Machine Learning Analysis of Ising Worms

Sam Foreman

University of Iowa

December 7, 2017

Overview

Introduction

- Present evidence that principal component analysis can be used to identify relevant features in the 2D Ising model.
- Demonstrate that the eigenvalue corresponding to the first principal component captures the logarithmic divergence of the specific heat at the critical temperature.
- Look at the behavior under renormalization group transformations (blocking).

Introduction

- Use an alternative representation of the Ising model, using graphs consisting of closed loops ("worms") instead of spins.
- **Worm algorithm**:
  - Avoids critical slowing down problem while remaining a local Metropolis scheme.
  - Preferable over the normal configuration space (spins) because paths are easily represented by products of tensors.
  - Apply ideas from TRG to perform iterated blocking.

Motivation

- Interest in applying ideas from machine learning to physical problems has grown tremendously over past few years.
- Mehta & Schwab showed that there exists an exact mapping between the Variational Renormalization Group and Deep Learning
- Melko et al. have successfully used neural networks to identify phase transitions in a variety of systems.
- Creating self-learning Monte Carlo (MC) methods via supervised learning for more efficient simulations.
- Using Restricted Boltzmann Machines to learn parameters of effective models for efficiently generating gauge configurations in lattice QCD simulations.

High Temperature Expansion

- We begin with the $2D$-Ising Hamiltonian on an $L \times L = V$ lattice

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j, \text{ with } s_i = \pm 1,$$

- And the partition function

$$Z = \sum_{\{s\}} e^{-\beta \mathcal{H}} = \sum_{\{s\}} \prod_{\langle i,j \rangle} e^{\beta J s_i s_j},$$

High Temperature Expansion

- Using the identity

$$e^{\beta s_i s_j} = \cosh\beta\,(1 + s_i s_j \tanh\beta)$$

we can rewrite the partition function as a product over bonds connecting nearest neighbor sites

$$\begin{aligned} Z &= \sum_{\{s\}} \prod_{\langle ij \rangle} \cosh\beta(1 + s_i s_j \tanh\beta) \\ &= (\cosh\beta)^{2V} \sum_{\{s\}} \prod_{\langle ij \rangle} (1 + s_i s_j \tanh\beta) \\ &= (\cosh(\beta))^{2V} \sum_{\{s\}} (1 + s_i s_j \tanh(\beta))(1 + s_k s_l \tanh(\beta)) \cdots \end{aligned}$$

High Temperature Expansion

- We can rewrite the sum $\sum_{\{s\}}$ using

$$\sum_{s_i=\pm 1} s_i = 0 \Longrightarrow \sum_{s_i=\pm 1} s_i^n = \begin{cases} 2 & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases}$$

- From which it follows that

$$\sum_{s_1=\pm 1,\, s_2=\pm 1,\ldots} s_1^{n_1} s_2^{n_2} \cdots = \begin{cases} 2^V & \text{if all } n_i \text{ are even} \\ 0 & \text{otherwise} \end{cases}$$

- **Only closed paths!**

High Temperature Expansion

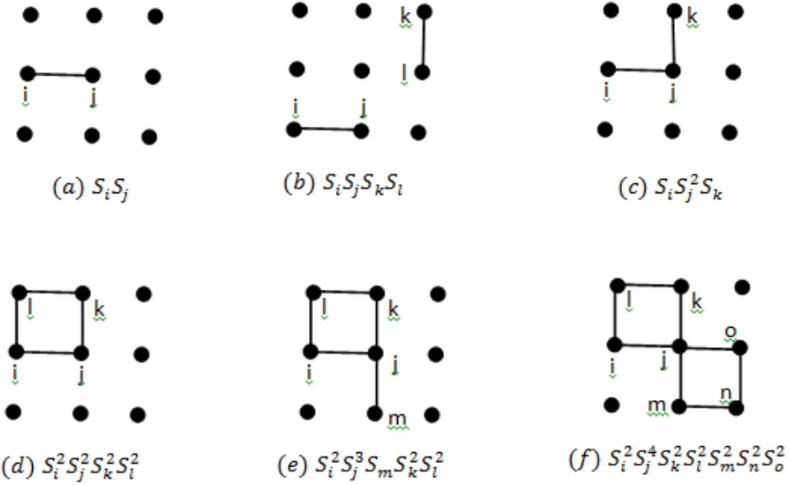


Figure: Image borrowed from http://nptel.ac.in/courses/115103028/31

High Temperature Expansion

- Finally,

$$Z = 2^V \left(\cosh\beta\right)^{2V} \sum_{N_b} n(N_b) \left(\tanh\beta\right)^{N_B}$$

- Where $n(N_b)$ is a combinatoric factor representing the number of distinct closed-loop configurations consisting of N_b active bonds.

High Temperature Expansion

- By definition,

$$\begin{aligned}\langle E\rangle &= -\frac{\partial}{\partial\beta}\ln Z\\ &= -\tanh(\beta)\left(2V + \frac{\langle N_b\rangle}{\sinh^2(\beta)}\right)\end{aligned}$$

- Calculate the **specific heat** using

$$\begin{aligned}\mathcal{C}_V &= \frac{1}{V}\frac{\partial\langle E\rangle}{\partial T}\\ &= \beta^2\left[\frac{1}{\cosh^2(\beta)}\frac{2V}{V} - \frac{4\cosh(2\beta)}{\sinh(2\beta)}\frac{\langle N_b\rangle}{V}\right.\\ &\qquad\left. + \left(\frac{2}{\sinh(2\beta)}\right)^2\frac{\langle(N_b-\langle N_b\rangle)^2\rangle}{V}\right]\end{aligned}$$

High Temperature Expansion

- The singularity near T_c has been calculated previously by Onsager:

$$\mathcal{C}_V = -\frac{2}{\pi}\left(\ln(1+\sqrt{2})\right)^2 \ln\left(|T - T_c|\right) + \text{regular}$$

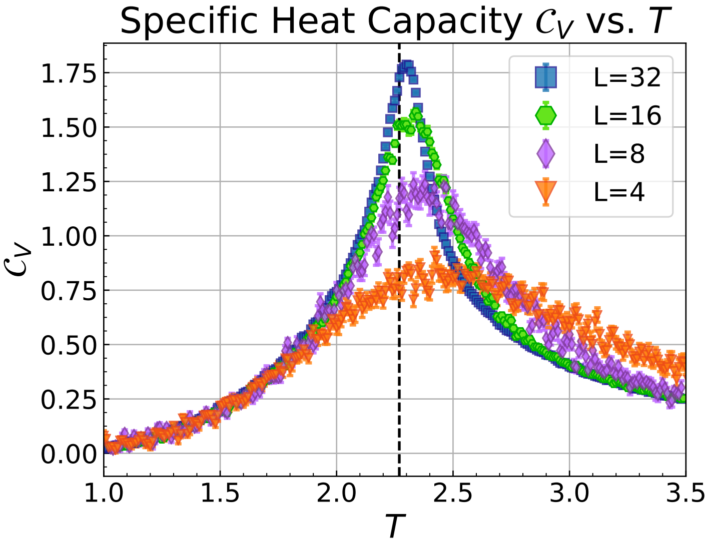
- The only possibly divergent part of our expression is the variance of N_b per unit volume:

$$\Delta^2_{N_b} \equiv \frac{\langle (N_b - \langle N_b \rangle)^2 \rangle}{V}$$

- Comparing to our result for $\mathcal{C}_V$, we have

$$\Delta^2_{N_b} = -\frac{2}{\pi}\ln(|T - T_c|)$$

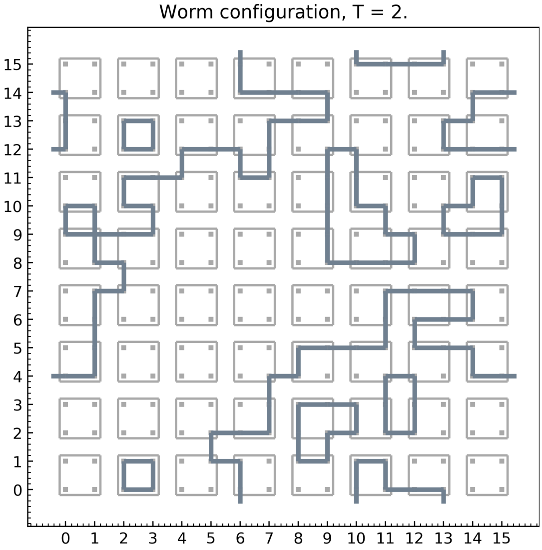
Specific Heat Results



Worm Algorithm

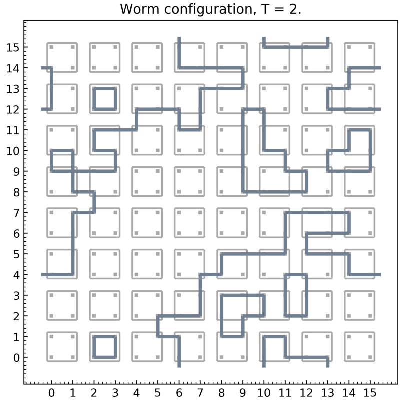
1. **Initialize:** Randomly select a starting point on the lattice i.
2. **Update:** Propose a move to a neighboring site i', selected at random.
   - If no link exists between i, i', create link with probability $P = \min\{1, \tanh\beta\}$
   - If link already exists, remove it with probability $P = 1$
   - If $i' = i$ (i.e. closed path), go to (1.) Otherwise, go to (2.)

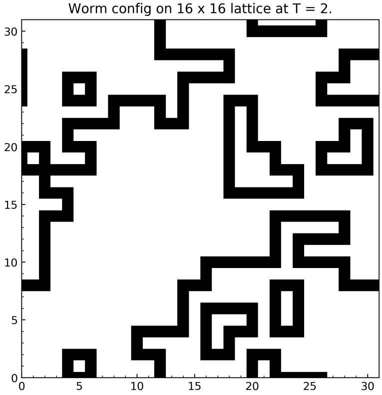
Sample Worm Configuration



Worms as Images

- We can represent a given worm configuration as a grayscale image of $2L \times 2L$ pixels.





Worms as Images

- At a particular temperature T, we have a collection of $N_{configs}$ worm configurations stored as images.
- We can flatten each of these images into a vector

$$\mathbf{v}^{(n)} = \left(v_0^{(n)}, v_1^{(n)}, \ldots, v_{4V}^{(n)}\right) \quad \text{with}$$

$$v_i^{(n)} \in \{0, 1\}, \quad \text{and} \quad n = 1, \ldots, N_{configs}$$

Worms as Images

- We can calculate the average number of occupied bonds by averaging over all configurations

$$
\begin{aligned}
\langle N_b \rangle &= \frac{1}{N_{configs}} \sum_{n=1}^{N_{configs}} N_b^{(n)} \\
&= \frac{1}{N_{configs}} \sum_{n=1}^{N_{configs}} \left(\sum_{j=bonds} v_j^{(n)} \right) \\
&= 2V \langle \mathbf{v}_b \rangle \implies \\
\langle \mathbf{v}_b \rangle &= \frac{\langle N_b \rangle}{2V}
\end{aligned}
$$

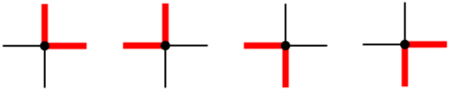
- Where $\langle \mathbf{v}_b \rangle$ represents the average greyscale value of pixels corresponding to bonds on the lattice.

Worms as Images

- By translation invariance, in the limit $N_{configs} \to \infty$, we have

$$\langle v_j \rangle = \begin{cases} \langle \mathbf{v}_b \rangle = \langle N_b \rangle / 2V & \text{if } j \text{ is a } \textbf{bond} \text{ pixel} \\ \langle \mathbf{v}_s \rangle = \left(\langle N_{sites,1} \rangle + \langle N_{sites,2} \rangle \right) / V & \text{if } j \text{ is a } \textbf{site} \text{ pixel} \\ 0 & \text{if } j \text{ is a plaquette} \end{cases}$$

- $N_{sites,1}$ is the number of sites visited once (connecting exactly two links)



- $N_{sites,2}$ is the number of sites visited twice (connecting exactly four links)

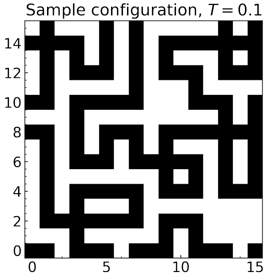
Worms as Images

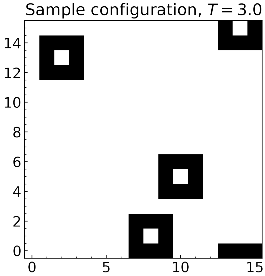
- At high T, we can make the approximation $\langle N_{sites,2} \rangle \simeq 0$, from which it follows that

$$\left\langle \sum_{j=bonds} v_j \right\rangle \simeq \left\langle \sum_{j=sites} v_j \right\rangle$$

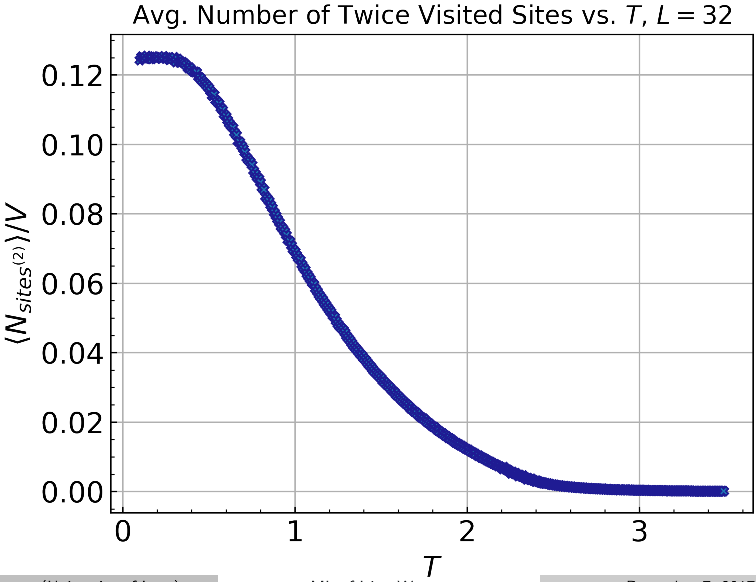
$$\Longrightarrow \langle \mathbf{v}_b \rangle \simeq \frac{1}{2} \langle \mathbf{v}_s \rangle$$

- In the average image, site pixels should be roughly twice as dark as bond pixels.

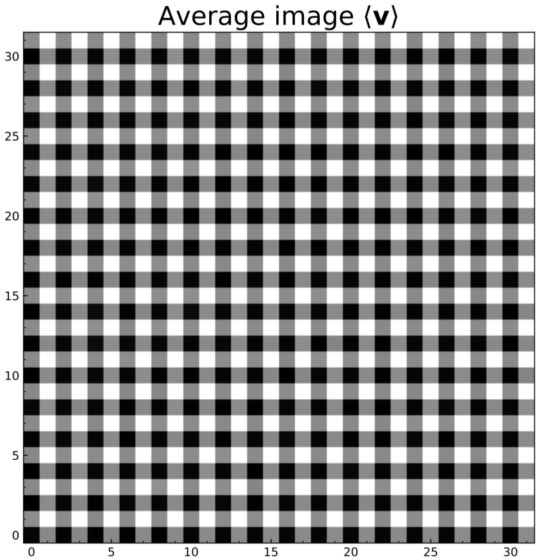




Worms as Images



Average Image



Principal Component Analysis (PCA)

- **Covariance matrix:**

$$C_{ij} = \left\langle \left(v_i - \langle \mathbf{v} \rangle_i\right)\left(v_j - \langle \mathbf{v} \rangle_j\right)^T \right\rangle$$
$$= \frac{1}{N_{configs}} \sum_{n=1}^{N_{configs}} \left(v_i^{(n)} - \langle \mathbf{v} \rangle_i\right)\left(v_j^{(n)} - \langle \mathbf{v} \rangle_j\right)^T,$$

- $v_i^{(n)}$ is the grayscale value of the i^{th} pixel in the n^{th} sample configuration
- $\langle \mathbf{v} \rangle_i$ is the grayscale value of the i^{th} pixel, averaged over the set of configurations.

PCA

- Eigenvalue decomposition of covariance matrix

$$C\mathbf{w}_i = \lambda_i \mathbf{w}_i, \quad \text{with} \quad \lambda_1 > \lambda_2 > \cdots$$

- $\mathbf{w}_i$ is called the i^{th} **principal component**.
- Intuitively, $\mathbf{w}_1$ points in the direction along which the data described by C varies most significantly.
- Our focus is on the first principal component, $\mathbf{w}_1$, along with its corresponding eigenvalue $\lambda_1 \equiv \lambda^{max}$.

PCA

Conjecture

We propose that the first principal component, $\mathbf{w}_1$ of the covariance matrix C is directly proportional to the average worm configuration (image) $\langle \mathbf{v} \rangle$, i.e.

$$\mathbf{w}_1 \propto \langle \mathbf{v} \rangle$$

PCA

- From this, we can extract a relationship between the eigenvalue corresponding to the first principal component, λ^{max} and the fluctuations Δ_{N_b} and Δ_{N_s}, (using $\langle \mathbf{v} \rangle^2 = 2V\langle \mathbf{v}_b \rangle^2 + V\langle \mathbf{v}_s \rangle^2$)

$$\begin{aligned}\lambda^{max} &= \mathbf{w}_1^T C \mathbf{w}_1 \\ &= \frac{1}{N_{configs}} \sum_{n=1}^{N_{configs}} \left[\langle \mathbf{v}_b \rangle^2 \left(\Delta_{N_b}^{(n)}\right)^2 + 2\langle \mathbf{v}_b \rangle \langle \mathbf{v}_s \rangle \Delta_{N_b}^{(n)} \Delta_{N_s}^{(n)} \right. \\ &\qquad \left. + \langle \mathbf{v}_s \rangle^2 \left(\Delta_{N_s}^{(n)}\right)^2 \right] \frac{1}{\left(2\langle \mathbf{v}_b \rangle^2 + \langle \mathbf{v}_s \rangle^2\right)}\end{aligned}$$

PCA

- In the high Temperature approximation where we neglect the contributions from twice-visited sites,

$$\langle \mathbf{v}_s \rangle \simeq 2\langle \mathbf{v}_b \rangle, \;\; \langle N_s \rangle \simeq \langle N_b \rangle, \quad \text{and } \Delta_{N_b} \simeq \Delta_{N_s}$$

we obtain the result

$$\begin{aligned} \lambda^{max} &\simeq \frac{3}{2} \frac{1}{N_{configs}} \sum_{n=1}^{N_{configs}} \left(\Delta_{N_b}^{(n)}\right)^2 \\ &= \frac{3}{2} \left\langle \Delta_{N_b}^2 \right\rangle \end{aligned}$$

PCA

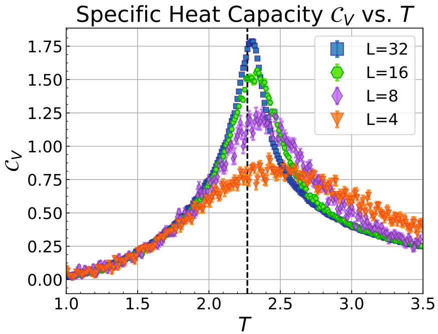
- From our earlier expression for $\mathcal{C}_V$, **near the critical temperature** $T_c = 1/\beta_c$ we can write

$$\mathcal{C}_V \simeq \beta_C^2 \left(\frac{2}{\sinh(2\beta_C)}\right)^2 \langle \Delta_{N_b}^2 \rangle \tag{1}$$

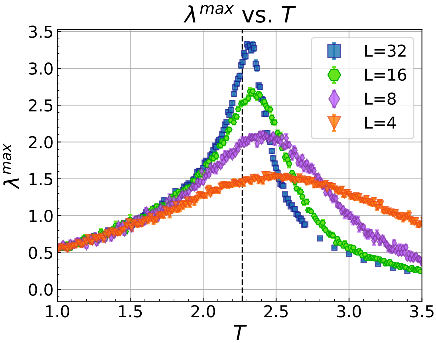
$$= \left(\ln\left(1+\sqrt{2}\right)\right)^2 \langle \Delta_{N_b}^2 \rangle \tag{2}$$

$$= \underbrace{\left(\ln\left(1+\sqrt{2}\right)\right)^2 \frac{2}{3}}_{\approx 0.52} \lambda^{max} \tag{3}$$

Conjecture



Specific Heat Capacity $\mathcal{C}_V$ vs. T
$\mathcal{C}_V$
T
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
1.0
1.5
2.0
2.5
3.0
3.5
L=32
L=16
L=8
L=4



λ^{max} vs. T
λ^{max}
T
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1.0
1.5
2.0
2.5
3.0
3.5
L=32
L=16
L=8
L=4

Results

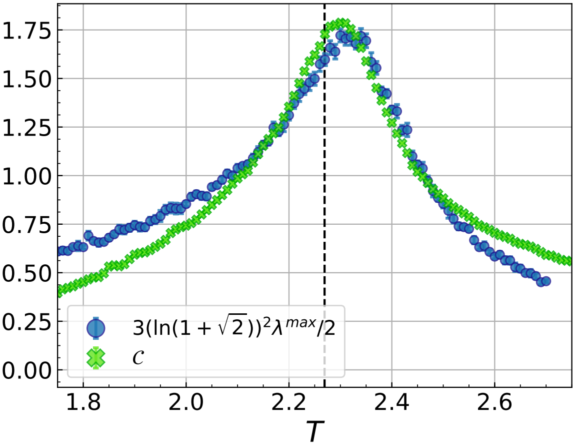
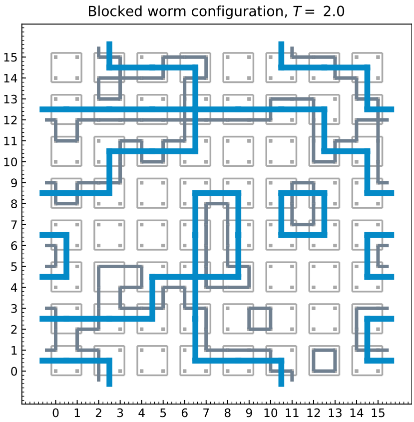


Figure: Relationship between $\mathcal{C}_V$ and λ^{max} illustrating (3) near criticality suggests an excellent agreement with our conjecture.

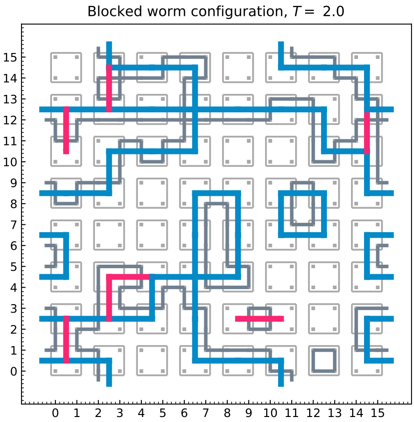
Blocking

- We implement a 'coarse-graining' renormalization approach, where the lattice is divided into blocks of 2×2 squares.
- Each 2×2 square is then 'blocked' into a single site, where the new external bonds in a given direction are determined by the number of active bonds exiting a given square.
- If a given block has exactly one external bond in a given direction, the blocked site retains this bond in the blocked configuration, otherwise it is ignored.
- However, if a given block has exactly two external bonds in a given direction, we are free to either include or neglect these bonds when constructing the blocked configuration.

Blocking

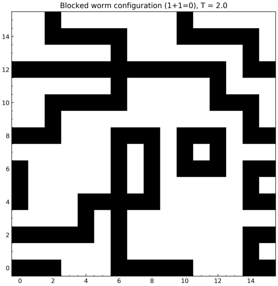


(a) $1 + 1 = 0$

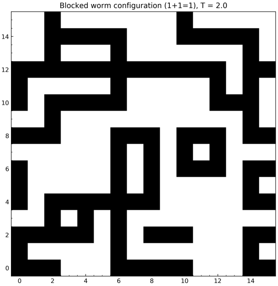


(b) $1 + 1 = 1, 2$

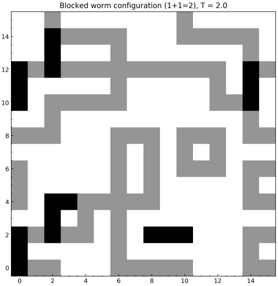
Blocking



(c) $1 + 1 = 0$



(d) $1 + 1 = 1$



(e) $1 + 1 = 2$

Blocking results

- We can plot Δ_{N_b} versus T for each of the different blocking procedures, as shown in Fig. 3.

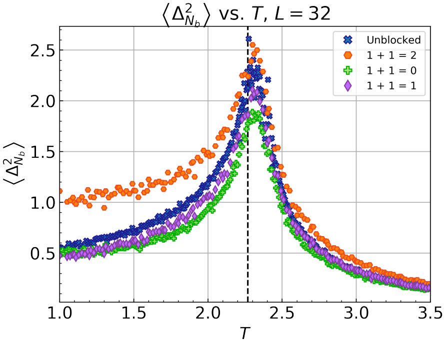
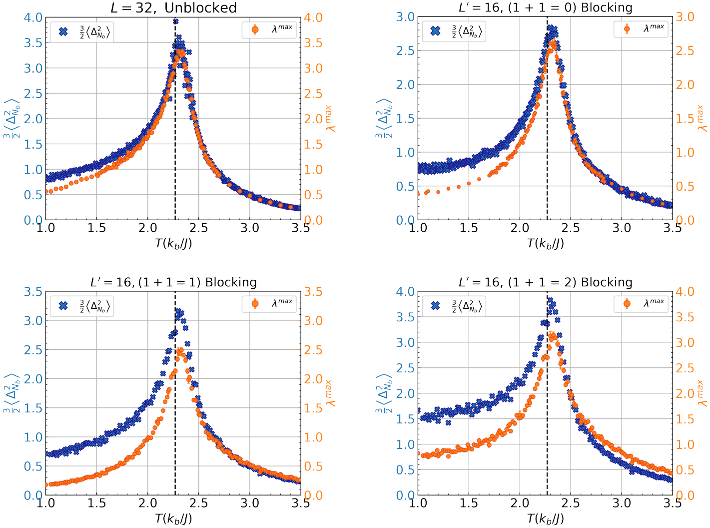


Figure: $\langle \Delta^2_{N_b} \rangle$ vs. T under each of the different blocking schemes.

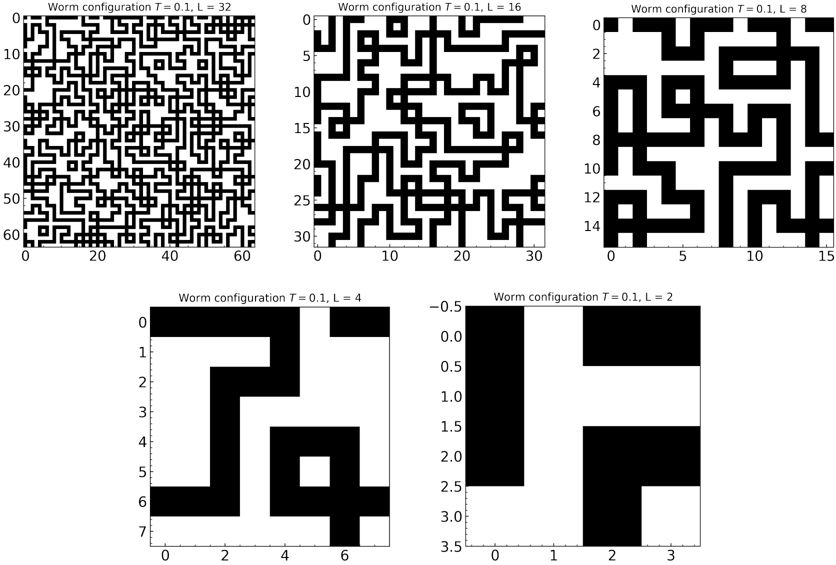
Blocking results



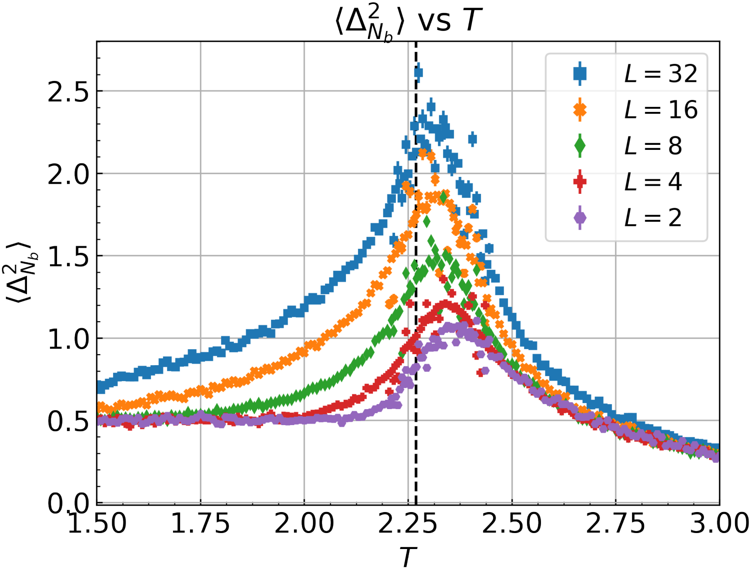
Iterated Blocking

- Using the $1 + 1 = 0$ blocking scheme, we maintain the 'closed-path' restriction under a single blocking step, so we can successively block a given configuration.
- Do these iterated blocking configurations retain information about the logarithmic divergence of $\mathcal{C}_V$?

Iterated Blocking



Iterated Blocking (Results)



Conclusion

- We successfully identified the logarithmic divergence of the specific heat $\mathcal{C}$ at the critical temperature.
- We were able to relate this divergent behavior to the fluctuations $\Delta^2_{N_b}$ in the average number of bonds $\langle N_b \rangle$ in configurations stored as images.
- From this analysis, we observe that PCA can also be used to successfully extract information about this divergence without any information about the underlying physical structure.
- The simplicity of this approach, applied to the 2D Ising model suggests that there may be additional unsupervised learning techniques that also be successful in discovering physical information about a wide-variety of physical systems.

Thanks for listening!

More information available at https://arxiv.org/abs/1710.02079