individual defects, as has been previously reported in many studies, we set out to investigate the effect of entire statistical distribution of defects on the formation of skyrmion lattices. Thereby, the defects were here modelled as a quenched random pinning field, following Gaussian statistical distribution with zero-mean, varying only in z ($H_{z,i} \sim \mathcal{N}(0, H_\sigma)$). In the future, however, we aim to refine our approach by modelling defects as fluctuations in anisotropy, or DMI[20, 204, 209, 210, 213, 214, 218, 220–227], in accordance with ab-initio studies[228].

In a comprehensive study for 'clean' systems, we observed an instantaneous switch to a hexagonal skyrmion lattice, whereas 'dirty' systems exhibited a gradual nucleation of skyrmions, forming a disordered array and blurring the first-order phase transition. After further analysis, we hypothesised the existence of a phase transition from the OSkL to DSkL, driven by the material defects. By mapping the spin-spin correlation (Q) of configurations along the field sweep, we identified the

transition boundary and found that larger skyrmions demonstrate heightened sensitivity to defects, with a lower critical defect level needed to disrupt the OSkL.

Through our calculations of Q, we distinguished OSkL and DSkL phases, however our values were sensitive to minute changes in spin texture and required further post-processing for clear phase differentiation, where the specific post-processing method differs depending on phase transition. For example, in determining $H_{\sigma,crit}$, we found the skyrmion nucleation boundary and corresponding spin-spin correlation across all defect levels. Then, we pinpointed the defect level with the greatest variability in Q and extrapolated to an infinite-sized lattice. In contrast, through deep learning (DL) methods we replaced the manual approach with backpropagation to identify macroscopic properties for phase classification.

Our study used variational autoencoders (VAE) to identify latent variables used to differentiate the majority of phases found in materials with defects. Of these phases, only OSKL and DSKL could not be differentiated using VAEs, primarily due to the use of convolutional layers, which resulted in a loss of long-range information. Consequently, our research highlights the necessity of careful network design. Layers that simplify the architecture could potentially discard crucial phase information, and an overly complex VAE network may struggle to learn the latent representation effectively.

However, the VAE parameters can differentiate OSkL from DSkL via the configuration at the skyrmion transition boundary, and therefore the critical H_σ can be determined. $H_{\sigma,crit}$ offers an approximate gauge for the permissible defect amount in the material while maintaining a tightly-packed skyrmion lattice, thereby promoting increased data density in prospective storage devices. Another crucial aspect underlying all devices, including the racetrack, logic gates and neuromorphic computing, is the ease and reliability of skyrmion transportation. We hypothesise that below $H_{\sigma,crit}$, skyrmions move along relatively predictable trajectories, shrinking and avoiding small defects when necessary. Conversely, in the presence of large defects above $H_{\sigma,crit}$, we expect significant deviation of skyrmion trajectories, even potentially causing skyrmion annihilation. However, the investigation into the dynamics of skyrmions in materials with defects is beyond the scope of our work and we leave this aspect to future studies.

Our final study involved solving the inverse problem of accurately identifying the underlying defect field solely from the spin configuration, achieved through training a U-Net-based network in a supervised manner. The combination of convolutional layers, max-pool layers, the symmetric network, and the skip connections are all crucial components of the U-Net, enabling accurate pixel-wise defect detection. To enhance the model's capabilities, we extended it with a secondary branch at the network bottleneck, enabling the prediction of an additional global parameter, such as lattice-wide DMI strength. This extended U-Net is capable of predicting any number of local and global Hamiltonian parameters, even in a polycrystalline material with regions of varying thicknesses, however larger datasets are required for training as the complexity increases. While our study focused on parameter estimation, the model's versatility allows for its potential extension to diverse problems requiring the prediction of both local and global properties, for example Lee et al. similarly repurposes U-Net to estimate the magnetic effective field from spin textures[275].

In conclusion, by studying statistical distributions of pinning sites, we observed significant changes in the phase boundaries, including a novel disorder-driven phase transition. We developed various methods to identify the different critical transitions, including the spin-spin correlation mapping and deep-learning techniques. Finally, due to impact of defects on skyrmion nucleation and dynamics in experimental systems, we solve the inverse problem of identifying defect locations using the U-Net model.

Appendix A

Appendix

A.1 Derivation of the CW-MF Model

In this chapter, we show the full derivation of the Curie-Weiss mean field (CW-MF) model for an Ising system.

The Ising model is given as,

$$\mathcal{H}_{ising} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\in i_{nn}} J_{ij}s_i s_j - \mu H\sum_{i}^{N} s_i, \tag{A.1}$$

with the following partition function,

$$Z_N = \mathrm{Tr}_{s_k} \exp\left[\beta\left(\frac{1}{2}J_{ij}s_i s_j + \mu H s_i\right)\right]. \tag{A.2}$$

The derivation of the Curie-Weiss model begins by decomposing the spin into the mean field, m_i, and the fluctuation about the mean field, $\delta s_i = (s_i - m_i)$,

$$s_i = m_i + \delta s_i. \tag{A.3}$$

Substituting this into the interaction term, gives,

$$s_i s_j = (m_i + \delta s_i)(m_j + \delta s_j) \tag{A.4}$$

$$\approx m_i m_j + m_i \delta s_j + m_j \delta s_i, \tag{A.5}$$

where, in the final line, we have neglected fluctuations about the mean field. Using the fact that the mean field is site-independent (due to the translational invariance of the model), and that $\sum_i m(s_i - m) = \sum_j m(s_j - m)$, the Ising spin Hamiltonian thus

reads,

$$\mathcal{H}(s) = -\frac{1}{2}\sum_{i}^{N}\sum_{j}^{N} J_{ij}(m^2 + 2m(s_i - m)) - \mu H \sum_{i} s_i \tag{A.6}$$

$$= \frac{1}{2}m^2 \sum_{i}^{N}\sum_{j}^{N} J_{ij} - Jm\sum_{i}^{N}\sum_{j}^{N} s_i - \mu H \sum_{i} s_i, \tag{A.7}$$

$$= \frac{1}{2}m^2 JzN - (Jmz + \mu H)\sum_{i}^{N} s_i, \tag{A.8}$$

where we have replaced $\sum_j$ by z, the coordination number. The partition function can therefore be written as,

$$Z_N = \mathrm{Tr}_{s_k}\left(e^{-\beta\mathcal{H}(s)}\right) = e^{\frac{\beta J m^2 zN}{2}}\mathrm{Tr}_{s_k}\left(\exp\left[\beta(Jmz + \mu H)s_i\right]\right), \tag{A.9}$$

with the notation simplified using Einstein summation. The exponential term in the trace is linear in s_i, so this expression is simply a sum over non-interacting spins,

$$Z_N = e^{\frac{\beta J m^2 zN}{2}} \sum_{s_1=\pm 1} ... \sum_{s_N=\pm 1} \left(\exp\left[\beta(Jmz + \mu H)s_i\right]\right) \tag{A.10}$$

$$= e^{\frac{\beta J m^2 zN}{2}} \left(\sum_{s_1=\pm 1} \exp\left[\beta(Jmz + \mu H)s_1\right]\right)^N \tag{A.11}$$

$$= e^{\frac{\beta J m^2 zN}{2}} \left(2\cosh\left(\beta(Jmz + \mu H)\right)\right)^N. \tag{A.12}$$

The free energy in the thermodynamic limit is,

$$F = -\lim_{N\to\infty}\frac{1}{\beta N}\ln Z_N = \frac{1}{2}Jzm^2 - \frac{1}{\beta}\ln\left[2\cosh\left(\beta(Jzm + \mu H)\right)\right], \tag{A.13}$$

and magnetisation is therefore,

$$m = -\frac{1}{\mu}\frac{\partial F}{\partial H} = \tanh\left(\beta(Jzm + \mu H)\right). \tag{A.14}$$

This is the self-consistency equation for magnetization.

A.2 Derivation of HS-MF Model

In Chapter 2, Section 2.4.2, we briefly outlined the steps taken to derive the mean field model using the Hubbard-Stratonovich identity. In this section, we show the full derivation for both the Ising model and the Heisenberg model.

A.2.1 Ising Model

We will first consider the Ising model given below,

$$\mathcal{H}_{ising} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\in i_{nn}} J_{ij}s_i s_j - \mu H\sum_{i}^{N} s_i. \tag{A.15}$$

The partition function for this model is,

$$Z_N = \mathrm{Tr}_{s_k}\exp\left[\beta\left(\frac{1}{2}J_{ij}s_i s_j + \mu H s_i\right)\right]. \tag{A.16}$$

The Hubbard-Stratonovich transformation explained in Section 2.4.2, is rewritten here for convenience,

$$\frac{1}{\sqrt{|A|}}e^{\frac{1}{2}\boldsymbol{b}_i(A^{-1})_{ij}\boldsymbol{b}_j} = \int_{-\infty}^{\infty}\prod_{i}^{N}\frac{\exp\left[-\frac{1}{2}x_i A_{ij}x_j + x_i\boldsymbol{b}_i\right]}{\sqrt{2\pi}}dx_i. \tag{A.17}$$

The HS transformation seeks to simplify the exponential term containing the bi-linear isotropic exchange interaction,

$$e^{\frac{1}{2}J_{ij}s_i s_j} = \frac{1}{\sqrt{|J|}}\int_{-\infty}^{\infty}\prod_{i}^{N}\frac{\exp\left[-\frac{1}{2}\psi_i(J^{-1})_{ij}\psi_j + \psi_i s_i\right]}{\sqrt{2\pi}}d\psi_i, \tag{A.18}$$

where we have used to following replacements,

$$\boldsymbol{b}_i \coloneqq s_i, \tag{A.19}$$

$$(A^{-1})_{ij} \coloneqq J_{ij}, \tag{A.20}$$

$$x_i \coloneqq \psi_i. \tag{A.21}$$

Substituting the exponential back into the partition function, we find,

$$Z_N = \frac{1}{\sqrt{(2\pi)^N |J|}} \int_{-\infty}^{\infty} \prod_i^N e^{-\frac{1}{2}\psi_i(\beta(J^{-1})_{ij})\psi_j} \mathrm{Tr}_{s_k} e^{\beta(\psi_i+\mu H)s_i)} d\psi_i. \tag{A.22}$$

The exponential has been made linear in terms of s_i and as such can be interpreted as a system of non-interacting spins. Therefore, this can be simplified,

$$\mathrm{Tr}_{s_k} e^{\beta(\psi_i+\mu H)s_i)} = \sum_{s_1=\pm 1} ... \sum_{s_N=\pm 1} e^{\beta(\psi_i+\mu H)s_i} \tag{A.23}$$

$$= \prod_i^N \sum_{s=\pm 1} e^{\beta(\psi_i+\mu H)s} \tag{A.24}$$

$$= \prod_i^N 2\cosh\left[\beta\left(\psi_i + \mu H\right)\right]. \tag{A.25}$$

We would like to rewrite this trace in exponential form for future simplifications. So using the fact that $x = e^{\ln x}$ and $\ln(xy) = \ln x + \ln y$, we see that,

$$\mathrm{Tr}_{s_k} e^{\beta(\psi_i+\mu H)s_i)} = \exp\left[\ln\left(\prod_i^N 2\cosh\left[\beta\left(\psi_i + \mu H\right)\right]\right)\right] \tag{A.26}$$

$$= \exp\left[\sum_i^N \ln\left(2\cosh\left[\beta\left(\psi_i + \mu H\right)\right]\right)\right]. \tag{A.27}$$

Thus we can write partition function as,

$$Z_N = \frac{1}{\sqrt{(2\pi)^N |J|}} \int_{-\infty}^{\infty} \prod_i^N e^{-\beta\mathcal{L}(\psi_i, J_{ij}, \mu H)} d\psi_i, \tag{A.28}$$

where $\mathcal{L}(\psi_i, J_{ij}, \mu H)$ is the action,

$$\mathcal{L}(\psi_i, J_{ij}, \mu H) = \frac{1}{2}\psi_i (J^{-1})_{ij} \psi_j - \frac{1}{\beta}\sum_i^N \ln\left(2\cosh\left[\beta\left(\psi_i + \mu H\right)\right]\right). \tag{A.29}$$

This was the result derived in the main thesis. This is a functional integral and the partition function is still describes the Ising model in an exact form. We will now apply the saddle point approximation, which assumes that the integrand can be approximated by its maximal value.

$$\int_{-\infty}^{\infty} \prod_i^N e^{-\beta\mathcal{L}(\psi_i, J_{ij}, \mu H)} d\psi_i \approx e^{-\beta\mathcal{L}(\bar{\psi}_i, J_{ij}, \mu H)} \tag{A.30}$$

where $\bar{\psi}_i$ is the value of the auxiliary field at the maximum of the FI. The partition function is now,

$$Z_N = \frac{1}{\sqrt{(2\pi)^N |J|}} e^{-\beta \mathcal{L}(\bar{\psi}_i, J_{ij}, \mu H)}. \tag{A.31}$$

We can calculate the value of $\bar{\psi}_i$ by maximising the exponential in terms of ψ_i, i.e. by minimising the action, $\mathcal{L}$,

$$\frac{\partial \mathcal{L}}{\partial \psi_i} = (J^{-1})_{ij}\psi_j - \sum_i \tanh\left(\beta(\psi_i + \mu H)\right) = 0. \tag{A.32}$$

With some rearranging and relabeling, the auxiliary field is given by,

$$\bar{\psi}_i = \sum_j J_{ij} \tanh\left(\beta(\bar{\psi}_j + \mu H)\right). \tag{A.33}$$

We would now like to calculate the local magnetisation at site, i, as a function of this auxiliary field. We know that magnetization can be found from the free energy by,

$$m_i = -\frac{\partial F}{\partial H} \tag{A.34}$$

Using the partition function, we can find the relationship between the free energy and the saddle point approximation,

$$Z = e^{-\beta \mathcal{F}} \approx e^{-\beta \mathcal{L}(\bar{\psi}_i, J_{ij}, \mu H)}, \tag{A.35}$$

$$\mathcal{F} = \mathcal{L}(\bar{\psi}_i, J_{ij}, \mu H). \tag{A.36}$$

Therefore, magnetization is calculated as,

$$m_i = -\frac{\partial}{\partial H}\mathcal{L}(\bar{\psi}_i, J_{ij}, \mu H) \tag{A.37}$$

$$= -\frac{\partial}{\partial H}\left(\frac{1}{2}\bar{\psi}_i (J^{-1})_{ij}\bar{\psi}_j - \frac{1}{\beta}\sum_i^N \ln\left(2\cosh\left[\beta\left(\bar{\psi}_i + \mu H\right)\right]\right)\right) \tag{A.38}$$

$$= \sum_i^N \tanh\left(\beta(\bar{\psi}_i + \mu H)\right). \tag{A.39}$$

We can rearrange this to find another equation for $\bar{\psi}_i$,

$$\bar{\psi}_i = \frac{1}{\beta}\sum_i^N \tanh^{-1} m_i - \mu H. \tag{A.40}$$

Dependence on $\bar{\psi}_i$ can be eliminated by equating Equation A.33 to Equation A.40, and substituting A.40 into the remaining ψ parameter.

$$\sum_j J_{ij} \tanh\left(\beta(\bar{\psi}_j + \mu H)\right) = \frac{1}{\beta}\sum_i^N \tanh^{-1} m_i - \mu H \tag{A.41}$$

$$\sum_j J_{ij} \tanh\left(\tanh^{-1} m_j\right) = \frac{1}{\beta}\sum_i^N \tanh^{-1} m_i - \mu H \tag{A.42}$$

$$\sum_j J_{ij} m_j = \frac{1}{\beta}\sum_i^N \tanh^{-1} m_i - \mu H \tag{A.43}$$

$$m_i = \sum_i^N \tanh\left(\beta(\mu H + \sum_j^N J_{ij} m_j)\right) \tag{A.44}$$

This is the self-consistency equation for magnetization derived via the Hubbard-Stratonovich method.

A.2.2 Heisenberg Model

Finding the Function Integral Partition Function

We will now apply the same transformation to the Heisenberg model, which is given below,

$$\mathcal{H}_{\text{heisenberg}} = -\frac{1}{2}\sum_i^N \sum_{j\in i_{nn}} \left[J_{ij}(\boldsymbol{s}_i \cdot \boldsymbol{s}_j) + \boldsymbol{D}_{ij} \cdot (\boldsymbol{s}_i \times \boldsymbol{s}_j)\right] - \mu \sum_i^N \boldsymbol{H}_i \cdot \boldsymbol{s}_i. \tag{A.45}$$

The isotropic exchange strength, J_{ij} and antisymmetric Dzyaloshinskii-Moriya vector, $\boldsymbol{D}_{ij}$ can be combined into a general matrix,

$$\mathbb{I}_{ij} = \begin{bmatrix} J_{ij}^{xx} & D_{ij,z}^{xy} & -D_{ij,y}^{xz} \\ -D_{ij,z}^{yx} & J_{ij}^{yy} & D_{ij,x}^{yz} \\ D_{ij,y}^{zx} & -D_{ij,x}^{zy} & J_{ij}^{zz} \end{bmatrix}, \tag{A.46}$$

This matrix is non-zero only for nearest neighbour interactions, and is zero otherwise. It is also evidently both symmetric and invertible[1]. The Hamiltonian is now

[1] These are the requirements needed in order to apply the HS transformation.

given as,

$$\mathcal{H}_{\text{heisenberg}} = -\frac{1}{2}(\boldsymbol{s}_i \cdot \mathbb{I}_{ij} \cdot \boldsymbol{s}_j) - \mu \boldsymbol{H}_i \cdot \boldsymbol{s}_i, \tag{A.47}$$

in which Einstein summation has been used to simplify the notation. The partition function for this model is,

$$Z_N = \text{Tr}_{s_k} \exp\left[\beta\left(\frac{1}{2}(\boldsymbol{s}_i \cdot \mathbb{I}_{ij} \cdot \boldsymbol{s}_j) - \mu \boldsymbol{H}_i \cdot \boldsymbol{s}_i\right)\right], \tag{A.48}$$

where the trace operation is a sum over all possible spin configurations, s_k. Applying the same HS identity from Equation A.17 on the exchange exponential term gives,

$$e^{\frac{1}{2}\boldsymbol{s}_i \cdot \beta \mathbb{I}_{ij} \cdot \boldsymbol{s}_j} = \frac{1}{\sqrt{|\beta\mathbb{I}|}} \int_{-\infty}^{\infty} \prod_i^N \frac{\exp\left[-\frac{1}{2}\boldsymbol{\psi}_i \cdot (\beta^{-1}\mathbb{I}_{ij}^{-1}) \cdot \boldsymbol{\psi}_j + \boldsymbol{\psi}_i \cdot \boldsymbol{s}_i\right]}{\sqrt{2\pi}} d\boldsymbol{\psi}_i, \tag{A.49}$$

where $\boldsymbol{\psi}_i$ are the auxiliary fields. With this, the partition function now reads[2],

$$Z_N = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{|\beta\mathbb{I}|}} \int_{-\infty}^{\infty} \prod_i^N e^{-\frac{1}{2}\boldsymbol{\psi}_i \cdot (\beta^{-1}\mathbb{I}_{ij}^{-1}) \cdot \boldsymbol{\psi}_j} \text{Tr}_{s_k} e^{(\boldsymbol{\psi}_i + \mu\beta\boldsymbol{H}_i) \cdot \boldsymbol{s}_i} d\boldsymbol{\psi}_i, \tag{A.50}$$

$$= \frac{1}{\sqrt{|\beta\mathbb{I}|}} \int_{-\infty}^{\infty} \prod_i^N e^{-\frac{1}{2}\boldsymbol{\phi}_i \cdot (\beta\mathbb{I}_{ij}^{-1}) \cdot \boldsymbol{\phi}_j} \text{Tr}_{s_k} e^{(\beta(\boldsymbol{\phi}_i + \mu\boldsymbol{H}_i)) \cdot \boldsymbol{s}_i} \frac{\beta d\boldsymbol{\phi}_i}{(2\pi)^{\frac{3}{2}}}, \tag{A.51}$$

$$= \frac{\beta^{\frac{3}{2}}}{\sqrt{|\mathbb{I}|}} \int_{-\infty}^{\infty} e^{-\beta\left(\frac{1}{2}\boldsymbol{\phi}_i \cdot (\mathbb{I}_{ij}^{-1}) \cdot \boldsymbol{\phi}_j - \frac{1}{\beta}\ln \text{Tr}_{s_k} exp(\beta(\boldsymbol{\phi}_i + \mu\boldsymbol{H}_i)) \cdot \boldsymbol{s}_i\right)} d\Omega, \tag{A.52}$$

$$= \frac{\beta^{\frac{3}{2}}}{\sqrt{|\mathbb{I}|}} \int_{-\infty}^{\infty} e^{-\beta\mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \boldsymbol{H}_i)} d\Omega. \tag{A.53}$$

In the second line, we made a substitution, $\boldsymbol{\phi}_i = \frac{\boldsymbol{\psi}_i}{\beta}$, then the third line simplifies the integral notation using a volume element, $d\Omega = (2\pi)^{-\frac{3}{2}} \prod_i^N d\boldsymbol{\phi}_i$, and evaluates the determinant, $|\beta\mathbb{I}|^{-\frac{1}{2}} = (|\beta I||\mathbb{I}|)^{-\frac{1}{2}} = \beta^{-\frac{3}{2}}|\mathbb{I}|^{-\frac{1}{2}}$. Here, I is the identity matrix and we assume that $\mathbb{I}$ is a 3×3 matrix. The initial substitution of $\boldsymbol{\phi}_i$ yielded an integration element $\prod_i^N \beta d\boldsymbol{\phi}_i = \beta^3 \prod_i^N d\boldsymbol{\phi}_i$, since there are 3 conjugate variables, $\boldsymbol{\phi}_i$, to match the 3 spin components. The constant factor thus simplifies to $\beta^{\frac{3}{2}}|\mathbb{I}|^{-\frac{1}{2}}$.

[2] In the Ising model, spins are constrained to take discrete values, $s_i \in \{-1, 1\}$, and thus a trace over non-interacting spins can be simplified further using the cosh identity. The Heisenberg model, however, has no such constraints and this results in a more complex partition function.

We have thus derived the function integral form of the partition function, and the action is given by,

$$\mathcal{L}(\boldsymbol{\phi}_i, \mathbb{I}_{ij}, \mu H) = \frac{1}{2}\boldsymbol{\phi}_i \cdot \mathbb{I}_{ij}^{-1} \cdot \boldsymbol{\phi}_j - \frac{1}{\beta} \ln \left[\mathrm{Tr}_{s_k} e^{\beta(\boldsymbol{\phi}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i} \right]. \tag{A.54}$$

Equations A.53 and A.54 can now be used to derive key properties of the magnetic system.

Finding magnetic moment and internal energy

From thermodynamic relations, the magnetic moment at site i can be found in terms of the partition function,

$$\boldsymbol{m}_i = -\frac{\partial \mathcal{F}}{\mu \partial \boldsymbol{H}_i} = \frac{1}{\mu\beta}\frac{1}{Z}\frac{\partial Z}{\partial \boldsymbol{H}_i}, \tag{A.55}$$

where $\mathcal{F}$ is the free energy given by, $\mathcal{F} = -\frac{1}{\beta}\ln Z$. Substituting our FI partition function (Equation A.53) into the above equation, we find that

$$\boldsymbol{m}_i = -\frac{1}{\mu}\frac{\int_{-\infty}^{\infty} d\Omega \frac{\partial \mathcal{L}}{\partial \boldsymbol{H}_i} e^{-\beta\mathcal{L}}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta\mathcal{L}}}. \tag{A.56}$$

Thus, the magnetic moment at i is completely specified by the action. The internal energy of the system can also be derived in similar fashion: from thermodynamics functions,

$$\mathcal{U} = -\frac{\partial \ln Z}{\partial \beta} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} \tag{A.57}$$

Once again, we substitute Equation A.53 into the above equation to get the following formula for internal energy,

$$\mathcal{U} = -\frac{3}{2\beta} + \frac{\int_{-\infty}^{\infty} d\Omega (\mathcal{L} + \beta \frac{\partial \mathcal{L}}{\partial \beta}) e^{-\beta\mathcal{L}}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta\mathcal{L}}} \tag{A.58}$$

Saddle Point Approximation

Having rewritten the partition function as an FI and derived key properties in terms of the action, we can now apply the saddle point approximation. As before, we

assume that the main contribution to $\bar{\boldsymbol{m}}_i$ and $\mathcal{U}$ is from the maximum of the exponential term, i.e. the minimum of the action, $\mathcal{L}$. The minimum of $\mathcal{L}$ in terms of the auxiliary variable is found by setting $\frac{\partial \mathcal{L}}{\partial \boldsymbol{\phi}_i} = 0$,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\phi}_i} = \mathbb{I}_{ij}^{-1}\boldsymbol{\phi}_j - \frac{1}{\beta}(\mathrm{Tr}_{s_k} e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i})^{-1}\beta \mathrm{Tr}_{s_k}\boldsymbol{s}_i e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i} = 0, \tag{A.59}$$

and therefore,

$$\bar{\boldsymbol{\phi}}_i = \sum_j \mathbb{I}_{ij}\frac{\mathrm{Tr}_{s_j}\boldsymbol{s}_j e^{\beta(\bar{\boldsymbol{\phi}}_j+\mu \boldsymbol{H}_j)\cdot \boldsymbol{s}_j}}{\mathrm{Tr}_{s_j} e^{\beta(\bar{\boldsymbol{\phi}}_j+\mu \boldsymbol{H}_j)\cdot \boldsymbol{s}_j}} \tag{A.60}$$

Thus the auxiliary field variables have been written as a set of non-linear equations coupled by the interaction matrix $\mathbb{I}_{ij}$. To find the magnetic moment in terms of these conjugate variables, we must first expand the action about the mean field solutions, $\boldsymbol{\phi}_i$,

$$\mathcal{L} = \bar{\mathcal{L}} + \frac{1}{2}\frac{\partial^2 \bar{\mathcal{L}}}{\partial \boldsymbol{\phi}_i \partial \boldsymbol{\phi}_j}\delta\boldsymbol{\phi}_i \delta \boldsymbol{\phi}_j + ... \tag{A.61}$$

Substituting this into our equation for magnetic moment, (Equation A.56), yields,

$$\bar{\boldsymbol{m}}_i = -\frac{1}{\mu}\frac{\int_{-\infty}^{\infty} d\Omega \left(\left(\frac{\partial \bar{\mathcal{L}}}{\partial \boldsymbol{H}_i}\right) + ...\right) e^{-\beta\bar{\mathcal{L}}+...}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta\bar{\mathcal{L}}+...}} \tag{A.62}$$

$$\approx -\frac{1}{\mu}\frac{\partial \bar{\mathcal{L}}}{\partial \boldsymbol{H}_i}\frac{\int_{-\infty}^{\infty} d\Omega e^{-\beta\bar{\mathcal{L}}+...}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta\bar{\mathcal{L}}+...}} = -\frac{1}{\mu}\frac{\partial \bar{\mathcal{L}}}{\partial \boldsymbol{H}_i} \tag{A.63}$$

The magnetic field derivative of the action gives,

$$\frac{\partial \bar{\mathcal{L}}}{\partial \boldsymbol{H}_i} = -\frac{1}{\beta}(\mathrm{Tr}_{s_k} e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i})^{-1}(\beta\mu)\mathrm{Tr}_{s_k}\boldsymbol{s}_i e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i} \tag{A.64}$$

$$= -\mu\frac{\mathrm{Tr}_{s_k}\boldsymbol{s}_i e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_k} e^{\beta(\boldsymbol{\phi}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}} \tag{A.65}$$

and therefore the magnetic moment is,

$$\bar{\boldsymbol{m}}_i = \frac{\mathrm{Tr}_{s_i}\boldsymbol{s}_i e^{\beta(\bar{\boldsymbol{\phi}}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_i} e^{\beta(\bar{\boldsymbol{\phi}}_i+\mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}. \tag{A.66}$$

In comparison to the equations over the field variables, (Equation A.60), we obtain the following relation,

$$\bar{\boldsymbol{\phi}}_i = \mathbb{I}_{ij}\bar{\boldsymbol{m}}_j \tag{A.67}$$

This can be used to conveniently eliminate dependence on the field variables in Equation A.66,

$$\bar{\boldsymbol{m}}_i = \frac{\mathrm{Tr}_{s_i} \boldsymbol{s}_i e^{\beta(\mathbb{I}_{ij}\bar{\boldsymbol{m}}_j + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_i} e^{\beta(\mathbb{I}_{ij}\bar{\boldsymbol{m}}_j + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}} \tag{A.68}$$

This is the set of self-consistent equations for magnetic moment, in a magnetic system with a more complex Hamiltonian. Performing a similar derivation on the internal energy, results in the mean-field Hamiltonian which describes the interactions of mean-field spins, $\bar{\boldsymbol{m}}_i$.

We start by inputting the Taylor expansion of the action into our derived formula for internal energy,

$$\mathcal{U} = -\frac{3}{2\beta} + \frac{\int_{-\infty}^{\infty} d\Omega (\bar{\mathcal{L}} + \beta \frac{\partial \bar{\mathcal{L}}}{\partial \beta}) e^{-\beta \bar{\mathcal{L}}}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta \bar{\mathcal{L}}}} \tag{A.69}$$

$$= -\frac{3}{2\beta} + (\bar{\mathcal{L}} + \beta \frac{\partial \bar{\mathcal{L}}}{\partial \beta}) \frac{\int_{-\infty}^{\infty} d\Omega e^{-\beta \bar{\mathcal{L}}}}{\int_{-\infty}^{\infty} d\Omega e^{-\beta \bar{\mathcal{L}}}} \tag{A.70}$$

$$\bar{\mathcal{U}} = \bar{\mathcal{L}} + \beta \frac{\partial \bar{\mathcal{L}}}{\partial \beta}, \tag{A.71}$$

where in the last line, the factor $-\frac{3}{2\beta}$ has been absorbed into $\bar{\mathcal{U}}$ since it is independent of the field variables. The beta-derivative of the action gives,

$$\frac{\partial \bar{\mathcal{L}}}{\partial \beta} = -\frac{1}{\beta} \frac{\mathrm{Tr}_{s_k} (\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i) \cdot \boldsymbol{s}_i e^{\beta(\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_k} e^{\beta(\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}} + \frac{1}{\beta^2} \ln \left[\mathrm{Tr}_{s_k} e^{\beta(\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i} \right] \tag{A.72}$$

and therefore internal energy is approximated by,

$$\bar{\mathcal{U}} = \frac{1}{2} \bar{\boldsymbol{\phi}}_i \cdot \mathbb{I}_{ij}^{-1} \cdot \bar{\boldsymbol{\phi}}_j - \frac{\mathrm{Tr}_{s_k} (\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i) \cdot \boldsymbol{s}_i e^{\beta(\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_k} e^{\beta(\bar{\boldsymbol{\phi}}_i + \mu \boldsymbol{H}_i)\cdot \boldsymbol{s}_i}} \tag{A.73}$$

$$= \frac{1}{2} \bar{\boldsymbol{m}}_i \cdot \mathbb{I}_{ij} \cdot \bar{\boldsymbol{m}}_j - \bar{\boldsymbol{m}}_i \cdot \mathbb{I}_{ij} \cdot \bar{\boldsymbol{m}}_j - \mu \boldsymbol{H}_i \cdot \bar{\boldsymbol{m}}_i \tag{A.74}$$

$$= -\frac{1}{2} \bar{\boldsymbol{m}}_i \cdot \mathbb{I}_{ij} \cdot \bar{\boldsymbol{m}}_j - \mu \boldsymbol{H}_i \cdot \bar{\boldsymbol{m}}_i \tag{A.75}$$

In the second line, we made the substitution, $\bar{\boldsymbol{\phi}}_i = \mathbb{I}_{ij} \bar{\boldsymbol{m}}_j$ and simplified the trace with Equation A.68. If we decompose the interaction matrix into its constituent parts, we find the mean field Hamiltonian over the mean field spin solution,

$$\bar{\mathcal{U}} \equiv \mathcal{H}_{MF} = -\frac{1}{2} \sum_{i}^{N} \sum_{j \in i_{nn}} \left[J_{ij} (\bar{\boldsymbol{m}}_i \cdot \bar{\boldsymbol{m}}_j) + \boldsymbol{D}_{ij} \cdot (\bar{\boldsymbol{m}}_i \times \bar{\boldsymbol{m}}_j) \right] - \mu \sum_{i}^{N} \boldsymbol{H}_i \cdot \bar{\boldsymbol{m}}_i \tag{A.76}$$

Differentiating this with respect to the mean field moment, we find

$$\mu \bar{\boldsymbol{H}}_i^e = \frac{\partial \mathcal{H}_{MF}}{\partial \bar{\boldsymbol{m}}_i} = J_{ij}\bar{\boldsymbol{m}}_j - \boldsymbol{D}_{ij} \times \bar{\boldsymbol{m}}_j - \mu \boldsymbol{H}_i. \tag{A.77}$$

This is the effective field term which acts on every spin, i, and is the argument of the exponentials in the mean field moment.

$$\bar{\boldsymbol{m}}_i = \frac{\mathrm{Tr}_{s_i} \boldsymbol{s}_i e^{\beta \mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}_i}}{\mathrm{Tr}_{s_i} e^{\beta \mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}_i}} \tag{A.78}$$

Thus we have derived a formula for the mean field spin in terms of a local effective field, and the algorithm for solving such an equation would be to simply calculate the effective field at point i, and then calculate the new moment from Equation A.78. Unfortunately calculating Equation A.78 would be resource intensive since this involves a trace over all the individual spins. Instead in the next section, we simplify this equation further using more mathematical tricks.

Iterative Algorithm

First, the sum over all spins can replaced with an integral, $\mathrm{Tr}_{s_i} \to \int d\boldsymbol{s}$

$$\bar{\boldsymbol{m}}_i = \frac{\int d\boldsymbol{s} \boldsymbol{s} e^{\beta \mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}}}{\int d\boldsymbol{s} e^{\beta \mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}}} \tag{A.79}$$

$$\mu \bar{\boldsymbol{H}}_i^e = J_{ij}\bar{\boldsymbol{m}}_j - \boldsymbol{D}_{ij} \times \bar{\boldsymbol{m}}_j - \mu \boldsymbol{H}_i \tag{A.80}$$

For further simplifications, we make the simple assumption that a stable spin configuration corresponds to the minimum of the thermodynamic internal energy (Equation A.76). Thus to evaluate stability, we can consider infinitesimal variations in the internal energy, $\mathcal{U} \to \mathcal{U} + \delta\mathcal{U}$, which are the result of variations in the magnetic moment, $\bar{\boldsymbol{m}}_i \to \bar{\boldsymbol{m}}_i + \delta\bar{\boldsymbol{m}}_i$. The variation in internal energy, calculated from Equation A.76 and only keeping first-order terms in $\delta\bar{\boldsymbol{m}}_i$, is given by $\delta\mathcal{U} = -\bar{\boldsymbol{H}}_i^e \cdot \delta\bar{\boldsymbol{m}}_i$. $\delta\bar{\boldsymbol{m}}_i$ can be considered as a small rotation in the vector, which in its most general form is $\delta\bar{\boldsymbol{m}}_i = \bar{\boldsymbol{m}}_i \times \delta\boldsymbol{\theta}$. This signifies a small rotation of $\bar{\boldsymbol{m}}_i$ about the axis specified by $\delta\boldsymbol{\theta}$. Therefore, internal energy can be written as $\delta\mathcal{U} = -\bar{\boldsymbol{H}}_i^e \cdot (\bar{\boldsymbol{m}}_i \times \delta\boldsymbol{\theta}) = (\bar{\boldsymbol{m}}_i \times \bar{\boldsymbol{H}}_i^e) \cdot \delta\boldsymbol{\theta}$. Internal energy minima corresponds to $\delta\mathcal{U}$ for any $\delta\boldsymbol{\theta}$, and thus we find the following

stability condition,

$$\bar{\boldsymbol{m}}_i \times \bar{\boldsymbol{H}}_i^e = 0 \qquad (\text{or } (\bar{\boldsymbol{m}}_i \parallel \bar{\boldsymbol{H}}_i^e)) \qquad i = 1, ..., N \tag{A.81}$$

This set of coupled equations indicates that in stable configurations the magnetic moment at site i is aligned to the effective field at that point, $\bar{\boldsymbol{H}}_i^e$. Exploiting this property and multiplying (dot product) Equation A.79 by the effective field, gives

$$\bar{\boldsymbol{H}}_i^e \cdot \bar{\boldsymbol{m}}_i = \frac{\int d\boldsymbol{s} \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s} e^{\beta\mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}}}{\int d\boldsymbol{s} e^{\beta\mu \bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s}}}. \tag{A.82}$$

Since we expect that in stable states, $\bar{\boldsymbol{H}}_i^e$ and $\bar{\boldsymbol{m}}_i$ are aligned, we can assume that $\bar{\boldsymbol{H}}_i^e \cdot \bar{\boldsymbol{m}}_i = |\bar{\boldsymbol{H}}_i^e||\bar{\boldsymbol{m}}_i|$ and that $\bar{\boldsymbol{H}}_i^e \cdot \boldsymbol{s} = \bar{\boldsymbol{H}}_i^e \cos\theta$, where θ is the angle between $\bar{\boldsymbol{H}}_i^e$ and $\boldsymbol{s}$, and $|\boldsymbol{s}| = 1$.

$$|\bar{\boldsymbol{H}}_i^e||\bar{\boldsymbol{m}}_i| = \frac{\int d\boldsymbol{s} \bar{\boldsymbol{H}}_i^e \cos\theta e^{\beta\mu \bar{\boldsymbol{H}}_i^e \cos\theta}}{\int d\boldsymbol{s} e^{\beta\mu \bar{\boldsymbol{H}}_i^e \cos\theta}} \tag{A.83}$$

Dividing both sides by $\bar{\boldsymbol{H}}_i^e$ and integrating over spherical coordinates, yields,

$$|\bar{\boldsymbol{m}}_i| = \mathbf{L}(\beta\mu|\bar{\boldsymbol{H}}_i^e|) \tag{A.84}$$

where $\mathbf{L}(x) = \coth(x) - x^{-1}$ is the Langevin function. To obtain the vector form, we note that that $\bar{\boldsymbol{m}}_i$ must align with the effective field, $\bar{\boldsymbol{H}}_i^e$,

$$\bar{\boldsymbol{m}}_i = \mathbf{L}(\beta\mu|\bar{\boldsymbol{H}}_i^e|)\frac{\bar{\boldsymbol{H}}_i^e}{|\bar{\boldsymbol{H}}_i^e|} \tag{A.85}$$

$$\mu\bar{\boldsymbol{H}}_i^e = J_{ij}\bar{\boldsymbol{m}}_j - \boldsymbol{D}_{ij} \times \bar{\boldsymbol{m}}_j - \mu\boldsymbol{H}_i \tag{A.86}$$

Equations A.85 and A.86 are the final results of the full HS derivation for the mean field model. They represent the simplest algorithm for computing the MF ground state: for every spin, $\bar{\boldsymbol{H}}_i^e$ is first computed for all the nearest-neighbour and single spin interactions, then the Langevin function is evaluated to find the new spin moment, $\bar{\boldsymbol{m}}_i$[3].

[3]From the main algorithm in Section 2.4.2, we see that the new spin moment actually only contains a small fraction of the moment calculated by $\mathbf{L}$ added to the old spin. This is to control the rate of convergence.

A.3 Tuning the Extended U-Net

In this section, we present the hyperparameter tuning results of the (1L, 1G) extended U-Net model, tested in Chapter 8. In particular, we tune the γ in the loss function, and the number of layers/nodes in the U-Net and Dense branches. The results are given in Tables A.1-A.3.

Hyperparameter Tuning of γ

γ	Validation Loss
0.1	0.0257813
1	0.0207061
2	0.0198380
10	8.0847902
2500	12.4036674

TABLE A.1: Loss on the validation set for different values for hyperparameter γ in the extended U-Net architecture.

Tuning U-Net Branch Architecture

#Layers/#Nodes - U-Net	Validation Loss
(16, 32)	8.1087703
(32, 64)	0.0457296
(8, 16, 32, 64)	1.4887293
(16, 32, 64, 128)	1.2743723
(32, 64, 128, 256)	2.0644283
(64, 128, 256, 512)	1.6617354
(16, 32, 64, 128, 256)	0.0156831
(32, 64, 128, 256, 512)	0.0225782
(64, 128, 256, 512, 1024)	0.0261728

TABLE A.2: Loss on the validation set for different number of layers and nodes in the U-Net branch of our architecture.

Tuning Dense Branch Architecture

#Layers/#Nodes - Dense	Validation Loss
(32, 16)	0.0202541
(64, 16)	0.0171578
(64, 32)	0.0185938
(128, 64)	0.0183323
(64, 32, 16)	0.0215993
(128, 64, 32)	0.0223834
(128, 64, 32, 16, 8)	0.0202563
(256, 128, 64, 32, 16)	0.0154065
(512, 256, 128, 64, 32)	0.0170458

TABLE A.3: Loss on the validation set for different number of layers and nodes in the Dense branch of our architecture.

Bibliography

Here are the references in citation order.

[1] J. Slonczewski, "Current-driven excitation of magnetic multilayers," *Journal of Magnetism and Magnetic Materials*, vol. 159, no. 1, pp. L1–L7, 1996, ISSN: 0304-8853. DOI: https://doi.org/10.1016/0304-8853(96)00062-5. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0304885396000625.

[2] L. Berger, "Emission of spin waves by a magnetic multilayer traversed by a current," *Phys. Rev. B*, vol. 54, pp. 9353–9358, 13 Oct. 1996. DOI: 10.1103/PhysRevB.54.9353. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.54.9353.

[3] D. Ralph and M. Stiles, "Spin transfer torques," *Journal of Magnetism and Magnetic Materials*, vol. 320, no. 7, pp. 1190–1216, 2008, ISSN: 0304-8853. DOI: https://doi.org/10.1016/j.jmmm.2007.12.019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304885307010116.

[4] X. S. Wang, H. Y. Yuan, and X. R. Wang, "A theory on skyrmion size," *Communications Physics*, vol. 1, no. 1, p. 31, Jul. 2018, ISSN: 2399-3650. DOI: 10.1038/s42005-018-0029-0. [Online]. Available: https://doi.org/10.1038/s42005-018-0029-0.

[5] N. Romming, A. Kubetzka, C. Hanneken, K. von Bergmann, and R. Wiesendanger, "Field-dependent size and shape of single magnetic skyrmions," *Phys. Rev. Lett.*, vol. 114, p. 177 203, 17 May 2015. DOI: 10.1103/PhysRevLett.114.177203. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.114.177203.

[6] A. Schwarz, *Topology for physicists*. Springer-Verlag, 1996.

[7] H.-B. Braun, "Topological effects in nanomagnetism: From superparamagnetism to chiral quantum solitons," *Advances in Physics*, vol. 61, no. 1, pp. 1–116, 2012. DOI: 10.1080/00018732.2012.663070. eprint: https://doi.org/10.1080/00018732.2012.663070. [Online]. Available: https://doi.org/10.1080/00018732.2012.663070.

[8] J. Lin, "Thermal modeling and management of dram memory systems," Ph.D. dissertation, Iowa State University, 2008. DOI: https://doi.org/10.31274/etd-180810-2677.

[9] J. Iyer, C. L. Hall, J. Shi, and Y. Huang, "System memory power and thermal management platforms built on intel centrino duo mobile technology," *Intel Technology Journal*, vol. 10, 2006.

[10] E. Farjallah, "Monitoring of temperature effects on cmos memories," 2018.

[11] S. Mühlbauer *et al.*, "Skyrmion lattice in a chiral magnet," *Science*, vol. 323, no. 5916, pp. 915–919, 2009, ISSN: 0036-8075. DOI: 10.1126/science.1166767. eprint: https://science.sciencemag.org/content/323/5916/915.full.pdf. [Online]. Available: https://science.sciencemag.org/content/323/5916/915.

[12] X. Z. Yu *et al.*, "Real-space observation of a two-dimensional skyrmion crystal," *Nature*, vol. 465, no. 7300, pp. 901–904, 2010, ISSN: 1476-4687. DOI: 10.1038/nature09124. [Online]. Available: https://doi.org/10.1038/nature09124.

[13] W. Münzer *et al.*, "Skyrmion lattice in the doped semiconductor $Fe_{1-x}Co_xSi$," *Phys. Rev. B*, vol. 81, p. 041 203, 4 Jan. 2010. DOI: 10.1103/PhysRevB.81.041203. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.81.041203.

[14] X. Z. Yu *et al.*, "Near room-temperature formation of a skyrmion crystal in thin-films of the helimagnet fege," *Nature Materials*, vol. 10, no. 2, pp. 106–109, Feb. 2011, ISSN: 1476-4660. DOI: 10.1038/nmat2916. [Online]. Available: https://doi.org/10.1038/nmat2916.

[15] M. N. Wilson *et al.*, "Stability and metastability of skyrmions in thin lamellae of Cu_2OSeO_3," *Phys. Rev. Research*, vol. 2, p. 013 096, 1 Jan. 2020. DOI: 10.

1103/PhysRevResearch.2.013096. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevResearch.2.013096.

[16] R. Ritz *et al.*, "Giant generic topological hall resistivity of mnsi under pressure," *Phys. Rev. B*, vol. 87, p. 134 424, 13 Apr. 2013. DOI: 10.1103/PhysRevB.87.134424. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.87.134424.

[17] M. Birch *et al.*, "Increased lifetime of metastable skyrmions by doping," Sep. 2018.

[18] A. Soumyanarayanan *et al.*, "Tunable room-temperature magnetic skyrmions in ir/fe/co/pt multilayers," *Nature Materials*, vol. 16, no. 9, pp. 898–904, Sep. 2017, ISSN: 1476-4660. DOI: 10.1038/nmat4934. [Online]. Available: https://doi.org/10.1038/nmat4934.

[19] G. Pradhan, B. Ojha, and S. Bedanta, "Effect of random anisotropy in stabilization of skyrmions and antiskyrmions," *Journal of Magnetism and Magnetic Materials*, vol. 528, p. 167 805, 2021, ISSN: 0304-8853. DOI: https://doi.org/10.1016/j.jmmm.2021.167805. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304885321000810.

[20] W. Legrand *et al.*, "Room-temperature current-induced generation and motion of sub-100 nm skyrmions," *Nano Letters*, vol. 17, no. 4, pp. 2703–2712, Apr. 2017, ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.7b00649. [Online]. Available: https://doi.org/10.1021/acs.nanolett.7b00649.

[21] D. Kahng and S. M. Sze, "A floating gate and its application to memory devices," *The Bell System Technical Journal*, vol. 46, no. 6, pp. 1288–1295, 1967. DOI: 10.1002/j.1538-7305.1967.tb01738.x.

[22] N. Romming *et al.*, "Writing and deleting single magnetic skyrmions," *Science*, vol. 341, no. 6146, pp. 636–639, 2013, ISSN: 0036-8075. DOI: 10.1126/science.1240573. eprint: https://science.sciencemag.org/content/341/6146/636.full.pdf. [Online]. Available: https://science.sciencemag.org/content/341/6146/636.

[23] S. Woo *et al.*, "Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets," *Nature Materials*, vol. 15, no. 5, pp. 501–506, May 2016, ISSN: 1476-4660. DOI: 10.1038/nmat4593. [Online]. Available: https://doi.org/10.1038/nmat4593.

[24] G. Yu *et al.*, "Room-temperature skyrmions in an antiferromagnet-based heterostructure," *Nano Letters*, vol. 18, no. 2, pp. 980–986, Feb. 2018, ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.7b04400. [Online]. Available: https://doi.org/10.1021/acs.nanolett.7b04400.

[25] O. Boulle *et al.*, "Room-temperature chiral magnetic skyrmions in ultrathin magnetic nanostructures," *Nature Nanotechnology*, vol. 11, no. 5, pp. 449–454, May 2016, ISSN: 1748-3395. DOI: 10.1038/nnano.2015.315. [Online]. Available: https://doi.org/10.1038/nnano.2015.315.

[26] G. Chen, A. Mascaraque, A. T. N'Diaye, and A. K. Schmid, "Room temperature skyrmion ground state stabilized through interlayer exchange coupling," *Applied Physics Letters*, vol. 106, no. 24, p. 242 404, 2015. DOI: 10.1063/1.4922726. eprint: https://doi.org/10.1063/1.4922726. [Online]. Available: https://doi.org/10.1063/1.4922726.

[27] C. Moreau-Luchaire *et al.*, "Additive interfacial chiral interaction in multilayers for stabilization of small individual skyrmions at room temperature," *Nature Nanotechnology*, vol. 11, no. 5, pp. 444–448, May 2016, ISSN: 1748-3395. DOI: 10.1038/nnano.2015.313. [Online]. Available: https://doi.org/10.1038/nnano.2015.313.

[28] G. Yu *et al.*, "Room-temperature creation and spin–orbit torque manipulation of skyrmions in thin films with engineered asymmetry," *Nano Letters*, vol. 16, no. 3, pp. 1981–1988, Mar. 2016, ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.5b05257. [Online]. Available: https://doi.org/10.1021/acs.nanolett.5b05257.

[29] W. Jiang *et al.*, "Blowing magnetic skyrmion bubbles," *Science*, vol. 349, no. 6245, pp. 283–286, 2015, ISSN: 0036-8075. DOI: 10.1126/science.aaa1442. eprint: https://science.sciencemag.org/content/349/6245/283.full.pdf. [Online]. Available: https://science.sciencemag.org/content/349/6245/283.

[30] M. Raju *et al.*, "The evolution of skyrmions in ir/fe/co/pt multilayers and their topological hall signature," *Nature Communications*, vol. 10, no. 1, p. 696, Mar. 2019, ISSN: 2041-1723. DOI: 10.1038/s41467-018-08041-9. [Online]. Available: https://doi.org/10.1038/s41467-018-08041-9.

[31] A. Schlenhoff *et al.*, "Magnetic nano-skyrmion lattice observed in a si-wafer-based multilayer system," *ACS Nano*, vol. 9, no. 6, pp. 5908–5912, Jun. 2015, ISSN: 1936-0851. DOI: 10.1021/acsnano.5b01146. [Online]. Available: https://doi.org/10.1021/acsnano.5b01146.

[32] J. F. Pulecio, A. Hrabec, K. Zeissler, R. M. White, Y. Zhu, and C. H. Marrows, *Hedgehog skyrmion bubbles in ultrathin films with interfacial dzyaloshinskii-moriya interactions*, 2016. arXiv: 1611.06869 [cond-mat.mtrl-sci].

[33] S. Zhang, J. Zhang, Y. Wen, E. M. Chudnovsky, and X. Zhang, "Creation of a thermally assisted skyrmion lattice in pt/co/ta multilayer films," *Applied Physics Letters*, vol. 113, no. 19, p. 192 403, 2018. DOI: 10.1063/1.5053983. eprint: https://doi.org/10.1063/1.5053983. [Online]. Available: https://doi.org/10.1063/1.5053983.

[34] S. Zhang, J. Zhang, Y. Wen, E. M. Chudnovsky, and X. Zhang, "Determination of chirality and density control of néel-type skyrmions with in-plane magnetic field," *Communications Physics*, vol. 1, no. 1, p. 36, Jul. 2018, ISSN: 2399-3650. DOI: 10.1038/s42005-018-0040-5. [Online]. Available: https://doi.org/10.1038/s42005-018-0040-5.

[35] J. Müller and A. Rosch, "Capturing of a magnetic skyrmion with a hole," *Phys. Rev. B*, vol. 91, p. 054 410, 5 Feb. 2015. DOI: 10.1103/PhysRevB.91.054410. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.91.054410.

[36] I. L. Fernandes, J. Chico, and S. Lounis, "Impurity-dependent gyrotropic motion, deflection and pinning of current-driven ultrasmall skyrmions in PdFe/ir(111) surface," *Journal of Physics: Condensed Matter*, vol. 32, no. 42, p. 425 802, Jul. 2020. DOI: 10.1088/1361-648x/ab9cf0. [Online]. Available: https://doi.org/10.1088/1361-648x/ab9cf0.

[37] W. Chen, L. Liu, and Y. Zheng, "Ultrafast ratchet dynamics of skyrmions by defect engineering in materials with poor conductivity under gigahertz magnetic fields," *Phys. Rev. Applied*, vol. 14, p. 064 014, 6 Dec. 2020. DOI: 10.1103/PhysRevApplied.14.064014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.14.064014.

[38] J. Castell-Queralt, L. González-Gómez, N. Del-Valle, A. Sanchez, and C. Navau, "Accelerating, guiding, and compressing skyrmions by defect rails," *Nanoscale*, vol. 11, pp. 12 589–12 594, 26 2019. DOI: 10.1039/C9NR02171J. [Online]. Available: http://dx.doi.org/10.1039/C9NR02171J.

[39] A. Bauer, M. Garst, and C. Pfleiderer, "History dependence of the magnetic properties of single-crystal $Fe_{1-x}Co_xSi$," *Phys. Rev. B*, vol. 93, p. 235 144, 23 Jun. 2016. DOI: 10.1103/PhysRevB.93.235144. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.93.235144.

[40] K. Karube *et al.*, "Skyrmion formation in a bulk chiral magnet at zero magnetic field and above room temperature," *Phys. Rev. Materials*, vol. 1, p. 074 405, 7 Dec. 2017. DOI: 10.1103/PhysRevMaterials.1.074405. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevMaterials.1.074405.

[41] N. Del-Valle, J. Castell-Queralt, L. González-Gómez, and C. Navau, "Defect modeling in skyrmionic ferromagnetic systems," en, *APL Mater.*, vol. 10, no. 1, p. 010 702, Jan. 2022.

[42] K. Buschow, *Concise Encyclopedia of Magnetic and Superconducting Materials* (Advances in Materials Sciences and Engineering). Elsevier Science, 2005, ISBN: 9780080457659. [Online]. Available: https://books.google.co.uk/books?id=N9mvytGEBtwC.

[43] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *CoRR*, vol. abs/1312.6114, 2014.

[44] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, 2015, pp. 234–241, ISBN: 9783319245744. DOI: 10.1007/978-3-319-24574-4_28. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-24574-4_28.

[45] E. Ising, "Beitrag zur theorie des ferromagnetismus," *Zeitschrift für Physik*, vol. 31, no. 1, pp. 253–258, Feb. 1925, ISSN: 0044-3328. DOI: 10.1007/BF02980577. [Online]. Available: https://doi.org/10.1007/BF02980577.

[46] L. Onsager, "Crystal statistics. i. a two-dimensional model with an order-disorder transition," *Phys. Rev.*, vol. 65, pp. 117–149, 3-4 Feb. 1944. DOI: 10.1103/PhysRev.65.117. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.65.117.

[47] W. Heisenberg, "Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen.," *Zeitschrift für Physik*, vol. 33, no. 1, pp. 879–893, Dec. 1925, ISSN: 0044-3328. DOI: 10.1007/BF01328377. [Online]. Available: https://doi.org/10.1007/BF01328377.

[48] L. Néel and R. Pauthenet, "Etude thermomagnétique d'un monocristal de Fe2O3-alpha," *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences*, vol. 234, no. 22, pp. 2172–2174, 1952. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02878466.

[49] I. Dzyaloshinsky, "A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics," *Journal of Physics and Chemistry of Solids*, vol. 4, no. 4, pp. 241–255, 1958, ISSN: 0022-3697. DOI: https://doi.org/10.1016/0022-3697(58)90076-3. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0022369758900763.

[50] T. Moriya, "Anisotropic superexchange interaction and weak ferromagnetism," *Phys. Rev.*, vol. 120, pp. 91–98, 1 Oct. 1960. DOI: 10.1103/PhysRev.120.91. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.120.91.

[51] V. E. Dmitrienko *et al.*, "Measuring the dzyaloshinskii–moriya interaction in a weak ferromagnet," *Nature Physics*, vol. 10, no. 3, pp. 202–206, Mar. 2014, ISSN: 1745-2481. DOI: 10.1038/nphys2859. [Online]. Available: https://doi.org/10.1038/nphys2859.

[52] A. N. Bogdaanov and D. A. Yablonskii, "Thermodynamically stable 'vortices' in magnetically ordered crystals. the mixed state of magnets," *JETP*, vol. 68, no. 1, p. 101, Jan. 1989. [Online]. Available: http://www.jetp.ac.ru/cgi-bin/e/index/e/68/1/p101?a=list.

[53] F. Bloch and G. Gentile, "Zur anisotropie der magnetisierung ferromagnetischer einkristalle," *Zeitschrift für Physik*, vol. 70, no. 5, pp. 395–408, May 1931, ISSN: 0044-3328. DOI: 10.1007/BF01339586. [Online]. Available: https://doi.org/10.1007/BF01339586.

[54] Y. Liu, D. J. Sellmyer, and D. Shindo, *Handbook of Advanced Magnetic Materials: Vol 1. Nanostructural Effects. Vol 2. Characterization and Simulation. Vol 3. Fabrication and Processing. Vol 4. Properties and Applications*. Springer US, 2006.

[55] N. A. Spaldin, *Magnetic materials fundamentals and device applications*. Cambridge Univ. Press, 2013.

[56] G. Bertotti, *Hysteresis in magnetism: for physicists, materials scientists, and engineers*. Academic Press, 2008.

[57] T. Ueno *et al.*, "Enhanced orbital magnetic moments in magnetic heterostructures with interface perpendicular magnetic anisotropy," *Scientific Reports*, vol. 5, no. 1, p. 14 858, Oct. 2015, ISSN: 2045-2322. DOI: 10.1038/srep14858. [Online]. Available: https://doi.org/10.1038/srep14858.

[58] R. Tilley, *Defects in Solids* (Special Topics in Inorganic Chemistry). Wiley, 2008, ISBN: 9780470380734. [Online]. Available: https://books.google.co.uk/books?id=I1YFofnHXusC.

[59] N. Del-Valle, J. Castell-Queralt, L. González-Gómez, and C. Navau, "Defect modeling in skyrmionic ferromagnetic systems," *APL Materials*, vol. 10, no. 1, p. 010 702, Jan. 2022, ISSN: 2166-532X. DOI: 10.1063/5.0072709. eprint: https://pubs.aip.org/aip/apm/article-pdf/doi/10.1063/5.0072709/19767149/010702_1_5.0072709.pdf. [Online]. Available: https://doi.org/10.1063/5.0072709.

[60] M. Ohring, *Materials Science of Thin Films Deposition and Structure*. Academic Press, 2002, ISBN: 978-0-12-524975-1. DOI: 10.1016/B978-0-12-524975-1.X5000-9.

[61] S. J. Poon and C. T. Ma, *Amorphous ferrimagnets: An ideal host for ultra-small skyrmions at room temperature*, 2019. arXiv: 1908.07003 [cond-mat.mtrl-sci].

[62] W.-Y. Choi, H.-W. Bang, S.-H. Chun, S. Lee, and M.-H. Jung, "Skyrmion phase in mnsi thin films grown on sapphire by a conventional sputtering," *Nanoscale Research Letters*, vol. 16, no. 1, p. 7, Jan. 2021, ISSN: 1556-276X. DOI: 10.1186/s11671-020-03462-2. [Online]. Available: https://doi.org/10.1186/s11671-020-03462-2.

[63] S. Woo *et al.*, "Current-driven dynamics and inhibition of the skyrmion hall effect of ferrimagnetic skyrmions in gdfeco films," *Nature Communications*, vol. 9, no. 1, p. 959, Mar. 2018, ISSN: 2041-1723. DOI: 10.1038/s41467-018-03378-7. [Online]. Available: https://doi.org/10.1038/s41467-018-03378-7.

[64] S. Woo *et al.*, "Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets," *Nature Materials*, vol. 15, no. 5, pp. 501–506, May 2016, ISSN: 1476-4660. DOI: 10.1038/nmat4593. [Online]. Available: https://doi.org/10.1038/nmat4593.

[65] Y. Hirata *et al.*, "Vanishing skyrmion hall effect at the angular momentum compensation temperature of a ferrimagnet," *Nature Nanotechnology*, vol. 14, no. 3, pp. 232–236, Mar. 2019, ISSN: 1748-3395. DOI: 10.1038/s41565-018-0345-2. [Online]. Available: https://doi.org/10.1038/s41565-018-0345-2.

[66] K. Chen, D. Lott, A. Philippi-Kobs, M. Weigand, C. Luo, and F. Radu, "Observation of compact ferrimagnetic skyrmions in dyco3 film," *Nanoscale*, vol. 12, pp. 18 137–18 143, 35 2020. DOI: 10.1039/D0NR02947E. [Online]. Available: http://dx.doi.org/10.1039/D0NR02947E.

[67] C. T. Ma, Y. Xie, H. Sheng, A. W. Ghosh, and S. J. Poon, "Robust formation of ultrasmall room-temperature neél skyrmions in amorphous ferrimagnets from atomistic simulations," *Scientific Reports*, vol. 9, no. 1, p. 9964, Jul. 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-46458-4. [Online]. Available: https://doi.org/10.1038/s41598-019-46458-4.

[68] Y. Imry and S.-k. Ma, "Random-field instability of the ordered state of continuous symmetry," *Phys. Rev. Lett.*, vol. 35, pp. 1399–1401, 21 Nov. 1975. DOI:

10.1103/PhysRevLett.35.1399. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.35.1399.

[69] D. Ryan, *Recent Progress in Random Magnets*. World Scientific, 1992, ISBN: 9789810208851. [Online]. Available: https://books.google.co.uk/books?id=BVPd7F06r8cC.

[70] E. M. Chudnovsky and D. A. Garanin, "Topological order generated by a random field in a 2d exchange model," *Phys. Rev. Lett.*, vol. 121, p. 017 201, 1 Jul. 2018. DOI: 10.1103/PhysRevLett.121.017201. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.121.017201.

[71] A. Berretti, "Some properties of random ising models," *Journal of Statistical Physics*, vol. 38, no. 3, pp. 483–496, Feb. 1985, ISSN: 1572-9613. DOI: 10.1007/BF01010473. [Online]. Available: https://doi.org/10.1007/BF01010473.

[72] M. Aizenman and J. Wehr, "Rounding of first-order phase transitions in systems with quenched disorder," *Phys. Rev. Lett.*, vol. 62, pp. 2503–2506, 21 May 1989. DOI: 10.1103/PhysRevLett.62.2503. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.62.2503.

[73] J. Bricmont and A. Kupiainen, "Phase transition in the 3d random field Ising model," *Communications in Mathematical Physics*, vol. 116, no. 4, pp. 539–572, 1988.

[74] J. Wehr, A. Niederberger, L. Sanchez-Palencia, and M. Lewenstein, "Disorder versus the mermin-wagner-hohenberg effect: From classical spin systems to ultracold atomic gases," *Phys. Rev. B*, vol. 74, p. 224 448, 22 Dec. 2006. DOI: 10.1103/PhysRevB.74.224448. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.74.224448.

[75] J. F. Fernandez, "Random fields generated by dilution in zero external field," *Europhysics Letters (EPL)*, vol. 5, no. 2, pp. 129–133, Jan. 1988. DOI: 10.1209/0295-5075/5/2/007. [Online]. Available: https://doi.org/10.1209/0295-5075/5/2/007.

[76] J. T. Graham, M. Maliepaard, J. H. Page, S. R. P. Smith, and D. R. Taylor, "Random-field effects on ising jahn-teller phase transitions," *Phys. Rev. B*,

vol. 35, pp. 2098–2101, 4 Feb. 1987. DOI: 10.1103/PhysRevB.35.2098. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.35.2098.

[77] D. Taylor, E. Zwartz, J. Page, and B. Watts, "Interpretation of jahn-teller phase transitions in mixed crystals using the random-field ising model," *Journal of Magnetism and Magnetic Materials*, vol. 54-57, pp. 57–58, 1986, ISSN: 0304-8853. DOI: https://doi.org/10.1016/0304-8853(86)90483-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/030488538690483X.

[78] A. Young, *Spin Glasses and Random Fields* (Directions in condensed matter physics). World Scientific, 1998, ISBN: 9789810232405. [Online]. Available: https://books.google.co.uk/books?id=ckJqDQAAQBAJ.

[79] R. Harris, M. Plischke, and M. J. Zuckermann, "New model for amorphous magnetism," *Phys. Rev. Lett.*, vol. 31, pp. 160–162, 3 Jul. 1973. DOI: 10.1103/PhysRevLett.31.160. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.31.160.

[80] R. Alben, J. J. Becker, and M. C. Chi, "Random anisotropy in amorphous ferromagnets," *Journal of Applied Physics*, vol. 49, no. 3, pp. 1653–1658, 1978. DOI: 10.1063/1.324881. eprint: https://doi.org/10.1063/1.324881. [Online]. Available: https://doi.org/10.1063/1.324881.

[81] G. Herzer, "The random anisotropy model," in *Properties and Applications of Nanocrystalline Alloys from Amorphous Precursors*, B. Idzikowski, P. Švec, and M. Miglierini, Eds., Dordrecht: Springer Netherlands, 2005, pp. 15–34, ISBN: 978-1-4020-2965-3.

[82] E. M. Chudnovsky and D. A. Garanin, "Skyrmion glass in a 2d heisenberg ferromagnet with quenched disorder," *New Journal of Physics*, vol. 20, no. 3, p. 033 006, Mar. 2018. DOI: 10.1088/1367-2630/aab576. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/aab576.

[83] R. Silva, R. Silva, W. Moura-Melo, and A. Pereira, "Skyrmion dynamics and skyrmionâ€"bimeron crossover in antiferromagnetic thin nanodisks with a random distribution of magnetic impurities," *Journal of Magnetism and Magnetic Materials*, vol. 546, p. 168 823, 2022, ISSN: 0304-8853. DOI: https://doi.

org/10.1016/j.jmmm.2021.168823. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304885321010362.

[84] H. Kahn, *Use of Different Monte Carlo Sampling Techniques*. Santa Monica, CA: RAND Corporation, 1955.

[85] D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, 2015, ISBN: 9781107074026. [Online]. Available: https://books.google.co.uk/books?id=CVreBAAAQBAJ.

[86] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953. DOI: 10.1063/1.1699114. eprint: https://doi.org/10.1063/1.1699114. [Online]. Available: https://doi.org/10.1063/1.1699114.

[87] D. Hinzke and U. Nowak, "Monte carlo simulation of magnetization switching in a heisenberg model for small ferromagnetic particles," *Computer Physics Communications*, vol. 121-122, pp. 334–337, 1999, Proceedings of the Europhysics Conference on Computational Physics CCP 1998, ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(99)00348-3. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010465599003483.

[88] P. Weiss, "L'hypothèse du champ moléculaire et la propriété ferromagnétique," *J. Phys. Theor. Appl.*, vol. 6, no. 1, pp. 661–690, 1907. DOI: 10.1051/jphystap:019070060066100. [Online]. Available: https://hal.archives-ouvertes.fr/jpa-00241247.

[89] N. M. A. 2020. "Pierre curie – biographical." (2020), [Online]. Available: https://www.nobelprize.org/prizes/physics/1903/pierre-curie/biographical/ (visited on 09/16/2012).

[90] B. Berg and M. Lüscher, "Definition and statistical distributions of a topological number in the lattice o(3) σ-model," *Nuclear Physics B*, vol. 190, no. 2, pp. 412–424, 1981, ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(81)90568-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/055032138190568X.

[91] G. Yin, Y. Li, L. Kong, R. K. Lake, C. L. Chien, and J. Zang, "Topological charge analysis of ultrafast single skyrmion creation," *Phys. Rev. B*, vol. 93, p. 174 403, 17 May 2016. DOI: 10.1103/PhysRevB.93.174403. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.93.174403.

[92] D. I. Cortes, "Computational simulations of complex chiral magnetic structures," Ph.D. dissertation, University of Southampton, Oct. 2017. [Online]. Available: https://eprints.soton.ac.uk/415834/.

[93] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006, ISBN: 0387310738.

[94] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016, ISBN: 0262035618.

[95] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989, ISSN: 0893-6080. DOI: https://doi.org/10.1016/0893-6080(89)90020-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0893608089900208.

[96] S. Linnainmaa, "Taylor expansion of the accumulated rounding error," *BIT*, vol. 16, no. 2, pp. 146–160, Jun. 1976, ISSN: 1572-9125. DOI: 10.1007/bf01931367. [Online]. Available: http://dx.doi.org/10.1007/BF01931367.

[97] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. DOI: 10.1109/5.726791.

[98] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, ISSN: 0001-0782. DOI: 10.1145/3065386. [Online]. Available: https://doi.org/10.1145/3065386.

[99] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:14124313.

[100] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9. DOI: 10.1109/CVPR.2015.7298594.

[101] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:206594692.

[102] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587. DOI: 10.1109/CVPR.2014.81.

[103] H. Gholamalinezhad and H. Khosravi, "Pooling methods in deep neural networks, a review," *ArXiv*, vol. abs/2009.07485, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221738915.

[104] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Computer Vision – ECCV 2014*, ser. Lecture notes in computer science, Cham: Springer International Publishing, 2014, pp. 740–755.

[105] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017. DOI: 10.1109/TPAMI.2016.2572683.

[106] D. Cireşan, A. Giusti, L. Gambardella, and J. Schmidhuber, "Neural networks for segmenting neuronal structures in em stacks," [Online]. Available: https://people.idsia.ch/~ciresan/data/ISBI2012.pdf.

[107] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, 2481ndash, 2495, Dec. 2017, ISSN: 1939-3539. DOI: 10.1109/tpami.2016.2644615. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2016.2644615.

[108] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Lecture Notes*

in Computer Science. Springer International Publishing, 2018, pp. 3–11, ISBN: 9783030008895. DOI: 10.1007/978-3-030-00889-5_1. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-00889-5_1.

[109] I. Jolliffe. Springer-Verlag, 2002, ISBN: 0387954422. DOI: 10.1007/b98835. [Online]. Available: http://dx.doi.org/10.1007/b98835.

[110] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, Nov. 1901, ISSN: 1941-5990. DOI: 10.1080/14786440109462720. [Online]. Available: http://dx.doi.org/10.1080/14786440109462720.

[111] H. Hotelling, "Analysis of a complex of statistical variables into principal components.," *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417–441, Sep. 1933, ISSN: 0022-0663. DOI: 10.1037/h0071325. [Online]. Available: http://dx.doi.org/10.1037/h0071325.

[112] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20 150 202, Apr. 2016, ISSN: 1471-2962. DOI: 10.1098/rsta.2015.0202. [Online]. Available: http://dx.doi.org/10.1098/rsta.2015.0202.

[113] J. Paisley, D. M. Blei, and M. I. Jordan, "Variational bayesian inference with stochastic search," in *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ser. ICML'12, Edinburgh, Scotland: Omnipress, 2012, pp. 1363–1370, ISBN: 9781450312851.

[114] T. Skyrme, "A unified field theory of mesons and baryons," *Nuclear Physics*, vol. 31, pp. 556–569, 1962, ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(62)90775-7. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0029558262907757.

[115] D. C. Wright and N. D. Mermin, "Crystalline liquids: The blue phases," *Rev. Mod. Phys.*, vol. 61, pp. 385–432, 2 Apr. 1989. DOI: 10.1103/RevModPhys.61.385. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.61.385.

[116] R. A. Battye, N. R. Cooper, and P. M. Sutcliffe, "Stable skyrmions in two-component bose-einstein condensates," *Phys. Rev. Lett.*, vol. 88, p. 080 401, 8 Feb. 2002. DOI: 10.1103/PhysRevLett.88.080401. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.88.080401.

[117] U. Al Khawaja and H. Stoof, "Skyrmions in a ferromagnetic bose-einstein condensate," *Nature*, vol. 411, no. 6840, pp. 918–920, 2001, ISSN: 1476-4687. DOI: 10.1038/35082010. [Online]. Available: https://doi.org/10.1038/35082010.

[118] L. S. Leslie, A. Hansen, K. C. Wright, B. M. Deutsch, and N. P. Bigelow, "Creation and detection of skyrmions in a bose-einstein condensate," *Phys. Rev. Lett.*, vol. 103, p. 250 401, 25 Dec. 2009. DOI: 10.1103/PhysRevLett.103.250401. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.103.250401.

[119] U. A. Khawaja and H. T. C. Stoof, "Skyrmion physics in bose-einstein ferromagnets," *Phys. Rev. A*, vol. 64, p. 043 612, 4 Sep. 2001. DOI: 10.1103/PhysRevA.64.043612. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.64.043612.

[120] D.-H. Lee and C. L. Kane, "Boson-vortex-skyrmion duality, spin-singlet fractional quantum hall effect, and spin-1/2 anyon superconductivity," *Phys. Rev. Lett.*, vol. 64, pp. 1313–1317, 12 Mar. 1990. DOI: 10.1103/PhysRevLett.64.1313. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.64.1313.

[121] S. L. Sondhi, A. Karlhede, S. A. Kivelson, and E. H. Rezayi, "Skyrmions and the crossover from the integer to fractional quantum hall effect at small zeeman energies," *Phys. Rev. B*, vol. 47, pp. 16 419–16 426, 24 Jun. 1993. DOI: 10.1103/PhysRevB.47.16419. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.47.16419.

[122] A. Schmeller, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, "Evidence for skyrmions and single spin flips in the integer quantized hall effect," *Phys. Rev. Lett.*, vol. 75, pp. 4290–4293, 23 Dec. 1995. DOI: 10.1103/PhysRevLett.

75.4290. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.75.4290.

[123] K. Yang, S. Das Sarma, and A. H. MacDonald, "Collective modes and skyrmion excitations in graphene $SU(4)$ quantum hall ferromagnets," *Phys. Rev. B*, vol. 74, p. 075 423, 7 Aug. 2006. DOI: 10.1103/PhysRevB.74.075423. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.74.075423.

[124] N. Nagaosa and Y. Tokura, "Topological properties and dynamics of magnetic skyrmions," *Nature Nanotechnology*, vol. 8, no. 12, pp. 899–911, Dec. 2013, ISSN: 1748-3395. DOI: 10.1038/nnano.2013.243. [Online]. Available: https://doi.org/10.1038/nnano.2013.243.

[125] A. K. Nayak *et al.*, "Magnetic antiskyrmions above room temperature in tetragonal heusler materials," *Nature*, vol. 548, no. 7669, pp. 561–566, Aug. 2017, ISSN: 1476-4687. DOI: 10.1038/nature23466. [Online]. Available: https://doi.org/10.1038/nature23466.

[126] R. Saha *et al.*, "Intrinsic stability of magnetic anti-skyrmions in the tetragonal inverse heusler compound mn1.4pt0.9pd0.1sn," *Nature Communications*, vol. 10, no. 1, p. 5305, Nov. 2019, ISSN: 2041-1723. DOI: 10.1038/s41467-019-13323-x. [Online]. Available: https://doi.org/10.1038/s41467-019-13323-x.

[127] X. Z. Yu *et al.*, "Biskyrmion states and their current-driven motion in a layered manganite," en, *Nat. Commun.*, vol. 5, no. 1, p. 3198, 2014.

[128] J. C. Loudon *et al.*, "Do images of biskyrmions show type-II bubbles?" en, *Adv. Mater.*, vol. 31, no. 16, e1806598, Apr. 2019.

[129] S. Zhang, F. Kronast, G. van der Laan, and T. Hesjedal, "Real-space observation of skyrmionium in a ferromagnet-magnetic topological insulator heterostructure," en, *Nano Lett.*, vol. 18, no. 2, pp. 1057–1063, Feb. 2018.

[130] X. Zhang *et al.*, "Control and manipulation of a magnetic skyrmionium in nanostructures," *Phys. Rev. B*, vol. 94, p. 094 420, 9 Sep. 2016. DOI: 10.1103/PhysRevB.94.094420. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.94.094420.

[131] P. Taylor, *Data growth worldwide 2010-2025*, Nov. 2023. [Online]. Available: https://www.statista.com/statistics/871513/worldwide-data-created/.

[132] G. E. Moore, "Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff.," *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33–35, Sep. 2006. DOI: 10.1109/N-SSC.2006.4785860.

[133] N. S. Kim *et al.*, "Leakage current: Moore's law meets static power," *Computer*, vol. 36, no. 12, pp. 68–75, Dec. 2003. DOI: 10.1109/MC.2003.1250885.

[134] K. K. Thornber, T. C. McGill, and C. A. Mead, "The tunneling time of an electron," *Journal of Applied Physics*, vol. 38, no. 5, pp. 2384–2385, 1967. DOI: 10.1063/1.1709888. eprint: https://doi.org/10.1063/1.1709888. [Online]. Available: https://doi.org/10.1063/1.1709888.

[135] P. Dey and J. N. Roy, *Spintronics: Fundamentals and Applications*. Springer Singapore, 2021, ISBN: 9789811600692. DOI: 10.1007/978-981-16-0069-2. [Online]. Available: http://dx.doi.org/10.1007/978-981-16-0069-2.

[136] V. K. Joshi, "Spintronics: A contemporary review of emerging electronics devices," *Engineering Science and Technology, an International Journal*, vol. 19, no. 3, pp. 1503–1513, 2016, ISSN: 2215-0986. DOI: https://doi.org/10.1016/j.jestch.2016.05.002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2215098615300501.

[137] A. Hirohata *et al.*, "Review on spintronics: Principles and device applications," *Journal of Magnetism and Magnetic Materials*, vol. 509, p. 166 711, 2020, ISSN: 0304-8853. DOI: https://doi.org/10.1016/j.jmmm.2020.166711. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304885320302353.

[138] E. Y. Tsymbal and I. Zutic, Eds., *Spintronics handbook, second edition: Spin transport and magnetism*, 2nd ed. London, England: CRC Press, Mar. 2021.

[139] M. N. Baibich *et al.*, "Giant magnetoresistance of (001)fe/(001)cr magnetic superlattices," *Physical Review Letters*, vol. 61, no. 21, pp. 2472–2475, Nov. 1988,

ISSN: 0031-9007. DOI: 10.1103/physrevlett.61.2472. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.61.2472.

[140] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange," *Physical Review B*, vol. 39, no. 7, pp. 4828–4830, Mar. 1989, ISSN: 0163-1829. DOI: 10.1103/physrevb.39.4828. [Online]. Available: http://dx.doi.org/10.1103/PhysRevB.39.4828.

[141] S. Meyer *et al.*, "Isolated zero field sub-10nm skyrmions in ultrathin co films," *Nature Communications*, vol. 10, no. 1, Aug. 2019, ISSN: 2041-1723. DOI: 10.1038/s41467-019-11831-4. [Online]. Available: http://dx.doi.org/10.1038/s41467-019-11831-4.

[142] G. Chen, A. Mascaraque, A. T. N'Diaye, and A. K. Schmid, "Room temperature skyrmion ground state stabilized through interlayer exchange coupling," *Applied Physics Letters*, vol. 106, no. 24, Jun. 2015, ISSN: 1077-3118. DOI: 10.1063/1.4922726. [Online]. Available: http://dx.doi.org/10.1063/1.4922726.

[143] S. Zhang, A. A. Baker, S. Komineas, and T. Hesjedal, "Topological computation based on direct magnetic logic communication," *Scientific Reports*, vol. 5, no. 1, p. 15 773, Oct. 2015, ISSN: 2045-2322. DOI: 10.1038/srep15773. [Online]. Available: https://doi.org/10.1038/srep15773.

[144] S. S. P. Parkin, M. Hayashi, and L. Thomas, "Magnetic domain-wall racetrack memory," *Science*, vol. 320, no. 5873, pp. 190–194, 2008, ISSN: 0036-8075. DOI: 10.1126/science.1145799. eprint: https://science.sciencemag.org/content/320/5873/190.full.pdf. [Online]. Available: https://science.sciencemag.org/content/320/5873/190.

[145] E. Y. Vedmedenko *et al.*, "The 2020 magnetism roadmap," *Journal of Physics D: Applied Physics*, vol. 53, no. 45, p. 453 001, Aug. 2020, ISSN: 1361-6463. DOI: 10.1088/1361-6463/ab9d98. [Online]. Available: http://dx.doi.org/10.1088/1361-6463/ab9d98.

[146] J. Slonczewski, "Current-driven excitation of magnetic multilayers," *Journal of Magnetism and Magnetic Materials*, vol. 159, no. 1–2, pp. L1–L7, Jun. 1996, ISSN: 0304-8853. DOI: 10.1016/0304-8853(96)00062-5. [Online]. Available: http://dx.doi.org/10.1016/0304-8853(96)00062-5.

[147] A. Thiaville, Y. Nakatani, J. Miltat, and Y. Suzuki, "Micromagnetic understanding of current-driven domain wall motion in patterned nanowires," *Europhysics Letters (EPL)*, vol. 69, no. 6, pp. 990–996, Mar. 2005, ISSN: 1286-4854. DOI: 10.1209/epl/i2004-10452-6. [Online]. Available: http://dx.doi.org/10.1209/epl/i2004-10452-6.

[148] T. Koyama *et al.*, "Observation of the intrinsic pinning of a magnetic domain wall in a ferromagnetic nanowire," *Nature Materials*, vol. 10, no. 3, pp. 194–197, Feb. 2011, ISSN: 1476-4660. DOI: 10.1038/nmat2961. [Online]. Available: http://dx.doi.org/10.1038/nmat2961.

[149] M. Hayashi, L. Thomas, C. Rettner, R. Moriya, X. Jiang, and S. S. P. Parkin, "Dependence of current and field driven depinning of domain walls on their structure and chirality in permalloy nanowires," *Physical Review Letters*, vol. 97, no. 20, Nov. 2006, ISSN: 1079-7114. DOI: 10.1103/physrevlett.97.207205. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.97.207205.

[150] A. Fert, V. Cros, and J. Sampaio, "Skyrmions on the track," *Nature Nanotechnology*, vol. 8, no. 3, pp. 152–156, Mar. 2013, ISSN: 1748-3395. DOI: 10.1038/nnano.2013.29. [Online]. Available: https://doi.org/10.1038/nnano.2013.29.

[151] R. Tomasello, E. Martinez, R. Zivieri, L. Torres, M. Carpentieri, and G. Finocchio, "A strategy for the design of skyrmion racetrack memories," *Scientific Reports*, vol. 4, no. 1, p. 6784, Oct. 2014, ISSN: 2045-2322. DOI: 10.1038/srep06784. [Online]. Available: https://doi.org/10.1038/srep06784.

[152] J. Iwasaki, M. Mochizuki, and N. Nagaosa, "Universal current-velocity relation of skyrmion motion in chiral magnets," *Nature Communications*, vol. 4, no. 1, p. 1463, Feb. 2013, ISSN: 2041-1723. DOI: 10.1038/ncomms2442. [Online]. Available: https://doi.org/10.1038/ncomms2442.

[153] F. Jonietz *et al.*, "Spin transfer torques in mnsi at ultralow current densities," *Science*, vol. 330, no. 6011, pp. 1648–1651, 2010, ISSN: 0036-8075. DOI: 10.1126/science.1195709. eprint: https://science.sciencemag.org/content/330/6011/1648.full.pdf. [Online]. Available: https://science.sciencemag.org/content/330/6011/1648.

[154] X. Z. Yu *et al.*, "Skyrmion flow near room temperature in an ultralow current density," *Nature Communications*, vol. 3, no. 1, p. 988, Aug. 2012, ISSN: 2041-1723. DOI: 10.1038/ncomms1990. [Online]. Available: https://doi.org/10.1038/ncomms1990.

[155] L. Thomas, R. Moriya, C. Rettner, and S. S. Parkin, "Dynamics of magnetic domain walls under their own inertia," *Science*, vol. 330, no. 6012, pp. 1810–1813, 2010, ISSN: 0036-8075. DOI: 10.1126/science.1197468. eprint: https://science.sciencemag.org/content/330/6012/1810.full.pdf. [Online]. Available: https://science.sciencemag.org/content/330/6012/1810.

[156] M. Tsoi, R. E. Fontana, and S. S. P. Parkin, "Magnetic domain wall motion triggered by an electric current," *Applied Physics Letters*, vol. 83, no. 13, pp. 2617–2619, 2003. DOI: 10.1063/1.1578165. eprint: https://doi.org/10.1063/1.1578165. [Online]. Available: https://doi.org/10.1063/1.1578165.

[157] I.-S. Hong and K.-J. Lee, "Magnetic skyrmion field-effect transistors," *Applied Physics Letters*, vol. 115, no. 7, p. 072 406, 2019. DOI: 10.1063/1.5110752. eprint: https://doi.org/10.1063/1.5110752. [Online]. Available: https://doi.org/10.1063/1.5110752.

[158] Y. Tchoe and J. H. Han, "Skyrmion generation by current," *Phys. Rev. B*, vol. 85, p. 174 416, 17 May 2012. DOI: 10.1103/PhysRevB.85.174416. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.85.174416.

[159] J. Sampaio, V. Cros, S. Rohart, A. Thiaville, and A. Fert, "Nucleation, stability and current-induced motion of isolated magnetic skyrmions in nanostructures," en, *Nat. Nanotechnol.*, vol. 8, no. 11, pp. 839–844, Nov. 2013.

[160] J. Iwasaki, M. Mochizuki, and N. Nagaosa, "Current-induced skyrmion dynamics in constricted geometries," en, *Nat. Nanotechnol.*, vol. 8, no. 10, pp. 742–747, Oct. 2013.

[161] M. Mochizuki and Y. Watanabe, "Writing a skyrmion on multiferroic materials," en, *Appl. Phys. Lett.*, vol. 107, no. 8, p. 082 409, Aug. 2015.

[162] P.-J. Hsu, A. Kubetzka, A. Finco, N. Romming, K. von Bergmann, and R. Wiesendanger, "Electric-field-driven switching of individual magnetic skyrmions," en, *Nat. Nanotechnol.*, vol. 12, no. 2, pp. 123–126, Feb. 2017.

[163] Z. Wang *et al.*, "Thermal generation, manipulation and thermoelectric detection of skyrmions," en, *Nat. Electron.*, vol. 3, no. 11, pp. 672–679, Oct. 2020.

[164] L.-M. Kern *et al.*, "Deterministic generation and guided motion of magnetic skyrmions by focused he+-ion irradiation," en, *Nano Lett.*, vol. 22, no. 10, pp. 4028–4035, May 2022.

[165] K. Fallon *et al.*, "Controlled individual skyrmion nucleation at artificial defects formed by ion irradiation," en, *Small*, vol. 16, no. 13, e1907450, Apr. 2020.

[166] F. Büttner *et al.*, "Field-free deterministic ultrafast creation of magnetic skyrmions by spin-orbit torques," en, *Nat. Nanotechnol.*, vol. 12, no. 11, pp. 1040–1044, Nov. 2017.

[167] X. Zhang *et al.*, "Skyrmions in magnetic tunnel junctions," en, *ACS Appl. Mater. Interfaces*, vol. 10, no. 19, pp. 16 887–16 892, May 2018.

[168] R. Tomasello, M. Ricci, P. Burrascano, V. Puliafito, M. Carpentieri, and G. Finocchio, "Electrical detection of single magnetic skyrmion at room temperature," *AIP Adv.*, vol. 7, no. 5, p. 056 022, May 2017.

[169] T. Schulz *et al.*, "Emergent electrodynamics of skyrmions in a chiral magnet," *Nature Physics*, vol. 8, no. 4, pp. 301–304, Apr. 2012, ISSN: 1745-2481. DOI: 10.1038/nphys2231. [Online]. Available: https://doi.org/10.1038/nphys2231.

[170] K. Hamamoto, M. Ezawa, and N. Nagaosa, "Purely electrical detection of a skyrmion in constricted geometry," en, *Appl. Phys. Lett.*, vol. 108, no. 11, p. 112 401, Mar. 2016.

[171] B. Zhang, W. Wang, M. Beg, H. Fangohr, and W. Kuch, "Microwave-induced dynamic switching of magnetic skyrmion cores in nanodots," en, *Appl. Phys. Lett.*, vol. 106, no. 10, p. 102 401, Mar. 2015.

[172] C. Heo, N. S. Kiselev, A. K. Nandy, S. Blügel, and T. Rasing, "Switching of chiral magnetic skyrmions by picosecond magnetic field pulses via transient topological states," en, *Sci. Rep.*, vol. 6, no. 1, Jun. 2016.

[173] P. Upadhyaya, G. Yu, P. K. Amiri, and K. L. Wang, "Electric-field guiding of magnetic skyrmions," *Phys. Rev. B*, vol. 92, p. 134 411, 13 Oct. 2015. DOI: 10.1103/PhysRevB.92.134411. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.92.134411.

[174] M. Schott *et al.*, "The skyrmion switch: Turning magnetic skyrmion bubbles on and off with an electric field," en, *Nano Lett.*, vol. 17, no. 5, pp. 3006–3012, May 2017.

[175] A. De Lucia, K. Litzius, B. Krüger, O. A. Tretiakov, and M. Kläui, "Multiscale simulations of topological transformations in magnetic-skyrmion spin structures," *Phys. Rev. B*, vol. 96, p. 020 405, 2 Jul. 2017. DOI: 10.1103/PhysRevB.96.020405. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.96.020405.

[176] S. Woo *et al.*, "Deterministic creation and deletion of a single magnetic skyrmion observed by direct time-resolved x-ray microscopy," en, *Nat. Electron.*, vol. 1, no. 5, pp. 288–296, May 2018.

[177] X. Zhang *et al.*, "Skyrmion-electronics: Writing, deleting, reading and processing magnetic skyrmions toward spintronic applications," en, *J. Phys. Condens. Matter*, vol. 32, no. 14, p. 143 001, Apr. 2020.

[178] W. Jiang, G. Chen, K. Liu, J. Zang, S. G. te Velthuis, and A. Hoffmann, "Skyrmions in magnetic multilayers," *Physics Reports*, vol. 704, pp. 1–49, 2017, Skyrmions in Magnetic Multilayers, ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2017.08.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0370157317302934.

[179] C. Reichhardt, D. Ray, and C. J. O. Reichhardt, "Collective transport properties of driven skyrmions with random disorder," *Phys. Rev. Lett.*, vol. 114, p. 217 202, 21 May 2015. DOI: 10.1103/PhysRevLett.114.217202. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.114.217202.

[180] M. Xu *et al.*, "Reconfigurable skyrmion logic gates with auto-annihilating skyrmion function," *Physica Scripta*, vol. 98, no. 10, p. 105 939, Sep. 2023, ISSN: 1402-4896. DOI: 10.1088/1402-4896/acf742. [Online]. Available: http://dx.doi.org/10.1088/1402-4896/acf742.

[181] D. Yu, H. Yang, M. Chshiev, and A. Fert, "Skyrmions-based logic gates in one single nanotrack completely reconstructed via chirality barrier," *National Science Review*, vol. 9, no. 12, Feb. 2022, ISSN: 2053-714X. DOI: 10.1093/nsr/nwac021. [Online]. Available: http://dx.doi.org/10.1093/nsr/nwac021.

[182] X. Gong, H. Y. Yuan, and X. R. Wang, "Current-driven skyrmion motion in granular films," *Phys. Rev. B*, vol. 101, p. 064 421, 6 Feb. 2020. DOI: 10.1103/PhysRevB.101.064421. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.101.064421.

[183] Y. Huang, W. Kang, X. Zhang, Y. Zhou, and W. Zhao, "Magnetic skyrmion-based synaptic devices," *Nanotechnology*, vol. 28, no. 8, 08LT02, Jan. 2017. DOI: 10.1088/1361-6528/aa5838. [Online]. Available: https://doi.org/10.1088/1361-6528/aa5838.

[184] S. Li, W. Kang, Y. Huang, X. Zhang, Y. Zhou, and W. Zhao, "Magnetic skyrmion-based artificial neuron device," *Nanotechnology*, vol. 28, no. 31, 31LT01, Jul. 2017. DOI: 10.1088/1361-6528/aa7af5. [Online]. Available: https://doi.org/10.1088/1361-6528/aa7af5.

[185] K. M. Song *et al.*, "Skyrmion-based artificial synapses for neuromorphic computing," *Nature Electronics*, vol. 3, no. 3, pp. 148–155, Mar. 2020, ISSN: 2520-1131. DOI: 10.1038/s41928-020-0385-0. [Online]. Available: https://doi.org/10.1038/s41928-020-0385-0.

[186] X. Zhang, M. Ezawa, and Y. Zhou, "Magnetic skyrmion logic gates: Conversion, duplication and merging of skyrmions," *Scientific Reports*, vol. 5, no. 1, Mar. 2015, ISSN: 2045-2322. DOI: 10.1038/srep09400. [Online]. Available: http://dx.doi.org/10.1038/srep09400.

[187] Z. Zhou, U. Guin, P. Li, and V. D. Agrawal, "Defect characterization and testing of skyrmion-based logic circuits," in *2021 IEEE 39th VLSI Test Symposium*

(VTS), IEEE, Apr. 2021. DOI: 10.1109/vts50974.2021.9441054. [Online]. Available: http://dx.doi.org/10.1109/VTS50974.2021.9441054.

[188] A. Crépieux and C. Lacroix, "Dzyaloshinsky–moriya interactions induced by symmetry breaking at a surface," *Journal of Magnetism and Magnetic Materials*, vol. 182, no. 3, pp. 341–349, 1998, ISSN: 0304-8853. DOI: https://doi.org/10.1016/S0304-8853(97)01044-5. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304885397010445.

[189] A. Thiaville, S. Rohart, É. Jué, V. Cros, and A. Fert, "Dynamics of dzyaloshinskii domain walls in ultrathin magnetic films," *EPL (Europhysics Letters)*, vol. 100, no. 5, p. 57 002, Dec. 2012. DOI: 10.1209/0295-5075/100/57002. [Online]. Available: https://doi.org/10.1209/0295-5075/100/57002.

[190] W. Jiang, G. Chen, K. Liu, J. Zang, S. Velthuis, and A. Hoffmann, "Skyrmions in magnetic multilayers," *Physics Reports*, vol. 704, Sep. 2017. DOI: 10.1016/j.physrep.2017.08.001.

[191] P. Milde *et al.*, "Unwinding of a skyrmion lattice by magnetic monopoles," *Science*, vol. 340, no. 6136, pp. 1076–1080, 2013, ISSN: 0036-8075. DOI: 10.1126/science.1234657. eprint: https://science.sciencemag.org/content/340/6136/1076.full.pdf. [Online]. Available: https://science.sciencemag.org/content/340/6136/1076.

[192] P. C. Carvalho *et al.*, "Correlation of interface interdiffusion and skyrmionic phases," en, *Nano Lett.*, vol. 23, no. 11, pp. 4854–4861, Jun. 2023.

[193] A. W. J. Wells, P. M. Shepley, C. H. Marrows, and T. A. Moore, "Effect of interfacial intermixing on the dzyaloshinskii-moriya interaction in pt/co/pt," *Phys. Rev. B*, vol. 95, p. 054 428, 5 Feb. 2017. DOI: 10.1103/PhysRevB.95.054428. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.95.054428.

[194] L. Rózsa, E. Simon, K. Palotás, L. Udvardi, and L. Szunyogh, "Complex magnetic phase diagram and skyrmion lifetime in an ultrathin film from atomistic simulations," *Phys. Rev. B*, vol. 93, p. 024 417, 2 Jan. 2016. DOI: 10.1103/PhysRevB.93.024417. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.93.024417.

[195] J. Hagemeister, N. Romming, K. von Bergmann, E. Y. Vedmedenko, and R. Wiesendanger, "Stability of single skyrmionic bits," *Nature Communications*, vol. 6, no. 1, p. 8455, Oct. 2015, ISSN: 2041-1723. DOI: 10.1038/ncomms9455. [Online]. Available: https://doi.org/10.1038/ncomms9455.

[196] H. Oike *et al.*, "Interplay between topological and thermodynamic stability in a metastable magnetic skyrmion lattice," *Nature Physics*, vol. 12, p. 62, Oct. 2015. [Online]. Available: https://doi.org/10.1038/nphys3506.

[197] W. Jiang *et al.*, "Magnetism. blowing magnetic skyrmion bubbles," en, *Science*, vol. 349, no. 6245, pp. 283–286, Jul. 2015.

[198] M. Garst, J. Waizner, and D. Grundler, "Collective spin excitations of helices and magnetic skyrmions: Review and perspectives of magnonics in non-centrosymmetric magnets," *J. Phys. D Appl. Phys.*, vol. 50, no. 29, p. 293 002, Jul. 2017.

[199] X. Yu *et al.*, "Variation of skyrmion forms and their stability in mnsi thin plates," *Phys. Rev. B*, vol. 91, p. 054 411, 5 Feb. 2015. DOI: 10.1103/PhysRevB.91.054411. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.91.054411.

[200] J. Zázvorka *et al.*, "Skyrmion lattice phases in thin film multilayer," *Advanced Functional Materials*, vol. 30, no. 46, p. 2 004 037, 2020. DOI: https://doi.org/10.1002/adfm.202004037. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.202004037. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.202004037.

[201] J. Brandão, D. A. Dugato, R. L. Seeger, J. C. Denardin, T. J. A. Mori, and J. C. Cezar, "Observation of magnetic skyrmions in unpatterned symmetric multilayers at room temperature and zero magnetic field," en, *Sci. Rep.*, vol. 9, no. 1, p. 4144, Mar. 2019.

[202] A. S. Ahmed *et al.*, "Chiral bobbers and skyrmions in epitaxial fege/si(111) films," *Phys. Rev. Mater.*, vol. 2, p. 041 401, 4 Apr. 2018. DOI: 10.1103/PhysRevMaterials.2.041401. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevMaterials.2.041401.

[203] P. Panjan, A. Drnovšek, P. Gselman, M. Čekada, and M. Panjan, "Review of growth defects in thin films prepared by PVD techniques," en, *Coatings*, vol. 10, no. 5, p. 447, May 2020.

[204] J.-V. Kim and M.-W. Yoo, "Current-driven skyrmion dynamics in disordered films," en, *Appl. Phys. Lett.*, vol. 110, no. 13, p. 132 404, Mar. 2017.

[205] W. Koshibae and N. Nagaosa, "Theory of current-driven skyrmions in disordered magnets," en, *Sci. Rep.*, vol. 8, no. 1, Apr. 2018.

[206] A. A. Ivanov and V. A. Orlov, "A comparative analysis of the mechanisms of pinning of a domain wall in a nanowire," en, *Phys. Solid State*, vol. 53, no. 12, pp. 2441–2449, Dec. 2011.

[207] M. Kläui *et al.*, "Direct observation of domain-wall pinning at nanoscale constrictions," en, *Appl. Phys. Lett.*, vol. 87, no. 10, p. 102 509, Sep. 2005.

[208] D. Suess, C. Vogler, F. Bruckner, P. Heistracher, F. Slanovc, and C. Abert, "Spin torque efficiency and analytic error rate estimates of skyrmion racetrack memory," en, *Sci. Rep.*, vol. 9, no. 1, p. 4827, Mar. 2019.

[209] J. Sampaio, V. Cros, S. Rohart, A. Thiaville, and A. Fert, "Nucleation, stability and current-induced motion of isolated magnetic skyrmions in nanostructures," en, *Nat. Nanotechnol.*, vol. 8, no. 11, pp. 839–844, Nov. 2013.

[210] J. Ding, X. Yang, and T. Zhu, "Manipulating current induced motion of magnetic skyrmions in the magnetic nanotrack," *J. Phys. D Appl. Phys.*, vol. 48, no. 11, p. 115 004, Mar. 2015.

[211] T. Nozaki *et al.*, "Brownian motion of skyrmion bubbles and its control by voltage applications," en, *Appl. Phys. Lett.*, vol. 114, no. 1, p. 012 402, Jan. 2019.

[212] W. Jiang *et al.*, "Direct observation of the skyrmion hall effect," en, *Nat. Phys.*, vol. 13, no. 2, pp. 162–169, Feb. 2017.

[213] X. Gong, H. Y. Yuan, and X. R. Wang, "Current-driven skyrmion motion in granular films," en, *Phys. Rev. B.*, vol. 101, no. 6, Feb. 2020.

[214] A. Salimath, A. Abbout, A. Brataas, and A. Manchon, "Current-driven skyrmion depinning in magnetic granular films," en, *Phys. Rev. B.*, vol. 99, no. 10, Mar. 2019.

[215] Z. Zhou, U. Guin, P. Li, and V. D. Agrawal, "Defect characterization and testing of skyrmion-based logic circuits," in *2021 IEEE 39th VLSI Test Symposium (VTS)*, San Diego, CA, USA: IEEE, Apr. 2021.

[216] G. Pradhan and S. Bedanta, *Effect of random anisotropy in stabilization of topological chiral textures*, 2020. arXiv: 2003.12481 [cond-mat.mes-hall].

[217] M. N. Potkina, I. S. Lobanov, and V. M. Uzdin, "Nonmagnetic impurities in skyrmion racetrack memory," *Nanosistemy Fiz. Him. Mat.*, vol. 11, no. 6, pp. 628–635, Dec. 2020.

[218] C. Navau, N. Del-Valle, and A. Sanchez, "Interaction of isolated skyrmions with point and linear defects," en, *J. Magn. Magn. Mater.*, vol. 465, pp. 709–715, Nov. 2018.

[219] D. Stosic, T. B. Ludermir, and M. V. Milošević, "Pinning of magnetic skyrmions in a monolayer co film on pt(111): Theoretical characterization and exemplified utilization," *Phys. Rev. B.*, vol. 96, no. 21, Dec. 2017.

[220] K. Everschor-Sitte, M. Sitte, T. Valet, A. Abanov, and J. Sinova, "Skyrmion production on demand by homogeneous DC currents," *New J. Phys.*, vol. 19, no. 9, p. 092 001, Sep. 2017.

[221] S.-Z. Lin, C. Reichhardt, C. D. Batista, and A. Saxena, "Particle model for skyrmions in metallic chiral magnets: Dynamics, pinning, and creep," *Phys. Rev. B Condens. Matter Mater. Phys.*, vol. 87, no. 21, Jun. 2013.

[222] C. Hanneken, A. Kubetzka, K. von Bergmann, and R. Wiesendanger, "Pinning and movement of individual nanoscale magnetic skyrmions via defects," *New J. Phys.*, vol. 18, no. 5, p. 055 009, May 2016.

[223] Y.-H. Liu and Y.-Q. Li, "A mechanism to pin skyrmions in chiral magnets," en, *J. Phys. Condens. Matter*, vol. 25, no. 7, p. 076 005, Feb. 2013.

[224] D. Toscano, S. A. Leonel, P. Z. Coura, and F. Sato, "Building traps for skyrmions by the incorporation of magnetic defects into nanomagnets: Pinning and scattering traps by magnetic properties engineering," en, *J. Magn. Magn. Mater.*, vol. 480, pp. 171–185, Jun. 2019.

[225] J. Mulkers, B. Van Waeyenberge, and M. V. Milošević, "Effects of spatially engineered Dzyaloshinskii-Moriya interaction in ferromagnetic films," *Phys. Rev. B.*, vol. 95, no. 14, Apr. 2017.

[226] J. Castell-Queralt, L. González-Gómez, N. Del-Valle, A. Sanchez, and C. Navau, "Accelerating, guiding, and compressing skyrmions by defect rails," en, *Nanoscale*, vol. 11, no. 26, pp. 12 589–12 594, Jul. 2019.

[227] S. Saha *et al.*, "Formation of néel-type skyrmions in an antidot lattice with perpendicular magnetic anisotropy," en, *Phys. Rev. B.*, vol. 100, no. 14, Oct. 2019.

[228] I. G. Arjana, I. Lima Fernandes, J. Chico, and S. Lounis, "Sub-nanoscale atom-by-atom crafting of skyrmion-defect interaction profiles," en, *Sci. Rep.*, vol. 10, no. 1, p. 14 655, Sep. 2020.

[229] S. A. Pathak and R. Hertel, "Geometrically constrained skyrmions," en, *Magnetochemistry*, vol. 7, no. 2, p. 26, Feb. 2021.

[230] V. P. Kravchuk *et al.*, "Multiplet of skyrmion states on a curvilinear defect: Reconfigurable skyrmion lattices," *Phys. Rev. Lett.*, vol. 120, no. 6, Feb. 2018.

[231] S. Luo *et al.*, "Reconfigurable skyrmion logic gates," en, *Nano Lett.*, vol. 18, no. 2, pp. 1180–1184, Feb. 2018.

[232] V. L. Carvalho-Santos *et al.*, "Skyrmion propagation along curved racetracks," en, *Appl. Phys. Lett.*, vol. 118, no. 17, Apr. 2021.

[233] L. González-Gómez, J. Castell-Queralt, N. Del-Valle, A. Sanchez, and C. Navau, "Analytical modeling of the interaction between skyrmions and extended defects," en, *Phys. Rev. B.*, vol. 100, no. 5, Aug. 2019.

[234] J. Müller and A. Rosch, "Capturing of a magnetic skyrmion with a hole," *Phys. Rev. B Condens. Matter Mater. Phys.*, vol. 91, no. 5, Feb. 2015.

[235] J. Castell-Queralt, L. González-Gómez, N. Del-Valle, and C. Navau, "Deterministic approach to skyrmionic dynamics at nonzero temperatures: Pinning sites and racetracks," en, *Phys. Rev. B.*, vol. 101, no. 14, Apr. 2020.

[236] M. Chauwin *et al.*, "Skyrmion logic system for large-scale reversible computation," en, *Phys. Rev. Appl.*, vol. 12, no. 6, Dec. 2019.

[237] M. T. Birch *et al.*, "History-dependent domain and skyrmion formation in 2D van der waals magnet Fe3GeTe2," en, *Nat. Commun.*, vol. 13, no. 1, p. 3035, May 2022.

[238] M. T. Birch *et al.*, "Control of stripe, skyrmion and skyrmionium formation in the 2D magnet Fe3xGeTe2 by varying composition," *2d Mater.*, vol. 11, no. 2, p. 025 008, Apr. 2024.

[239] K. S. Burch, D. Mandrus, and J.-G. Park, "Magnetism in two-dimensional van der waals materials," en, *Nature*, vol. 563, no. 7729, pp. 47–52, Nov. 2018.

[240] C. Gong and X. Zhang, "Two-dimensional magnetic crystals and emergent heterostructure devices," en, *Science*, vol. 363, no. 6428, eaav4450, Feb. 2019.

[241] H. Li, S. Ruan, and Y.-J. Zeng, "Intrinsic van der waals magnetic materials from bulk to the 2D limit: New frontiers of spintronics," en, *Adv. Mater.*, vol. 31, no. 27, e1900065, Jul. 2019.

[242] M. Gibertini, M. Koperski, A. F. Morpurgo, and K. S. Novoselov, "Magnetic 2D materials and heterostructures," en, *Nat. Nanotechnol.*, vol. 14, no. 5, pp. 408–419, May 2019.

[243] Z. Fei *et al.*, "Two-dimensional itinerant ferromagnetism in atomically thin Fe3GeTe2," en, *Nat. Mater.*, vol. 17, no. 9, pp. 778–782, Sep. 2018.

[244] H. J. Deiseroth, K. Aleksandrov, C. Reiner, L. Kienle, and R. K. Kremer, "Fe3gete2 and ni3gete2 – two new layered transition-metal compounds: Crystal structures, hrtem investigations, and magnetic and electrical properties," *European Journal of Inorganic Chemistry*, vol. 2006, pp. 1561–1567, 2006. [Online]. Available: https://api.semanticscholar.org/CorpusID:97260784.

[245] H. L. Zhuang, P. R. C. Kent, and R. G. Hennig, "Strong anisotropy and magnetostriction in the two-dimensional stoner ferromagnet Fe_3GeTe_2," *Phys. Rev. B*, vol. 93, p. 134 407, 13 Apr. 2016. DOI: 10.1103/PhysRevB.93.134407. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.93.134407.

[246] C. Tan *et al.*, "Hard magnetic properties in nanoflake van der waals Fe3GeTe2," en, *Nat. Commun.*, vol. 9, no. 1, Dec. 2018.

[247] H. Wang *et al.*, "Characteristics and temperature-field-thickness evolutions of magnetic domain structures in van der waals magnet Fe3GeTe2 nanolayers," en, *Appl. Phys. Lett.*, vol. 116, no. 19, p. 192 403, May 2020.

[248] B. Ding *et al.*, "Observation of magnetic skyrmion bubbles in a van der waals ferromagnet Fe3GeTe2," en, *Nano Lett.*, vol. 20, no. 2, pp. 868–873, Feb. 2020.

[249] A. Kotani, H. Nakajima, K. Harada, Y. Ishii, and S. Mori, "Field-temperature phase diagram of magnetic bubbles spanning charge/orbital ordered and metallic phases in $La_{1-x}Sr_xMnO_3$ $(x = 0.125)$," *Phys. Rev. B*, vol. 95, p. 144 403, 14 Apr. 2017. DOI: 10.1103/PhysRevB.95.144403. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.95.144403.

[250] I. Lemesh *et al.*, "Current-induced skyrmion generation through morphological thermal transitions in chiral ferromagnetic heterostructures," en, *Adv. Mater.*, vol. 30, no. 49, e1805461, Dec. 2018.

[251] R. Gruber *et al.*, "Skyrmion pinning energetics in thin film systems," en, *Nat. Commun.*, vol. 13, no. 1, p. 3144, Jun. 2022.

[252] H. Yin *et al.*, "Defect-correlated skyrmions and controllable generation in perpendicularly magnetized CoFeB ultrathin films," en, *Appl. Phys. Lett.*, vol. 119, no. 6, p. 062 402, Aug. 2021.

[253] J. C. Gallagher *et al.*, "Robust zero-field skyrmion formation in fege epitaxial thin films," *Phys. Rev. Lett.*, vol. 118, p. 027 201, 2 Jan. 2017. DOI: 10.1103/PhysRevLett.118.027201. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.118.027201.

[254] Y. Imry and M. Wortis, "Influence of quenched impurities on first-order phase transitions," *Phys. Rev. B Condens. Matter*, vol. 19, no. 7, pp. 3580–3585, Apr. 1979.

[255] S. B. Roy, "First order magneto-structural phase transition and associated multi-functional properties in magnetic solids," en, *J. Phys. Condens. Matter*, vol. 25, no. 18, p. 183 201, May 2013.

[256] P. N. Timonin, "Thermodynamics of a first-order phase transition smeared by frozen disorder," *Phys. Rev. B Condens. Matter Mater. Phys.*, vol. 69, no. 9, Mar. 2004.

[257] D. Linsley, J. Kim, V. Veerabadran, C. Windolf, and T. Serre, "Learning long-range spatial dependencies with horizontal gated recurrent units," in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, ser. NIPS'18, Montréal, Canada: Curran Associates Inc., 2018, pp. 152–164.

[258] T. Lüddecke and A. S. Ecker, "CNNs efficiently learn long-range dependencies," in *NeurIPS 2020 Workshop SVRHM*, 2020. [Online]. Available: https://openreview.net/forum?id=dPwyQnHUVvw.

[259] L. Wang, "Discovering phase transitions with unsupervised learning," *Phys. Rev. B*, vol. 94, p. 195 105, 19 Nov. 2016. DOI: 10.1103/PhysRevB.94.195105. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.94.195105.

[260] W. Hu, R. R. P. Singh, and R. T. Scalettar, "Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination," *Phys. Rev. E*, vol. 95, p. 062 122, 6 Jun. 2017. DOI: 10.1103/PhysRevE.95.062122. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.95.062122.

[261] S. J. Wetzel, "Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders," *Phys. Rev. E*, vol. 96, p. 022 140, 2 Aug. 2017. DOI: 10.1103/PhysRevE.96.022140. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.96.022140.

[262] C. Alexandrou, A. Athenodorou, C. Chrysostomou, and S. Paul, "The critical temperature of the 2D-Ising model through deep learning autoencoders," en, *Eur. Phys. J. B*, vol. 93, no. 12, Dec. 2020.

[263] N. Walker, K.-M. Tam, and M. Jarrell, "Deep learning on the 2-dimensional ising model to extract the crossover region with a variational autoencoder," en, *Sci. Rep.*, vol. 10, no. 1, p. 13 047, Aug. 2020.

[264] D. Yevick, "Variational autoencoder analysis of ising model statistical distributions and phase transitions," en, *Eur. Phys. J. B*, vol. 95, no. 3, Mar. 2022.

[265] H. Y. Kwon *et al.*, "Magnetic hamiltonian parameter estimation using deep learning techniques," en, *Sci. Adv.*, vol. 6, no. 39, eabb0872, Sep. 2020.

[266] D. Wang *et al.*, "Machine learning magnetic parameters from spin configurations," en, *Adv. Sci. (Weinh.)*, vol. 7, no. 16, p. 2 000 566, Aug. 2020.

[267] M. Kawaguchi *et al.*, "Determination of the Dzyaloshinskii-Moriya interaction using pattern recognition and machine learning," en, *Npj Comput. Mater.*, vol. 7, no. 1, Jan. 2021.

[268] W. S. Lee, T. Song, and K.-M. Kim, "Deep learning methods for hamiltonian parameter estimation and magnetic domain image generation in twisted van der waals magnets," 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:267750189.

[269] Z. Tang *et al.*, "Thickness dependence of magnetic anisotropy and domains in amorphous Co40Fe40B20 thin films grown on PET flexible substrates," en, *J. Magn. Magn. Mater.*, vol. 426, pp. 444–449, Mar. 2017.

[270] D. D. Lam *et al.*, "MgO overlayer thickness dependence of perpendicular magnetic anisotropy in CoFeB thin films," en, *J. Korean Phys. Soc.*, vol. 62, no. 10, pp. 1461–1464, May 2013.

[271] C. Eyrich, "Exchange Stiffness in Thin-Film Cobalt Alloys," Ph.D. dissertation, Simon Fraser University, Canada, Jan. 2012.

[272] V. Ramasubramani, B. D. Dice, E. S. Harper, M. P. Spellings, J. A. Anderson, and S. C. Glotzer, "Freud: A software suite for high throughput analysis of particle simulation data," *Computer Physics Communications*, vol. 254, p. 107 275, 2020, ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2020.107275. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010465520300916.

[273] A. S. Silva *et al.*, "Dynamical behaviour of ultrathin [CoFeB (tCoFeB)/Pd] films with perpendicular magnetic anisotropy," en, *Sci. Rep.*, vol. 11, no. 1, p. 43, Jan. 2021.

[274] A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard artifacts," *Distill*, 2016. DOI: 10.23915/distill.00003. [Online]. Available: http://distill.pub/2016/deconv-checkerboard.

[275] D. B. Lee, H. G. Yoon, S. M. Park, J. W. Choi, H. Y. Kwon, and C. Won, "Estimating the effective fields of spin configurations using a deep learning technique," en, *Sci. Rep.*, vol. 11, no. 1, p. 22 937, Nov. 2021.