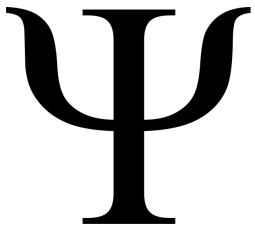

Distinguishing phases and detecting local and non-local order using t-SNE and Monte Carlo methods

Matthew Duschenes

An essay submitted
for partial fulfilment of
Perimeter Scholars International

June, 2018

Distinguishing phases and detecting local and non-local order using t-SNE and Monte Carlo methods

Matthew Duschenes

Supervisors: Prof. Roger Melko and Dr. Lauren Hayward Sierens

The use of computer simulations in condensed matter and many-body physics is in the process of expanding from traditional numerical calculations such as Monte Carlo sampling, to incorporate the emerging field of machine learning and the training of algorithms, with the continued objective of better understanding the underlying patterns and structure in systems with many degrees of freedom. A subset of learning methods include dimensional reduction techniques, where higher dimensional datasets are reduced to machine-learned lower dimensional representations. These techniques have recently been applied by J. Carrasquilla and R. Melko (Nature Physics **13**, 431 (2017)) to detect phases in systems with continuous phase transitions in 2 spatial dimensions. We extend this method of machine-learned dimensional reduction by applying Principal Component Analysis (PCA) and t-distributed Stochastic Neighbour Embedding (t-SNE) to both a q-state Pott's model, which has a first order phase transition for $q > 4$, as well as a classical $\mathbb{Z}_2$ gauge theory, which has topological order. Both PCA and t-SNE methods are shown to differentiate phases in models with first order transitions well, and each method reveals distinct features in the datasets. The linear PCA method is further shown to be inadequate at identifying phases in models with topological order, whereas the non-linear t-SNE method is able to distinguish between phases in such systems with non-local order, by exploiting more of the underlying model structure during the learning process. Future objectives are discussed in applying these promising methods to observing the coexistence of phases directly at first order transitions, as well as investigating further the ability of these methods to represent a model's underlying local or non-local order.

1 Introduction

Simulations in physics have the objective of finding, representing and sampling possible states of a given system, in such a way that any inferred results can be considered valid alternatives to actual experiment. Specifically in many-body physics problems, where the number of states scales exponentially with the system size, effective sampling is essential to understanding large, interacting systems. The established field of Monte Carlo sampling has shown to be particularly effective at probing a wide range of models, from canonical Ising models, to lattice gauge theory models, to quantum fermionic models [1].

The main guiding principal behind the ability of these simulations to deliver accurate and precise results is their use of the underlying physical phenomena and symmetries present in the model being studied itself. The implementation and study of these simulation methods therefore offers insight into not only how the system states can be sampled, but why particular states manifest themselves like they do.

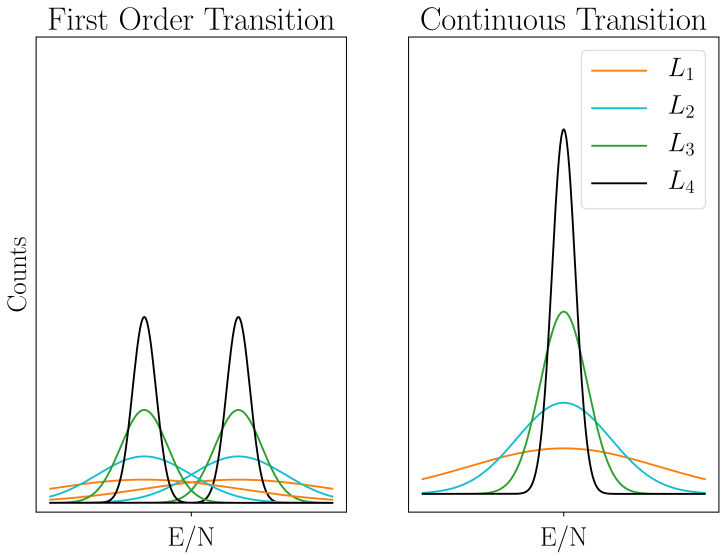

Figure 1: Scaling of Boltzmann distributed energy per spin histograms at the critical temperature for first order and continuous transitions. L_i are arbitrary length scales and $L_1 < L_2 < L_3 < L_4$.

Simulations are described in terms of the update procedure used, to transition from a given sampled system state to a subsequent valid state during a single update iteration. The update procedure can be thought of as displacing along a Monte Carlo time-like axis, which for the scope of this work, will be discretized. Two defining, and related characterizations of update procedures are the scope of the procedure,

and the resulting correlation between samples. Identical in nature to operators acting on elements of the system, updates are operations that are either local in scope, acting on single system elements during an iteration, or non-local in scope, acting on multiple elements [2]. The correlation between samples is related to both individual updates, and the underlying correlations between system elements.

The systems to be studied in this work are generalizations of the Ising spin model on a lattice, with the states of the system being characterized by a Boltzmann distribution, and with correlations between sites on the lattice described by a correlation length ξ. These systems have distinct phases, characterized by a specific *order* parameter σ that differs on either side of a defined critical point. Models can be classified by whether their order parameters are local and can be calculated solely with local measurements of individual lattice sites, or whether their order parameters are non-local and are calculated with measurements across the entire lattice.

Directly at a critical point the system undergoes a phase transition, and the order of the phase transition is defined differently for local and non-local systems [2]. In local systems, the phase transition order is defined by the order of the derivative at which the energy of the system suffers a discontinuity. In non-local systems, the phase transition order is defined to be topological, where the order parameter can be calculated using different specific sections of the lattice, and will always yield the same value. In spin models with continuous phase transitions, the second or a higher derivative of the energy diverges, leading to a $\xi \to \infty$ as $T \to T_c$ behaviour [1]. In first order phase transitions, the first derivative of the energy diverges, and ξ grows large, but remains finite at some ξ_c.

This critical behaviour of ξ, for both continuous and first order transitions presents several challenges in numerical simulations around the critical point. Due to numerical constraints, these systems are simulated on finite lattices, leading to finite size effects. The nature of these effects is outside the scope of this work, however there are excellent derivations for the dependence of observables on system size L and methods of using this scaling to calculate critical scaling of other observables, in Landau and Binder's text [1].

However one relevant finite size scaling effect when determining the effectiveness of a given simulation, is the study of the histogram of calculated energy samples around the critical point. These histograms can demonstrate the nature of the phase transition and an example figure is shown in Fig. 1, where L_i are arbitrary length scales, increasing in length with i.

In the thermodynamic limit, first order phase transitions, with their discontinuous first derivatives of the energy, give rise to two separate sharply peaked energy histograms, representing the two phases on either side of the critical point. The finite, but large divergence of ξ for these first order systems signifies that the length scale of the system at the critical point in small system sizes may in fact be $\xi_c >> L$. ξ will therefore appear almost the same as the infinite ξ would in the case of a con-

tinuous transition [1] and t he two peaks of the histogram will merge closer together, leading to the transition appearing almost continuous. When the system is cooled or heated at temperatures near T_c, there may also be hysteresis effects present in the histograms, as the system transitions between energy peaks [1].

However continuous transitions, with their continuous first derivatives of the energy, give rise to singly peaked energy histograms at the critical point [2]. The infinite divergence of ξ for these transitions, means the length scale will be cut off at the system size and $\xi_c \sim L$. The energy histogram at the critical point remains singly peaked, however become less sharply peaked for smaller systems sizes, due to simulations being more likely to access less likely states during the simulation [2].

In addition to finite size effects, the nature of the divergence of ξ around the critical point in continuous systems presents the additional effect of critical slowing down, where the effective number of samples from the simulation required to achieve statistically uncorrelated results grows as a power of ξ. Global update procedures have been shown to limit many of these effects, however proving the existence of a computationally feasible global update scheme has been shown to be difficult, and sometimes may not even exist [3].

The main objective in this field of simulations is therefore the development of such algorithms that obey and take advantage of the physical constraints of the model being studied, limit the described divergent effects and ultimately produce insight into the systems being studied that cannot be achieved by theoretical or experimental means.

When considering these challenges, the field of machine learning comes to mind. This emerging field similarly represents states of a system, by recognizing patterns and physical structure in models to produce new representations of particular configurations [4]. Specific implementations of machine learning include supervised learning, where known datasets with labels are used to train the parameters of an algorithm for optimal labelling on unknown datasets, as well as unsupervised learning, where unknown datasets are directly processed with algorithms, that converge to optimal representations.

The use of machine learning methods in the field of physics has been so far been concerned with extending and making use of established sampling techniques. In the specific case of using sampled data with known system parameters, such as the temperature, there have been many applications of using supervised learning to confirm results of critical behaviour [4], and to even extend sampling procedures by learning optimization parameters and information about Hamiltonians before generating samples [5, 6]. Phase identification in particular fits well into the philosophy of supervised learning, where phases, or the temperature ranges at which they occur, are explicit labels for datasets and learned labels are easily interpretable. Various supervised algorithms have been shown by Carrasquilla and Melko to be very successful at learning the states of the two-dimensional Ising model, and to ultimately calculate the critical temperature [4].

However in most simulations, data, with large numbers of degrees of freedom that scale exponentially with system size, will be generated that are in unknown states and have no labels for training. Novel applications of unsupervised learning is therefore required to be able to efficiently learn representations that can depict phenomena such as critical behaviour. With these objectives in mind, the specifics of machine learning concepts in this work will be constrained to those in the class of dimensional reduction algorithms, as an extension of the work by Carrasquilla and Melko [4]. These unsupervised algorithms represent the data elements of datasets as elements of some higher dimensional space, and then determine an optimal lower dimensional representation for simplified analysis [7].

It is therefore proposed that the critical behaviour of systems, and the presence of local and non-local order be studied, by representing Monte Carlo generated samples of lattice configurations in a plane using dimensional reduction techniques. In this work, it will be shown that the Monte Carlo update procedures, the representations of data in lower dimensions, and the behaviour of Ising like systems are fundamentally related, and only by studying each of these areas, can the critical behaviour be fully understood.

2 Critical behaviour of spin models

There are many proposed models for discrete spin systems on hyper-lattices, with spin values at each site σ_i, system size L and dimension d, with either local or non-local order. Here, we study three specific classical models: the Ising model, the $\mathbb{Z}_2$ lattice gauge theory, and the q-state Pott's model, each on a $d = 2$ dimensional square lattice with periodic boundary conditions. In particular, first order phase transitions occurring in the q-state Pott's model, when $q > 4$, will be analyzed in depth.

2.1 Ising model

The conventional binary ferromagnetic Ising model in d dimensions has $N = L^d$ lattice sites, and with local nearest neighbour interactions and no external magnetic field has a Hamiltonian of the form:

$$H_{\text{Ising}} = -J \sum_{<ij>} \sigma_i \sigma_j, \tag{1}$$

where $\sigma_i \in \{-1, 1\}$, and J is an energy scale , in units where $k_{\text{b}} \equiv 1$. This Ising model in $d = 2$ dimensions has a continuous phase transition at $\frac{T_{\text{c}}}{J} = \frac{2}{\log 1+\sqrt{2}} \approx 2.2692$ [8], and a local order parameter $\langle\sigma\rangle$ of the net magnetization per spin:

$$\langle\sigma_{\text{Ising}}\rangle = \langle m\rangle = \langle \frac{1}{N} \sum_i \sigma_i \rangle. \tag{2}$$

2.2 $\mathbb{Z}_2$ lattice gauge theory

A classical $\mathbb{Z}_2$ lattice gauge theory in d dimensions has $N = dL^d$ lattice sites, and has non-local interactions $B(p)$ based on defining the lattice structure in terms of plaquettes p. Its Hamiltonian is of the form:

$$H_{\text{gauge}} = -J \sum_{p} B(p), \tag{3}$$

where the interaction terms are the product of spins around each plaquette [9]:

$$B(p) = \prod_{i \in p} \sigma_i. \tag{4}$$

This classical lattice gauge theory has a phase transition directly at $\frac{T_c}{J} = 0$, however there is no local order parameter. Instead a non-local Wilson loop of spins around a closed contour C, typically wrapping around an entire axis of the lattice, must be calculated to determine the presence of topological order:

$$\langle \sigma_{\text{gauge}} \rangle = W(C) = \prod_{i \in C} \sigma_i. \tag{5}$$

When the system is in the topological state at $T = 0$, the Wilson loop takes a constant ± 1 value along all equivalent Wilson loops, such as along all x-axis or all y-axis contours, whereas it is not constant for these same contours in the disordered finite temperature state [9].

2.3 Pott's model

A generalization of the Ising model, is the ferromagnetic q-state Pott's model [10, 11], which assumes a unit magnitude spin, aligned along one of q discrete angles,

$$\sigma_i = \exp\left(i \frac{2\pi s_i}{q} \right), \tag{6}$$

where $s_i \in \{1, 2, ..., q\}$. This model has local order, as discussed below, and it's order parameter is proportional to the net magnetization along one spin direction. This model has been shown experimentally to describe critical behaviour in deposition of monolayers on substrates, such as N_2 onto graphite for $q = 4$, and theoretically to describe polymer vulcanization processes for non-integer $0 < q < 1$ [12].

The denoted *standard* ferromagnetic q-state Pott's model for nearest neighbour interactions is considered, with no external magnetic field [10]:

$$H_{\text{Potts}} = -J \sum_{<ij>} \delta_{s_i s_j}. \tag{7}$$

We define the temperature dependent scale,

$$K = \frac{J}{T} = \beta J. \tag{8}$$

This Hamiltonian has the interesting nearest neighbour interaction term of a delta function that compares strictly the value of the spin states. This indicates there is a symmetry in the system in that each of the q values are equivalent and there is a q fold degeneracy in the ordered phase of the system, where the system could be aligned with any of the q possible angles.

There is a mapping between the Ising model and the $q = 2$ Pott's models of:

$$\begin{aligned} \sigma_i \sigma_j &= 2\delta_{s_i s_j} - 1, \text{and} \\ T_{\mathrm{c}}^{\mathrm{Ising}} &= 2 \cdot T_{\mathrm{c}}^{\mathrm{Potts}}, \end{aligned} \tag{9}$$

and in correspondance with this mapping is the essential Onsager-like result of the critical temperature for the Pott's model in $d = 2$ on a square lattice: [8]:

$$T_{\mathrm{c}} = \frac{1}{\log\left(1 + \sqrt{q}\right)}. \tag{10}$$

This result describes how the $T_{\mathrm{c}}(q)$ monotonically decreases towards $T_{\mathrm{c}} = 0$ as q increases, and indicates that systems described by different q values may act fundamentally differently at a given temperature depending what phase they are in.

2.4 Dimension and q dependence of critical behaviour

The Pott's model and its critical behaviour have been investigated at several points in the 20^{th} century [10, 11, 12, 13] and several theorems, as well as more general conjectures have been developed. The aspect of the critical behaviour relevant to this essay is the relationship between the system dimension d and the q-value, and their resulting effect on the type of phase transition. Figure 2, from Wu [12] depicts the conjectured upper and lower critical dimensions $d_{u,l}(q)$: the respective q dependent dimensions at which fluctuations can be neglected (and therefore mean field theory can apply), and the dimensions at which temperature dependent critical behaviour is not observed.

Using the known critical dimensions for the $q = 2$ Ising model of $d_{\mathrm{u}}(2) = 4$ and $d_{\mathrm{l}}(2) = 1$, as well as renormalization group results that predict $d_{\mathrm{u}}(4) = 2$ and the conjecture that $d_{\mathrm{u}}(3) = 3$, the solid curve for the upper critical dimension shown in Fig. 2 is assumed. Similarly, due to the Ising model exhibiting no critical behaviour when $d \leq 1$, the lower critical dimension is assumed to be $d_{\mathrm{l}}(q) = 1$ for all q. Of particular interest is the nature of the phase transition in the $q - d$ plane. Mean field theory results indicate that for a given system dimension $d \geq d_{\mathrm{u}}(q)$, for $q \leq 2$, transitions are continuous, and for $q > 2$, the transitions are first order [12, 11].

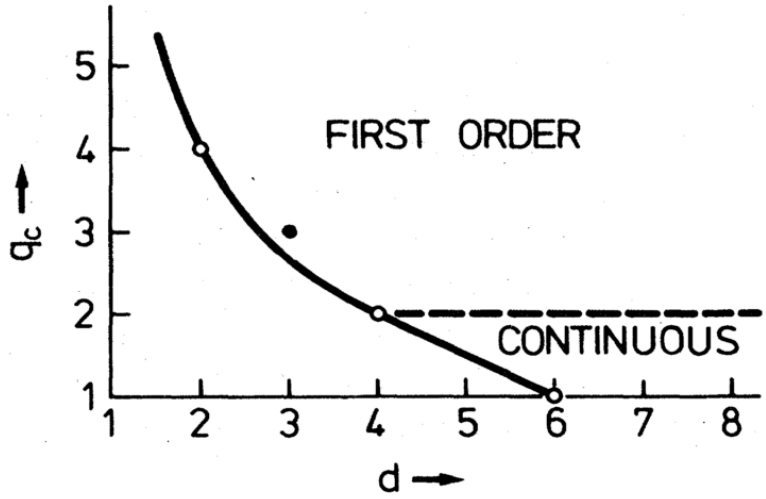

Figure 2: Conjectured upper and lower critical dimensions as a function of the q-value for a ferromagnetic Pott's model, taken from Wu [12]. For a given q, the solid curve denotes that systems with $d \geq d_u(q)$ exhibit behaviour that is independent of fluctuations of the order parameter, and systems with $d \leq d_l(q) = 1$ exhibit no phase transition. The dashed curve denotes that for a given system dimension $d > d_u(q)$, for $q \leq 2$, transitions are continuous, and for $q > 2$, transitions are first order.

This essay will focus on various q-state systems in $d = 2$ dimensions, and when $q > 4$, the transitions studied will be expected to be first order and these systems will therefore not be a member of the Ising model universality class. The form of the Pott's Hamiltonian however indicates that it is still expected that an order parameter related to the net magnetization will behave similarly to the case of the Ising model, where it is non zero at temperatures below T_c, and zero otherwise [8].

2.5 Estimation of observables

For general spin models, common observables to calculate are related to derivatives of a free energy F, written as a function of the partition function Z with total system size N:

$$Z = \sum_{\{\sigma_i\}} \mathrm{e}^{-\beta H(\{\sigma_i\})}, \tag{11}$$

$$F \equiv -\frac{1}{\beta} \log Z. \tag{12}$$

From this definition, the Hamiltonian can be Legendre transformed with pairs of conjugate variables, a fictitious field λ_i and an observable $\mathcal{O}_i$. The expectation of these observables, as the field goes to zero, as well as general susceptibilities can be defined according to the fluctuation theorem [2]. Lower case observables refer to observables per lattice site in the system, and expectation values $\langle\cdot\rangle$ are estimated as the average over Monte Carlo samples in Eq. 24:

$$H \to H - \lambda_i \mathcal{O}_i \tag{13}$$

$$\langle \mathcal{O}_i \rangle = +\frac{1}{\beta} \frac{\partial \log Z}{\partial \lambda_i}\bigg|_{\lambda_i \to 0} \tag{14}$$

$$\chi_i \equiv N\beta \left(\langle \mathcal{O}_i^2 \rangle - \langle O_i \rangle^2 \right). \tag{15}$$

The Pott's model defines its magnetization m_μ, which ranges from 0 to 1, order parameter σ and magnetic susceptibility χ_μ differently from the Ising model, where $\mu \in \{1, ..., q\}$. These quantities take a form related to the interaction term in the Hamiltonian, as well as the degeneracy of the q possible spin states [2]:

$$m_\mu \equiv \frac{1}{N} \sum_i \delta_{s_i,\mu}, \tag{16}$$

$$\langle \sigma^1_{\text{Potts}} \rangle \equiv \frac{q\langle m_1 \rangle - 1}{q-1}, \;\; \langle \sigma^2_{\text{Potts}} \rangle \equiv \frac{1}{q-1} \sum_{\mu=2}^{q} \frac{q\langle m_\mu \rangle - 1}{q-1}, \tag{17}$$

$$\chi^1_{\text{Potts}} \equiv N\beta \sum_{\mu=2}^{q} \left(\langle m_1^2 \rangle - \langle m_1 \rangle^2 \right), \;\; \chi^2_{\text{Potts}} \equiv N\beta \frac{1}{q-1} \sum_{\mu=2}^{q} \left(\langle m_\mu^2 \rangle - \langle m_\mu \rangle^2 \right). \tag{18}$$

When performing simulations, observables per spin of the energy E, specific heat $c = N\beta^2 \left(\langle E^2 \rangle - \langle E \rangle^2 \right)$, and order parameter σ_1 will be calculated to determine the accuracy of the simulations. The behaviour of the energy and order parameter can be evaluated analytically in the limits of the lattice system being completely ordered or disordered, for comparison to simulated expectation values.

When the Pott's model system is completely ordered, the spins will align along one of q directions and so the nearest neighbour interactions will be proportional to the number of spins N and the coordination number, defined as the number of bonds at a given lattice site. In a d dimensional cubic lattice, the coordination number is simply $2d$, and the interactions can be rewritten as a sum over all bonds, with an extra factor of $\frac{1}{2}$, to not double-count bonds.

When the system is completely disordered, there is still a uniform $\frac{1}{q}$ probability of neighbouring spins being equal, and so the sum of non zero interaction terms is proportional to the total expected number of bonds in the system. The limits on the possible expected energy values per spin for the Pott's model are therefore:

$$-dJ \leq \langle E \rangle \leq -\frac{dJ}{q}. \tag{19}$$

Similarly, the limits for the expected order parameter $\langle \sigma_{\text{Potts}} \rangle \equiv \langle \left| \sigma^1_{\text{Potts}} \right| \rangle$, where the absolute value is chosen to avoid expectations averaging to zero, can be deduced. When the system is completely ordered, the spins will align along one of q directions. Choosing the net alignment to be along the $\mu = 1$ direction for σ_1, the expected value will be:

$$\begin{aligned}
\langle \sigma_{\text{Potts}} \rangle &= \frac{q}{q-1} \frac{1}{N_{\text{samples}}} \sum_{t}^{N_{\text{samples}}} \left| m_1^{(t)} - \frac{1}{q} \right| \\
&= \frac{q}{q-1} \left[\frac{1}{N_{\text{samples}}} \cdot \left| 1 - \frac{1}{q} \right| \cdot \frac{N_{\text{samples}}}{q} \; + \; \frac{1}{N_{\text{samples}}} \cdot \left| 0 - \frac{1}{q} \right| \cdot \frac{(q-1)N_{\text{samples}}}{q} \right] \\
&= \frac{2}{q}.
\end{aligned} \tag{20}$$

The disorded state will have an expected order parameter of 0 and so the limits on the expected order parameter per spin for the Pott's mode will be:

$$0 \leq \langle \sigma_{\text{Potts}} \rangle \leq \frac{2}{q}. \tag{21}$$

3 Monte Carlo simulations

3.1 Importance sampling

Due to the $O(q^{L^d})$ scaling of operations required to compute all spin configurations of q state Potts' model systems of size L, the behaviour of these systems must be calculated numerically, using various Monte Carlo site sampling algorithms. These sampling algorithms implement the concept of importance sampling, where the true, or *nominal* distributions $p(X_i = X) \equiv p_i$ for random variables $X_i \sim p$ are estimated using *importance* distributions q_i during expectation value calculations [14]:

$$\langle \mathcal{O} \rangle = \frac{\sum_i \mathcal{O}_i p_i}{\sum_i p_i} = \frac{\sum_i \mathcal{O}_i q_i^{-1} p_i q_i}{\sum_i q_i^{-1} p_i q_i}. \tag{22}$$

These expectation values can then be approximated by

$$\langle \mathcal{O} \rangle \approx \frac{\sum'_i \mathcal{O}_i q_i^{-1} p_i}{\sum'_i q_i^{-1} p_i}, \tag{23}$$

where $\sum'_i$ represents a sum over a reduced subset of the whole probability space, sampled from the q_i distribution. The observables $\mathcal{O}_i$ are now weighted by the

additional factor of $q_i^{-1}p_i$, corresponding to taking into account that these samples are sampled from a distribution that is approximating the sampling from the true distribution. q_i is meant to be chosen such that samples are generated strictly from desired regions of the true distribution's support [14].

The Monte Carlo procedure samples observables from a distribution $q_i = Cp_i$, equal to the true distribution up to a possibly unknown normalization constant C, meaning that states are generated with the same likelihood with which they would appear in experiment. This further yields estimated observables with a minimal sampling variance that have expectations in the form of a simple to calculate arithmetic mean in Eq. 24 [14].

3.2 Correlations in Monte Carlo sampling

The goal of Monte Carlo methods is therefore to stochastically sample configurations of a system in Monte Carlo time-steps $t = \{1, 2, ..., N_{\mathrm{MC}}\}$, in order to estimate the expected value of observables $\mathcal{O}$:

$$\langle \mathcal{O} \rangle \approx \bar{\mathcal{O}} = \frac{1}{N_{\mathrm{MC}}} \sum_{t}^{N_{\mathrm{MC}}} \mathcal{O}_t. \tag{24}$$

The effectiveness of these simulations can be quantified by the precision of $\bar{\mathcal{O}}$, given by its variance:

$$\sigma^2_{\bar{\mathcal{O}}} = \frac{\sigma^2_{\mathcal{O}_j}}{N_{\mathrm{MC}}} \tau^{\mathrm{int}}_{\bar{\mathcal{O}}}, \tag{25}$$

where the integrated and exponential autocorrelation times $\tau_{\bar{\mathcal{O}}}$ can be defined in terms of the correlation function $\mathcal{C}(j)$ for the observables across Monte Carlo time steps:

$$\begin{aligned} \tau^{\mathrm{int}}_{\bar{\mathcal{O}}} &= 1 + 2 \sum_{t}^{N_{\mathrm{MC}}} C(t) \Big(1 - \frac{t}{N_{\mathrm{MC}}} \Big), \\ \mathcal{C}(j) &= \frac{\langle \mathcal{O}_i \mathcal{O}_{i+j} \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle}{\langle \mathcal{O}_i^2 \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle} \sim \exp\left(- \frac{j}{\tau^{\mathrm{exp}}_{\bar{\mathcal{O}}}} \right). \end{aligned} \tag{26}$$

At the critical temperature, the integrated autocorrelation times scale with the spatial correlation length ξ of the system as $\tau^{\mathrm{int}}_{\bar{\mathcal{O}}} \sim \xi^z$ for a positive (critical) exponent z. This critical slowing down effect implies a greater effective $N^{\mathrm{eff}}_{\mathrm{MC}} = \tau^{\mathrm{int}}_{\bar{\mathcal{O}}} \cdot N_{\mathrm{MC}}$ samples are required to achieve the same number of uncorrelated samples [15].

These algorithms must be efficient and generic enough to simulate various spin models, while also determining behaviours and making use of symmetries that are unique to a model. Numerous versions of these algorithms currently exist [3, 16, 17], however they must be adapted for the specific model, and chosen system parameters.

The stochastic Monte Carlo simulations must further satisfy the properties of being ergodic, meaning that every possible state must be accessible within a finite

number of samples, markovian, meaning each iteration's sampling procedure only depends on the previous iteration, and stable, meaning the sampling procedure will produce samples from a distribution $\mathbb{P}_\eta$ which converges to the system's true distribution $\mathbb{P}_\eta^{\text{eqb}}$. In the case of solid state systems, including the spin models being studied, this distribution is the Boltzmann distribution. Satisfying these constraints guarantees that equilibrated final configurations will be generated and it can be shown that each of these constraints are satisfied under the assumption of detailed balance [2]:

$$\frac{w_{\nu\to\eta}}{w_{\eta\to\nu}} = \frac{\mathbb{P}_\eta}{\mathbb{P}_\nu} = \frac{\mathbb{P}_\eta^{\text{eqb}}}{\mathbb{P}_\nu^{\text{eqb}}}, \tag{27}$$

where $w_{\eta\to\nu}$ are transition probabilities between states; $\mathbb{P}_\eta$ are the probability distributions of being in a certain state, which converge to an equilibrium distribution.

A more intuitive description defines the transition probabilities $w_{\eta\to\nu}$ to be the product of the *proposal* probability of performing a specific update $g_{\eta\to\nu}$, such as proposing to flip a certain spin, and the *acceptance* probability of retaining this updated state $a_{\eta\to\nu}$. This splitting of the update procedure into two stages aids in both the design of the algorithms, and determining whether they satisfy detailed balance [2].

The standard local update procedure is the single spin-flip Metropolis algorithm, which chooses any spin to potentially flip with proposal probability $\frac{1}{N}$ and has acceptance probability $a_{\eta\to\nu} = \min(1, \frac{\mathbb{P}_\nu}{\mathbb{P}_\eta})$.

For the given Pott's model on a two dimensional square lattice, several non-local cluster update algorithms that satisfy this detailed balance are appropriate due to the duality between the Boltzmann distribution of the configurations and the concepts of clusters and bonds between nodes in graph theory [17, 18]. This duality is illustrated by the Fortuin-Kastelyn cluster decomposition of the system's partition function Z:

$$Z = \sum_{\{\sigma_i\}} \mathrm{e}^{-\beta H(\{\sigma_i\})} = \sum_{\{\sigma_i\}} \prod_{<ij>} \sum_{n_{ij}} [\delta_{\sigma_i\sigma_j} n_{ij} p + (1 - n_{ij})(1 - p)]. \tag{28}$$

This equivalent form of the partition function assigns activation functions n_{ij} to each bond between pairs of sites in the lattice. For unaligned spins, $n_{ij} = 0$ and for aligned spins, bonds are activated, $n_{ij} = 1$, with probability p, and are deactivated, $n_{ij} = 0$, with probability $1 - p$, where p depends on the form of the Hamiltonian [17]. This duality makes the concept of updating clusters of spins, as opposed to updating individual spins, more intuitive and is used by Fortuin and Kastelyn to provide the mathematical insight into why detailed balance is satisfied [17].

The Wolff cluster update algorithm [16] is selected as the specific cluster update procedure used in this work, where a random site is chosen, and a single cluster is formed from neighbouring sites. Neighbouring sites, and their subsequent neighbours, are added to the cluster if they have an active bond with the chosen site.

The cluster update is then always accepted with acceptance probability 1 [2]. In the case of the Pott's model, the bond activation probability is $p = 1 - \mathrm{e}^{-K}$, where K is the energy scale for nearest neighbour interactions in Eq. 8.

Unlike other cluster update algorithm schemes, such as the Swendsen-Wang algorithm [3] which update multiple clusters per Monte Carlo step, the Wolff algorithm will be faster at higher temperatures when a single cluster will only span small portions of the whole lattice, and for d=2 dimensions, equilibrates within similar timescales [16].

Both local Metropolis and non-local Wolff updates will therefore be implemented to generate sample configurations of the Pott's model at various q, L, and T values above and below T_{c}. These configurations will allow the critical behaviour of the Pott's model to be investigated, and will be used as data in further analysis with subsequent numerical methods.

4 Representations of data using dimensional reduction

Given samples of system configurations, the explicit data of spin values on a lattice may be able to be visually analysed to determine information about the system such as its phase. However, further numerical methods must be implemented to formalize a more robust approach that can do similar pattern recognition by iteratively learning information. Extending the work of Carrasquilla and Melko, we propose using numerical methods called dimensional reduction techniques [4] for direct analysis of the system configurations.

Several numerical methods exist to reduce representations of data from N dimensions, to $\tilde{N} << N$ dimensions, in order to perform visual analysis of patterns in data in $\tilde{N} = 2$ or 3 dimensions. These algorithms are a sub-class of machine learning methods, where training is performed on a dataset in order to learn its structure and to generate an optimal lower dimensional representation. Similar to other machine learning methods, these algorithms contain several hyper-parameters, which must be fine-tuned by the user, are generally unique to the dataset being studied, and are not explicity learned during training. We will focus on the theory and application of Principal Component Analysis (PCA), and t-distributed Stochastic Neighbour Embedding (t-SNE) dimensional reduction techniques.

4.1 Principal component analysis

An established method of dimensional reduction is Principal Component Analysis (PCA), which performs a linear transformation on datasets $X = [x_1, ..., x_M]^T \in \mathbb{R}^{M \times N}$ for M datasets [19]. First, X is converted to $\tilde{X}$, whose columns have zero mean. $\tilde{X}$ is then linearly transformed through an eigenvalue relation to its corre-

sponding diagonalized matrix of its eigenvalues D [19]:

$$X \to \tilde{X} = X - \langle X \rangle_{\text{rows}} \tag{29}$$

$$\tilde{X}^T \tilde{X} = P^T D P, \tag{30}$$

where P is some invertible matrix. D is subsequently sorted such that its diagonal eigenvalue entries $\{\lambda_n\}_{n=1}^{N}$ are in descending order.

Solutions to this eigenvalue problem allow a new transformed representation Y, with its first $\tilde{N} < N$ principal columns being used for analysis:

$$Y = \tilde{X} P. \tag{31}$$

A useful quantity to determine the effectiveness of using a small $\tilde{N} << N$ to represent datasets is the explained variance ratio $r_n \equiv \frac{\lambda_n}{\sum_m^N \lambda_m}$, which depicts the relative importance of each component at representing the dataset [19].

To demonstrate this, we reproduce work done by Wang [19] and consider classical Ising model samples at a range of temperatures above and below $\frac{T_c}{J} \approx 2.269$, as well as classical $\mathbb{Z}_2$ lattice gauge theory samples at either finite temperatures, or at $T = T_c = 0$. Plots of their explained variance ratios for the first 10 components are shown in Fig. 3.

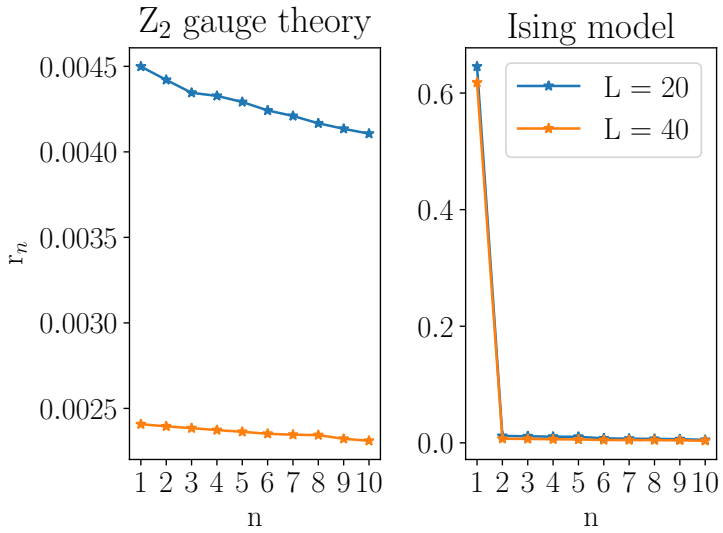

Figure 3: The explained variance ratio r_n for the first 10 principal components of a PCA analysis of classical $\mathbb{Z}_2$ lattice gauge theory and Ising model datasets, for system sizes L = 20, 40.

The Ising model, with its local nearest neighbour interactions, has an underlying structure where the order parameter, and thus the phase of the total system can be

determined from strictly local measurements. The observation that its configurations can be represented almost entirely by its first principal component in Fig. 3 is direct evidence that the lack of global structure in a model translates into redundancy in the information contained in the original configuration representations.

However the gauge theory has non-local structure and therefore the Wilson loop order parameter, and thus the phase of the total system cannot be determined from strictly local measurements. Therefore there is significantly less redundancy in the information contained in the original configuration representations and many of the principal components are required to effectively represent this data. The minimum number of components required is to be determined, however due to the Wilson loop requiring on the order of L measurements, $O(L)$ components is conjectured to be enough to learn about this non-local order. The ability of this linear PCA method to accurately represent models with local or non-local structure is suggested to be directly related to these variance ratio curves, and this can be confirmed by applying PCA to various spin models.

4.2 t-distributed stochastic neighbour embedding

A more recent method of dimensional reduction is the non-linear technique of t-distributed Stochastic neighbour embedding (t-SNE), developed by van der Maaten and Hinton [7]. Input datasets $X \in \mathbb{R}^{M\times N}$ are transformed into Gaussian distributions, based on their distance to all other data points $x_i \in \mathbb{R}^N$:

$$x_i, x_j \to p_{ji} = \frac{\exp\left(\frac{-|x_i - x_j|^2}{2\alpha_i^2}\right)}{\sum_{k\neq l}\exp\left(\frac{-|x_k - x_l|^2}{2\alpha_i^2}\right)}, \tag{32}$$

where the variance α_i is a hyper-parameter to be determined, and $p_{ii} \equiv 0$. Lower dimensional representations $y_i \in \mathbb{R}^{\tilde{N}<N}$ are to be learned through an optimization procedure and are represented as t-distributions [7]:

$$y_i, y_j \to q_{ji} = \frac{1 + |y_i - y_j|^2)^{-1}}{\sum_{k\neq l}(1 + |y_k - y_l|^2)^{-1}}. \tag{33}$$

There are several critical features of these representations that allow t-SNE to be an effective method of dimensional reduction, and for a complete explanation, please refer to [7]. The overall purpose of these representations is for the scaling of distances between datapoints to be preserved, particularly for datapoints that are moderately separated relative to the full dataset range. The lower dimensional representations y_i are not represented as Gaussian distributions due to t-distributions having greater tail amplitudes at larger distances from the centre of the distribution. his allows moderate distances between the higher dimensional datapoints to remain separated by moderate distances in the lower dimensional datapoints [7].

4.3 Training procedures

The hyper-parameters to be fine-tuned for optimal training include standard machine learning parameters such as the number of training iterations and learning rates, as well as an essential parameter of the perplexity $\mathcal{P}$ of the original dataset, defined as the average information content of the Gaussian representation:

$$\mathcal{P}(\alpha_i) = 2^{H(X;\alpha_i)}, \tag{34}$$

where $H(x) = -\sum_{i,j} p_{ji}(\alpha_i) \log p_{ji}(\alpha_i)$ is the entropy of this distribution. A fixed $\mathcal{P}$ is used to determine the variances α_i for the Gaussian distribution, and this fixed information content of p_{ji} relative to the information present in the initially randomized q_{ji} are important initial conditions in the convergence of the algorithm.

The original SNE method involved $p_{ji} \neq p_{ij}$ in their asymmetric form in Eq. 32, and used a cost function $C(y_i)$, chosen to be the sum of the Kullback-Leibler (KL) divergences $KL(p,q) \equiv \sum_i p_i \log\left(\frac{p_i}{q_i}\right)$ over all datapoints:

$$C(y_i)_{\text{asymmetric}} = \sum_i KL(P_i|Q_i), \text{where} \tag{35}$$

$$KL(P_i, Q_i) \equiv \sum_j p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right)$$

However van der Maaten and Hinton made the important development of symmetrizing $p_{ji} \to \frac{p_{ji}+p_{ij}}{2M}$ [7]. This simplifies the cost function to be a KL divergence over the joint distributions p_{ij}, q_{ij} and yields a simple form of its gradient in Eq. 37:

$$\begin{aligned} C(y_i)_{\text{symmetric}} &= KL(P,Q) \\ &= \sum_i \sum_j p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right), \end{aligned} \tag{36}$$

$$\frac{\partial C}{\partial y_i} = 4 \sum_j (q_{ji} - q_{ji})(y_i - y_j)(1 + |y_i - y_j|^2)^{-1}. \tag{37}$$

This cost function is minimized using stochastic gradient descent and other optimization schemes discussed in [7], in order to determine the lower dimensional representations y_i that optimally represent the higher dimensional x_i. A binary search algorithm is used to find the original representation variances α_i.

The simple form of the cost function gradient, and its $y_i - y_j$ terms intuitively explain how the algorithm attempts to find the lower dimensional representations that are as close together as possible. Concurrently, the $\sim 1/|y_i - y_j|^2$ scaling in the t-distributions and then in the gradient for large distances between y_i indicates that when $\tilde{N} = 2$ dimensions, these distributions will be almost invariant to changes

in scale, allowing larger distances to be maintained, and subsequently more distinct clusters to form.

This new form of the gradient is also significantly easier to optimize and ultimately takes advantage of the asymmetry of the KL divergence, ensuring that there is a large cost associated with using small q_{ij} (widely separated y_i) to model large p_{ji} (closely separated x_i), and a smaller cost associated with using large q_{ij} (closely separated y_i) to model small p_{ji} (widely separated x_i) [7].

The KL divergence has interesting properties related to its definition as a measure of relative entropy, which offer insight into why it allows t-SNE to effectively represent the true p_{ij} by the q_{ij}. In particular, defining a distance measure between two probability distributions P and Q with support over Ω as:

$$\delta(P,Q) \equiv \sup_{A\in\Omega} \|P(A) - Q(A)\| = \frac{1}{2}\|P - Q\|_1 = \frac{1}{2}\sum_{w\in\Omega} \|P(w) - Q(w)\|, \tag{38}$$

and citing Pinsker's inequality [20] reveals that minimizing the KL divergence is identical to minimizing the distance between two distributions:

$$\delta(P,Q) \leq \sqrt{\frac{1}{2}KL(P,Q)}. \tag{39}$$

This inequality validates the choice of the KL divergence cost function and accurate optimization will result in lower dimensional representations that have similar information content to the higher dimensional representations.

The formalism behind these two dimensional reduction methods demonstrates that they have the potential to both learn patterns in datasets, such as phases in spin configurations, and to make use of symmetries present in configurations, similar to how non-local Monte Carlo sampling makes use of bonds present between sites. These methods are now applied to various spin models to determine their capabilities, their limitations and how their lower dimensional representations can be interpreted in terms of known critical behaviour.

5 Numerical analysis of models with local and non-local order

5.1 Generation of uncorrelated model configurations

The previous theory of Monte Carlo simulations was applied for a general q-state Pott's model on a square lattice with N sites. The simulation was implemented with Python and C-Python languages, making use of the extensive *numpy* python package for large arrays, while still having the computational speed-up associated with the C language. Monte Carlo simulations were ran for various system parameters to determine the minimal sampling required to produce uncorrelated samples.

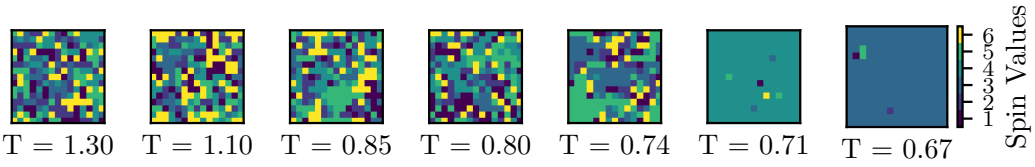

Figure 4: Sample spin configurations for a $q = 6$, $L = 15$ Pott's model, with 85% Metropolis updates and 15% Wolff updates per sweep. $T_c(q = 6) \approx 0.8076$.

Defining a single sweep to be N updates, Monte Carlo simulations consisting of $N_{\text{eqb}} \sim 10^5$ equilibrium sweeps, to ensure accurate configurations, and $N_{\text{meas}} \sim 10^4$ measurement sweeps, sampling every 5 sweeps, were conducted. Initially expectation values of the global Wolff updates were compared to the Metropolis updates to ensure the accurate data was being generated. Final simulations were conducted with each sweep containing an initial phase of some percentage of N Metropolis updates, and then the remaining sweeps being Wolff updates.

The use of Wolff cluster updates improves the ergodicity of the system by flipping between the degenerate ordered states at low temperatures. However, this has the consequence of the simulation not stabilizing to one particular ordered phase, but instead forcing the system into each possible state at some point during the simulation. This results in sampled magnetization values averaging to zero, and the absolute value definition in Eq. 20 is therefore justified in order to calculate what would be observed if the system were to cool naturally into a single phase. Sample final spin configurations for a temperature sweep between 1.3 and 0.67 can be referred to in Fig. 4, for $q = 6$ and $L = 20$. As the temperate decreases below $T_c(q = 6) \approx 0.8076$, the critical behaviour of ordered clusters aligned in one of the q degenerate directions is observed, eventually choosing one particular direction to fully align.

A final calculation to ensure the simulations are producing accurate uncorrelated configurations was to calculate the expected energy, order parameter and specific heat per spin from the generated Monte Carlo samples at each temperature. The resulting curves around the critical temperature for $q = 2$, 3, 4, 6, 10, and $L = 10$ are shown in Fig. 5. The energy curve demonstrates the expected continuous transition for $q = 2$, which then gets significantly steeper, suggesting a first order transition as q increases past 4 until 10. The finite size effects of the specific heat not diverging, but increasing to a large finite value at the critical point for this small $L = 10$ lattice are particularly apparent. The limiting values for energy and magnetization are further shown to agree with the bounds in Equations 19 and 21.

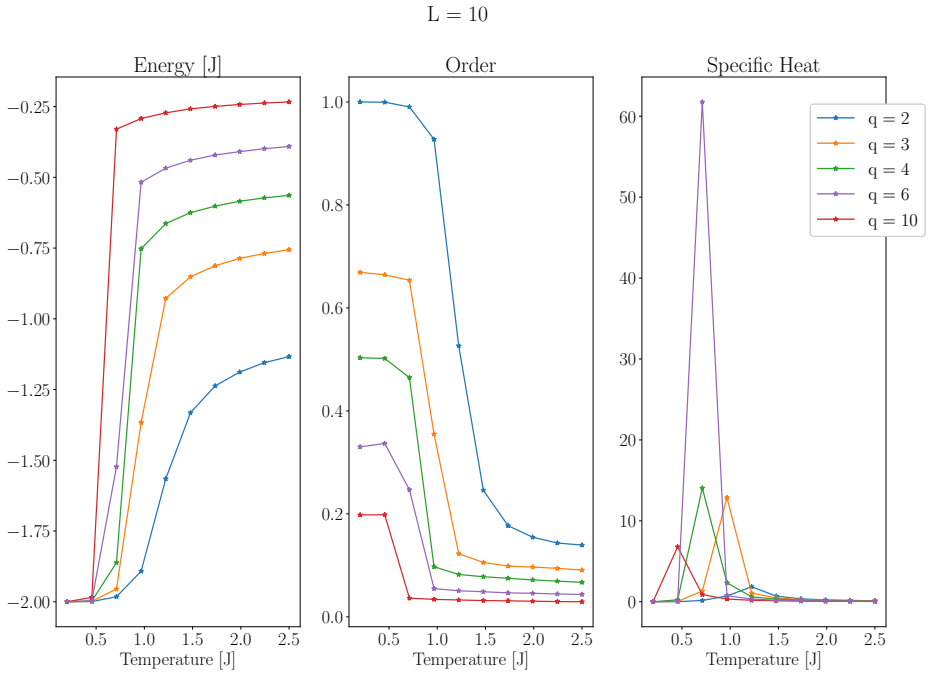

Figure 5: Simulated expectation values of observables per spin, for an $L = 10$, $q = 2, 3, 4, 6, 10$ Pott's model.

5.2 PCA and t-SNE representations of systems with first order phase transitions

The configurations generated for various L and q for the Pott's model, in conjunction with existing spin configurations data for the Ising model and the $\mathbb{Z}_2$ lattice gauge theory are subsequently reduced to $\tilde{N} = 2$ dimensions with PCA and t-SNE. Both the Ising and Pott's model configurations are at a range of temperatures above and below T_c, and the gauge theory configurations are labelled as either being in the zero temperature topological, or high temperature disordered phase.

The PCA and t-SNE algorithms were implemented as modified versions of the original code from [7], which originally elected to first process the input with PCA, and to then select the first N' dimensions for analysis with t-SNE. It also stated that t-SNE is robust to perplexities on the order of 20 - 50. The resulting reduced representations with various algorithm parameters are shown in Figures 6 to 12.

Figures 6 and 7 demonstrate the ability of both PCA and t-SNE to detect degenerate phases around first order phase transitions and are able to distinguish the $q + 1$ clusters. Similar to in the case of the Ising model, the PCA representation is almost one-dimensional and the first principal component contains almost all of the information about the configurations. Therefore even though this Pott's model is

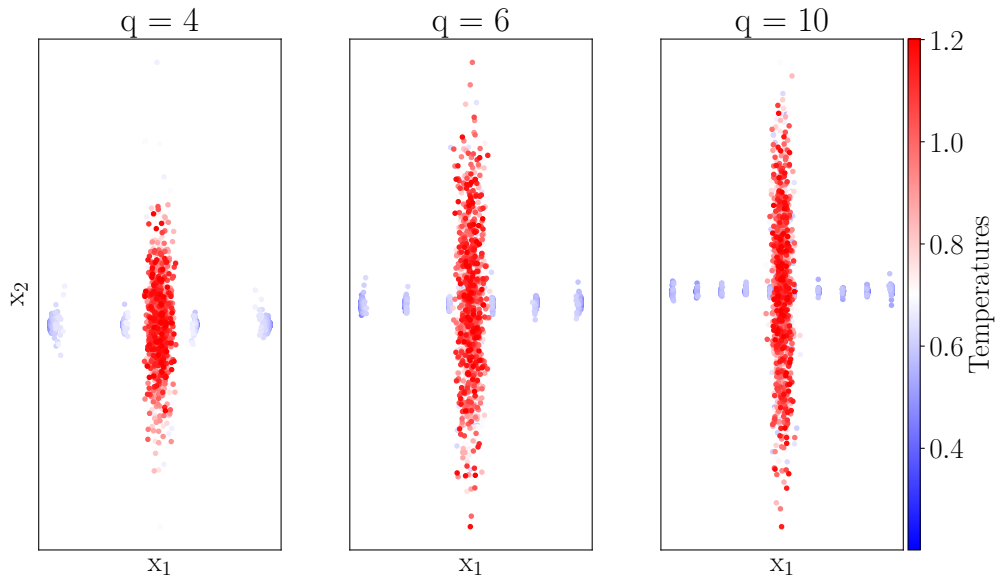

Figure 6: PCA representation of Pott's model configurations, for $L = 50$, $q = 4$, 6, 10.

q-fold degenerate in the ordered phase, it can be inferred that there is still significant redundancy in the original representations; the strictly local, and linear order parameter is still the dominant feature that is being learned by PCA [19].

The application of the t-SNE method to the Pott's model in Fig. 7 also distinguishes the $q+1$ phases fairly distinctly and even reveals structure of a temperature gradient within individual clusters. However there are two ordered spin state clusters that are hidden within the edges of the disordered high temperature cluster. This effect can likely be attributed to the Monte Carlo sampling not taking into account the degeneracy of the q ordered phases. Spin configurations were sampled at equal numbers of temperatures above and below the critical point, meaning each possible low temperature phase had significantly less samples available for training compared to the high temperature phase.

Ensuring that each possible phase of the spin configurations are sampled uniformly and ergodically should improve the amount, and variation of information available for training. Therefore even if there is only local order in this model, this q-fold degeneracy still proves difficult to distinguish with the algorithm parameters selected. The $q = 4$ clusters however are particularly defined, and this can be attributed to there being less degeneracy and more information about each possible cluster available to be learned during training. Further, unlike in PCA, the clusters found are not along one dimension, indicating that t-SNE is not similarly learning the dominant principal components, but a different feature entirely.

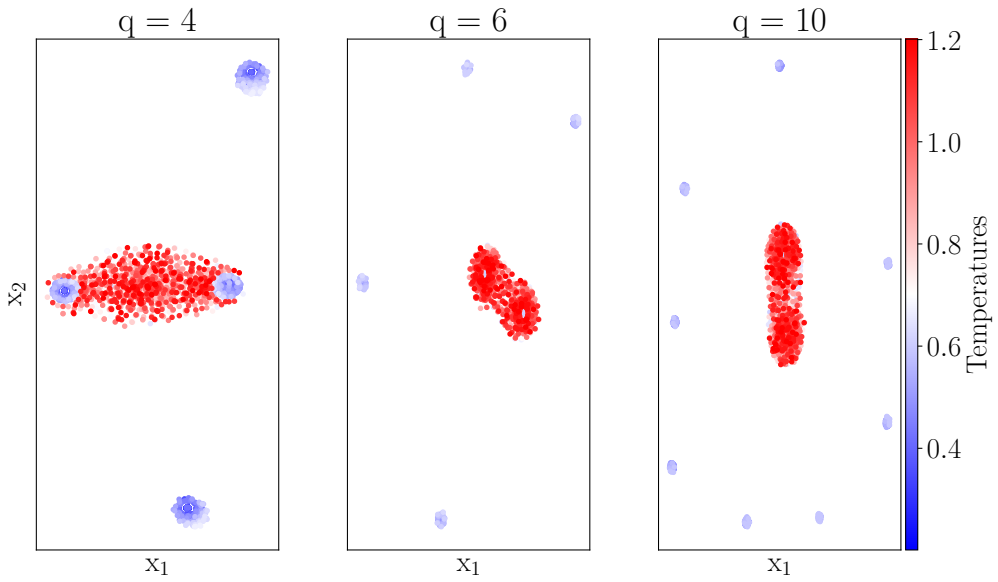

Figure 7: t-SNE representation of Pott's model configurations, with PCA pre-processing of 100 dimensions, for $L = 50$, $q = 4$, 6, 10.

5.3 t-SNE representations of systems with continuous phase transitions

Figures 8 and 9 demonstrate the effect of initially pre-processing datasets with PCA, before performing the t-SNE analysis, in the case of the Ising model with a continuous phase transition. Figure 8, where the first 50 dimensions of the PCA processed data were used in t-SNE and the perplexity is 30.0 is reproduced from work by Carrasquilla and Melko [4]. The produced representations not only show clear distinct clusters for the $2 + 1$ phases, they further show that the elements of individual clusters are aligned along a temperature gradient, indicating that t-SNE is learning information about the temperature dependence of the configurations, in addition to the phase.

Figure 9, where PCA is not initially used, and the perplexity is 40.0, shows significantly more structure of the datasets. Instead of clusters, the spiral behaviour distinguishes the $2 + 1$ phases and reveals a different underlying structure to the dataset, that clustering in PCA, or t-SNE with linear processing is unable to reveal. The spiral structure has interesting features of having two symmetrical spiraling curves, and each curve has two branches that meet at very low temperatures.

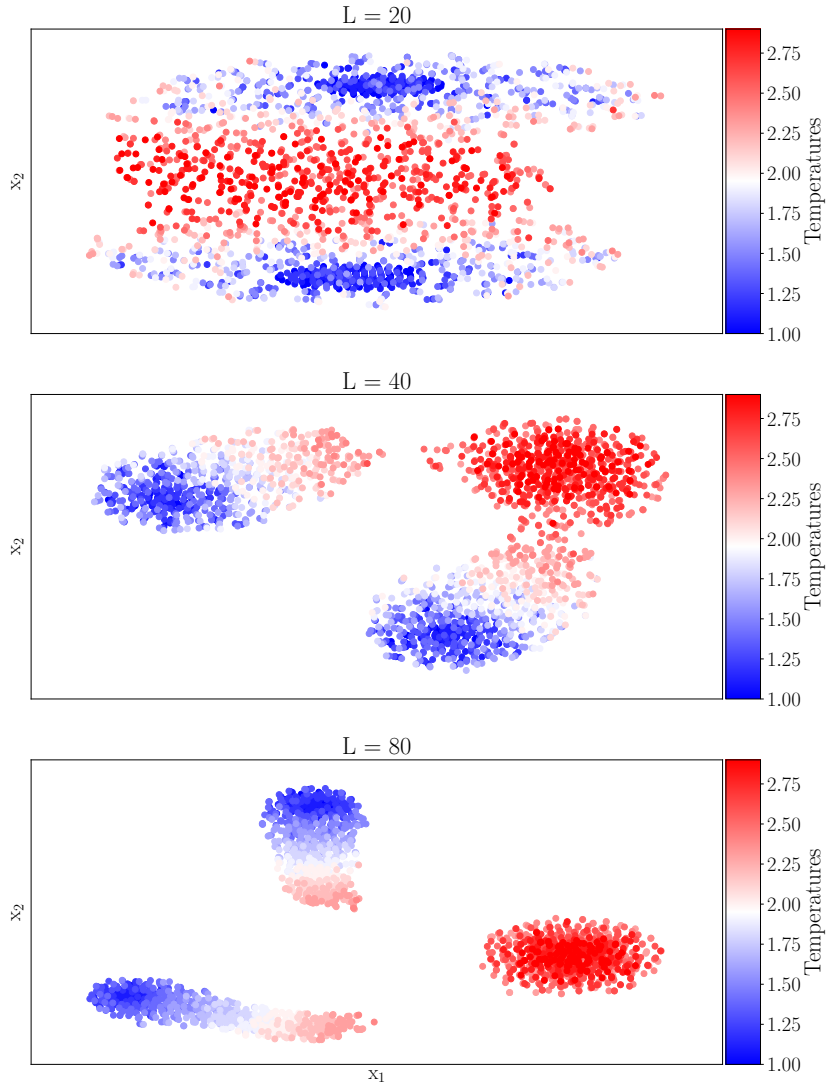

Figure 8: t-SNE representation of Ising model configurations, with pre-processing with PCA of 50 dimensions, for $L = 20, 40, 80$.

Either the two spiraling curves, or the two branches within the curves shown in Fig. 9 could represent the clustering of the two ordered low temperature phases. Further analysis would be required to determine what specific features correspond to the physical states being represented, and, due to these very distinct features, whether there is any evidence of the t-SNE algorithm converging to local minima. The high temperature phase is also shown to wrap around the outer low temperature spirals, suggesting further interaction between the two distinct phases.

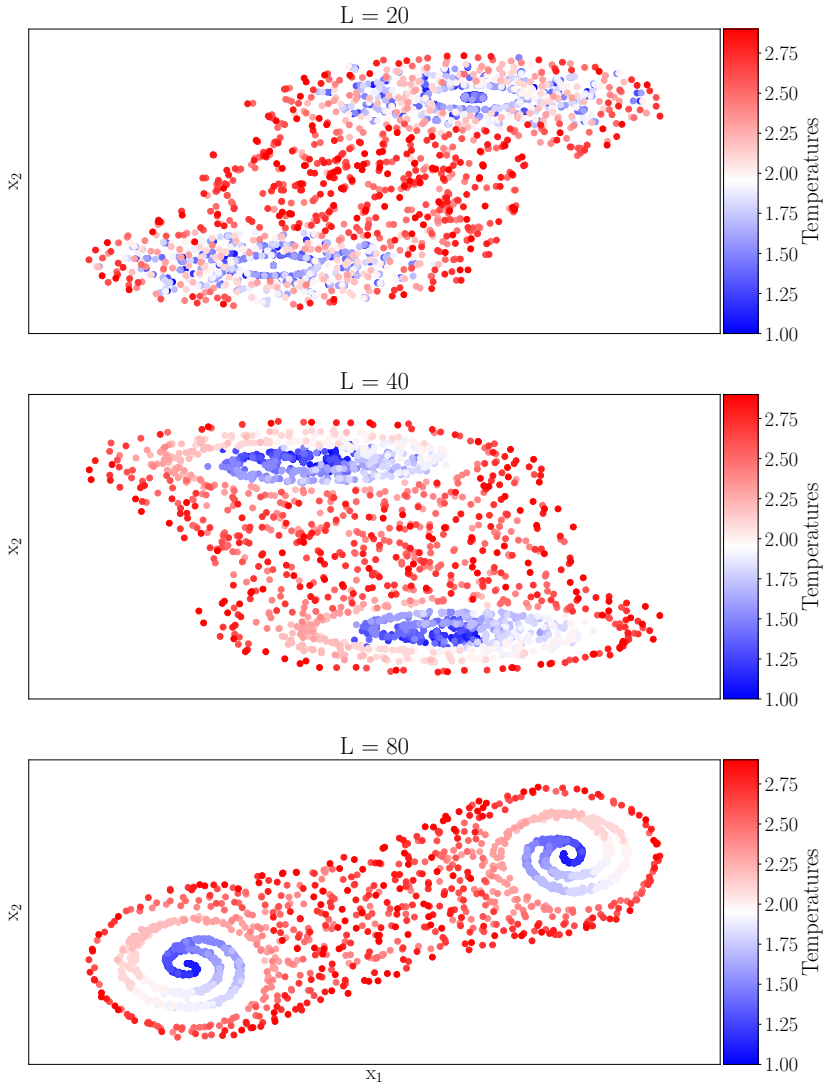

Figure 9: t-SNE representation of Ising model configurations, without pre-processing with PCA, for $L = 20, 40, 80$.

In both of these figures, there are significant finite size effects, and the structures, such as the spirals in Fig. 9, that t-SNE is learning as lower dimensional representations are much clearer for the $L = 80$ configurations. The presence of either clustering, or the spiral structure also suggests that there is a trade-off between the information that t-SNE is able to learn and convey about a particular system, and the ease of interpretability of these lower dimensional representations.

The spiral data structures are very distinct from the previous cluster structures obtained when PCA is involved in the training, and it is not as intuitive as to whether

t-SNE is still trying to learn how to distinguish the phases. It has been shown by Wang, that linear PCA attempts to learn the order parameter for systems with local order, which can likely be attributed to the order parameter being proportional to the net magnetization, a linear combination of local spin measurements [19]. Comparing the structures produced by t-SNE with those by PCA, the one-dimensional clustering in PCA is much easier to interpret as phase differentiation, however it is possible that t-SNE is trying to learn a much more complicated physical property, potentially a higher order correlation function. Therefore methods of interpreting and quantifying the information contained in these new types of structures should be developed, and a relationship between these structures, and which physical parameters are learned should be established.

5.4 PCA and t-SNE representations of systems with topological order

Figures 10 to 12 demonstrate the ability of PCA and t-SNE to distinguish topological phases. When only PCA, or t-SNE with initial PCA pre-processing is conducted in Figures 10 and 11, topological phases are unable to be detected. In the case of PCA, the linear method is unable to distinguish features past first order in the representation and the underlying non-local behaviour cannot be observed.

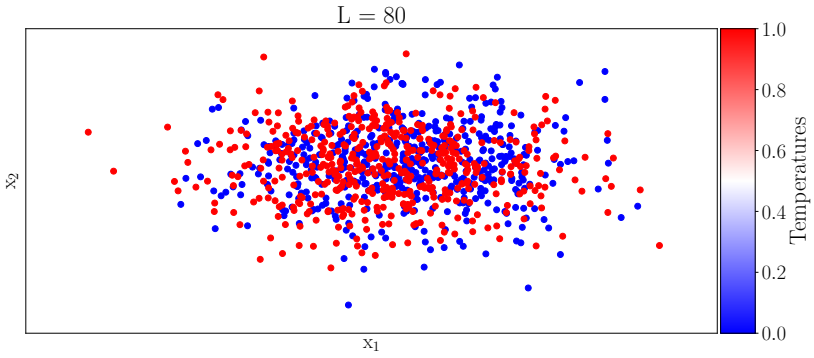

Figure 10: PCA representation of lattice gauge theory configurations, for $L = 80$.

Similarly in the case of t-SNE using the first 50 components of PCA, the algorithm cannot detect the non-local structure without significantly more information, potentially $O(L)$ components as per the discussion regarding PCA variance ratios, and the two phases are unable to be distinguished. Given the circular structure formed, the non-linear corrections to the representation appear to partially learn some feature during training compared to the PCA method.

When strictly t-SNE is used in Fig. 12, non-linearly transforming all of the information contained in the configurations is shown to be adequate to identify the

non-local structure of the model. However, it is important to distinguish that the temperature labels, although not explicitly used in the training process, are essential at identifying these phases which are not spatially separated, but are more tightly clustered at greater systems sizes. Optimizing the training process, and having configurations for a full range of temperatures approaching $T = 0$ could potentially allow further structure to be identified, and spatial separation may even be able to be achieved.

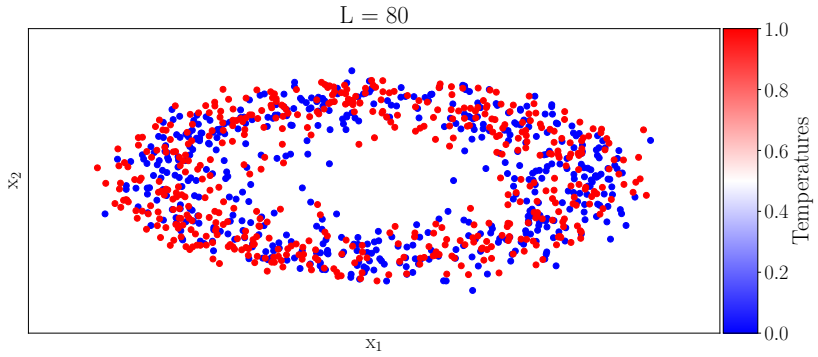

Figure 11: t-SNE representation of lattice gauge theory configurations, with pre-processing with PCA of 50 dimensions, for $L = 80$.

6 Conclusions

Magnetic spin models on square lattices with periodic boundary conditions, including an Ising model with a continuous phase transition, a q-state Pott's model with a first order phase transition for $q > 4$, and a $\mathbb{Z}_2$ lattice gauge theory with topological order were used to study the ability of PCA and t-SNE dimensional reduction techniques to distinguish phases around a critical point, and to detect the presence of local and non-local order.

PCA has shown to be quite effective at distinguishing degenerate ordered phases near continuous and first order phase transitions. The use of explained variance ratios of the principal components further quantifies both the redundancy in the information contained in sample configurations, and the ability of this method to convey the state of the system with few principal components. These variance ratios may be an indicator of strictly local order in a system if only one principal component is dominant in the representation.

t-SNE has shown to not only be able to distinguish degenerate ordered phases near continuous and first order phase transitions, but it can also identify phases near topological transitions. This non-linear method provides increased numbers of corrections to the lower dimensional representations, compared to finding the linear

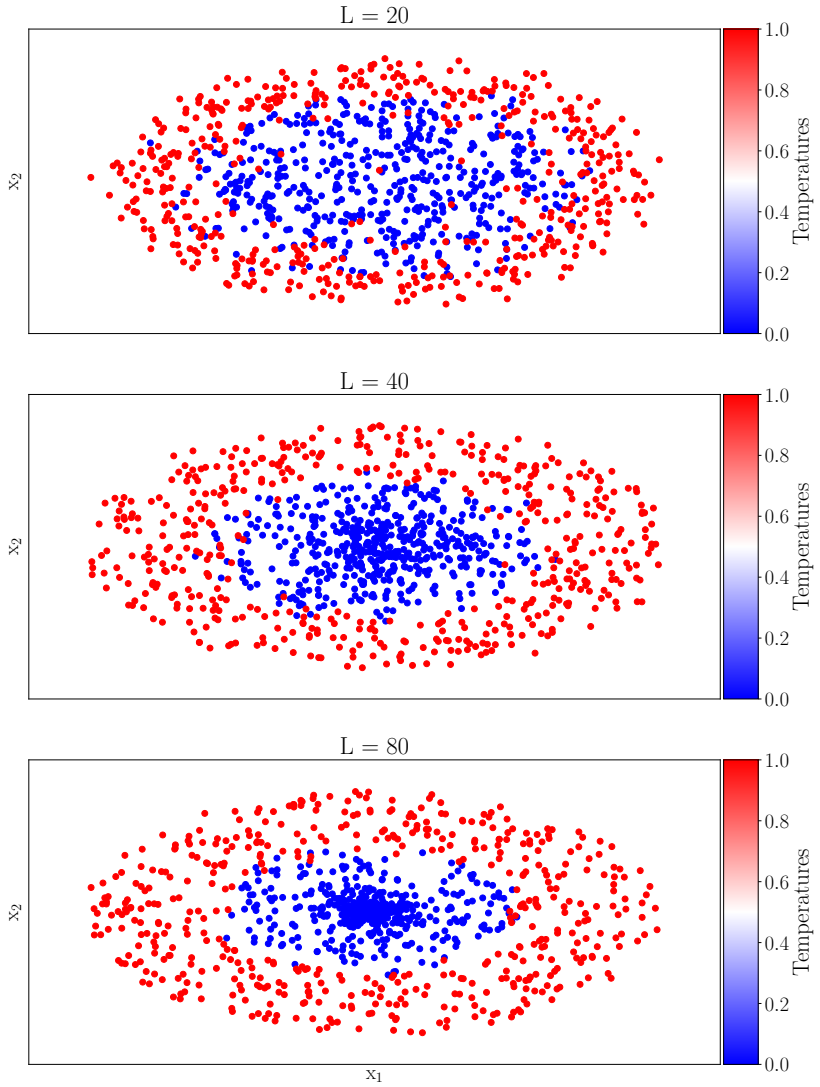

Figure 12: t-SNE representation of lattice gauge theory configurations, without pre-processing with PCA, for $L = 20, 40, 80$.

components, allowing more distinct structures and clustering to occur. Particularly when there is no initial linear processing, it uses as much of the information content in the initial configurations as is available during the training process. Its ability to detect non-local order in topological systems suggests it is highly applicable to a wide range of systems and should be applied to quantum systems with more complicated correlations and phases.

Both PCA and t-SNE have shown to be promising at analyzing phase transitions and each method is appropriate at distinguishing certain types of features of a model.

It has also been shown that both methods rely heavily on having enough information about each possible state of the physical parameter being learned: ergodic and uniform sampling of input configurations is likely essential to the success of the training. There is therefore an open question of the minimum information, and the variance of such information necessary to be input for these dimensional reduction methods to still learn a complete description of the phases of the system. Future emphasis should therefore be placed on determining the limits of which classes of systems can be represented. and answering this particular question.

More specific suggestions for future work regarding the computational aspects of the implementation of these methods include investigating, both theoretically and through simulation, the impact of the initial algorithm conditions and training procedures. In particular, the perplexity value that determines the higher dimensional representation variance in t-SNE appears to be an important parameter that determines how much information is supplied to the optimization procedure. It is also noted that the current t-SNE method defines higher and lower dimensional distributions in terms of the element-wise distance between configurations. It is suggested that this element-wise distance metric removes some of the translational invariance in the lattice systems, and instead a different distance metric, perhaps a convolution between all elements of all configurations, should be implemented to better reflect the symmetries present. The impact of these initial algorithm conditions on whether degenerate phases, or whether the presence of non-local order can always be represented by t-SNE would be beneficial in determining whether these methods can be generalized to more complex models.

Subsequent analysis is further required in order to determine whether actual physical parameters are being learned during the training process, or whether the cluster structures are formed from purely mathematical means. The linear method of PCA has been shown to learn the approximately linear order parameter in local systems [19], and it should be confirmed whether t-SNE is learning a higher order correlation function in these spin models, or some different parameter entirely.

Particularly when PCA is not initially used before t-SNE, the spiral data structure for the Ising model, and the distinction of the topological phases for the gauge theory model suggest a higher order, and non-linear physical quantity is being represented. These more complicated structures may indicate that more subtle physical parameters have been learned, however it has been shown that these structures are currently difficult to interpret in terms of the physical model.

Finally, these methods should be applied to investigating the coexistence of phases right at first order and topological phase transitions. As discussed, the effect of the double peaked energy histograms at first order phase transitions merging together in smaller system sizes, as well as hysteresis effects may radically affect what can be learned by PCA and t-SNE. Whether these methods can distinguish phases in configurations directly at a critical point would determine whether these methods would be appropriate for studying more complex phase transitions. The

calculated Monte Carlo observables show a clear first order transition for $q > 4$ and this model, in conjunction with obtained samples at $T_c = 0$ for the $\mathbb{Z}_2$ lattice gauge theory, would be appropriate models to expand on this work and further explore the applicability of t-SNE at detecting non-trivial structure.

Acknowledgements

The author would like to acknowledge and sincerely thank their supervisors Professor Roger Melko and Dr. Lauren Hayward Sierens , who's encouragement, passion and excitement about this field has inspired them to pursue computational physics well beyond this essay. It was such a pleasure to work with each member of Professor Melko's group, whose enthusiasm and insight fostered an exciting and motivating environment to learn everything about numerical methods. In particular, the useful discussions with Anna Golubev regarding machine learning methods, Stavros Efthymiou on calculating observables, and Juan Carrasquilla concerning his insight into the relationship between KL divergence and maximum likelihood, were all thoroughly enjoyed and appreciated.

Bibliography

[1] D. Landau and B. K., *A Guide to Monte Carlo Simulations in Statistical Physics*, 3 ed. (Cambridge University Press, United Kingdom, 2009).

[2] M. Newman and G. Barkema, *Monte Carlo Methods in Statistical Physics*, 1 ed. (Oxford University Press, New York, 1999).

[3] R. Swendsen, "Nonuniversal Critical Dynamics in Monte Carlo Simulations," Phys. Rev. Lett. **58**, 86 (1987).

[4] J. Carrasquilla and R. Melko, "Machine learning phases of matter," Nature Physics **13**, 431 (2017).

[5] L. Huang and L. Wang, "Accelerated Monte Carlo simulations with restricted Boltzmann machines," Phys. Rev. B **95**, 035105 (2017).

[6] J. Liu *et al.*, "Self-learning Monte Carlo method," Phys. Rev. B **95**, 041101 (2017).

[7] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," Journal of Machine Learning Research **1**, 1 (2008).

[8] K. Binder, "Static and Dynamic Critical Phenomena of the Two-Dimensional q-State Potts Model," Journal of Statistical Physics **24**, 69 (1981).

[9] J. Kogut, "An introduction to lattice gauge theory and spin systems," Reviews of Modern Physics **51**, 659 (1979).

[10] R. Baxter, S. Kelland, and F. Wu, "Potts model at the critical temperature," Journal of Physics C: Solid State Physics **6**, 445 (1973).

[11] F. Wu, "Potts model of magnetism," Journal of Applied Physics Low Temperature Physics **55**, 2421 (1984).

[12] F. Wu, "The Potts model," Reviews of Modern Physics **54**, 235 (1982).

[13] E. Domany *et al.*, "Classification of continuous order-disorder transitions in absorbed monolayers," Phys. Rev. B **18**, 2209 (1978).

[14] E. Anderson, "Monte Carlo Methods and Importance Sampling,", 1999.

[15] W. Janke, *Computational Physics* (Springer, Berlin, 1996), pp. 10–43.

[16] U. Wolff, "Collective Monte Carlo Updating for Spin Systems," Phys. Rev. Lett. **62**, 361 (1989).

[17] C. M. Fortuin and P. Kasteleyn, "On the Random Cluster Model," Physica **57**, 536 (1972).

[18] Y. Tomita and Y. Okabe, "Probability-Changing Cluster Algorithm: Study of Three-Dimensional Ising Model and Percolation Problem," Journal of the Physical Society of Japan **71**, 1570 (2002).

[19] L. Wang, "Discovering phase transitions with unsupervised learning," Phys. Rev. B **94**, 195105 (2016).

[20] Q. Liu and A. Ihler, "Distributed Estimation, Information Loss and Exponential Families,", 2014.