Machine learning inspired analysis of the Ising model transition[†]

Samuel Foreman[‡]

Department of Physics and Astronomy, The University of Iowa,
Computational Sciences Division, Argonne National Laboratory
E-mail: samuel-foreman@uiowa.edu

Joel Giedt

Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute
E-mail: giedtj@rpi.edu

Yannick Meurice

Department of Physics and Astronomy, The University of Iowa
E-mail: yannick-meurice@uiowa.edu

Judah Unmuth-Yockey

Department of Physics, Syracuse University
E-mail: jfunmuth@syr.edu

We illustrate how principal component analysis of simulation data represented as images generated from the worm algorithm, a method to sample the strong coupling contributions, can be used to identify the critical temperature T_c in the Ising model. It is shown that the eigenvalue corresponding to the first principal component of the covariance matrix obtained from pixel ensembles scales logarithmically as one approaches T_c, in a way that is similar to the specific heat. We then illustrate how to block the resulting worm configurations under renormalization group transformations. It is found that curves for the variance of the average number of bonds can be scaled appropriately to illustrate universal behavior under the renormalization group transformation, and that discrepancies can be understood as an effect of approximations.

The 36th Annual International Symposium on Lattice Field Theory - LATTICE2018
22-28 July, 2018
Michigan State University, East Lansing, Michigan, USA.

1. Introduction

Using the example of configurations generated with the worm algorithm for the two-dimensional Ising model, we propose renormalization group (RG) transformations, inspired by the tensor RG (TRG) [2, 3, 4], that can be applied to arbitrary sets of images. We discuss how quantities of interest behave under successive applications of these blocking transformations and suggest ways to obtain data collapse and compare results with the two-state TRG approximation near the fixed point. We begin with the well-known Ising model in two dimensions with hamiltonian and partition function

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j, \quad Z = \sum_{\{\sigma_i\}} e^{-\beta H/T} \tag{1.1}$$

with $\beta = 1/T$, and $\sigma_i = \pm 1$. For simplicity, we choose units in which $J/k_B = 1$. The partition function can be rewritten using the character expansion [5]:

$$\exp(\beta \sigma) = \cosh(\beta) + \sigma \sinh(\beta) = \cosh(\beta)(1 + \sigma \tanh(\beta)). \tag{1.2}$$

This allows us to reformulate the model in terms of integer variables that live on the links of the lattice. In doing so we are left with the restriction that the sum of the link variables associated with each site must be even, i.e. only those configurations consisting of closed loops will contribute to the partition function. In terms of these link variables, for a lattice of volume $V = L \times L$ the partition function is now

$$Z = 2^V \cosh(\beta) \sum_{N_b} (\tanh(\beta))^{N_b} \mathcal{N}(N_b) \tag{1.3}$$

where N_b is the total number of occupied links and $\mathcal{N}(N_b)$ denotes the number of legal graphs with N_b links. We can use the fact that $\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta}$ to relate $\langle N_b \rangle \propto \langle E \rangle$, and obtain an expression for the fluctuation in the average number of bonds

$$\langle \Delta_{N_b}^2 \rangle \equiv \langle (N_b - \langle N_b \rangle)^2 \rangle, \tag{1.4}$$

which is analogous to the specific heat per site in the original theory. From the logarithmic singularity of the specific heat, we find that

$$\frac{\langle \Delta_{N_b}^2 \rangle}{V} = \frac{-2}{\pi} \ln(|T - T_c|) + \cdots. \tag{1.5}$$

Contributions to Z are sampled using the Worm Algorithm [6], which can be represented as two-dimensional grayscale images of size $2L \times 2L$, each pixel being equal to one half of the lattice spacing. Each site, link, and plaquette of the lattice gets mapped to an individual pixel, which will take on a nonzero value if and only if it corresponds to an occupied link in the lattice. An example of this representation can be seen in Fig 1.

2. Principal component analysis and criticality

Principal component analysis (PCA) is a standard method to analyze sets of images. It is often used to reduce the dimensionality of a dataset to the most important "directions". For our example, at a given value of β, our data of interest consists of N_{configs} images $\{\vec{v}\}_{n=1}^{N_{\text{configs}}}$, each of
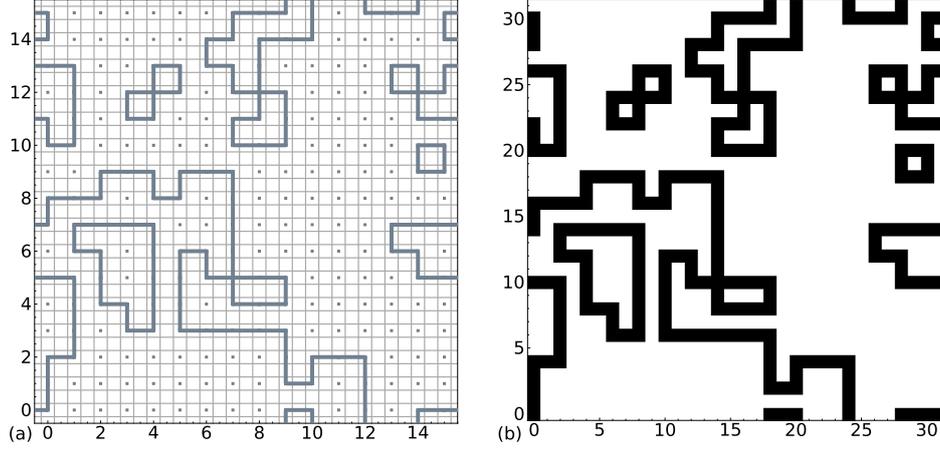
Figure 1: (a) Legal worm configuration on an $L \times L$ lattice with periodic boundary conditions and; (b) its equivalent representation as a $2L \times 2L$ black and white pixel image. *Figure reproduced with permission from [1], @2018 APS.*

which is a single row vector $\vec{v}^n \in \mathbb{R}^{4V}$, with each $v_i \in \{0, 1\}$. We can construct the covariance matrix describing this data set

$$S_{ij} = \frac{1}{N_{\text{configs}}} \sum_{n=1}^{N_{\text{configs}}} \left(v_i^n - \langle v_i \rangle\right) \left(v_j^n - \langle v_j \rangle\right), \tag{2.1}$$

where indices $i, j = 1, \ldots, 4V$ label the pixels of the nth image $\vec{v}^n$, and $\langle v_i \rangle$ denotes the average value of the jth pixel across all images. PCA then extracts the solutions to $S u_\alpha = \lambda_\alpha u_\alpha$ and orders them in decreasing magnitude of λ_α. Under mild assumptions and an easily justified approximation [1], we found a relation between the largest PCA eigenvalue $\lambda_{\max}$ and the logarithmic divergence of the specific heat, namely

$$\lambda_{\max} \approx \frac{3}{2} \frac{\langle \Delta_{N_b}^2 \rangle}{V} \approx -\frac{3}{\pi} \ln(|T - T_c|) \tag{2.2}$$

An independent confirmation of the approximate validity of this relationship can be seen in Fig 2.

3. TRG coarse-graining

We can use the constraints imposed by the character expansion to build a tensor

$$T_{xx'yy'}^{(i)}(\beta) = [\tanh(\beta)]^{(n_x + n_{x'} + n_y + n_{y'})/2} \times \delta_{n_x + n_{x'} + n_y + n_{y'}, 0 \bmod 2}. \tag{3.1}$$

located at the ith site of the lattice, with integer variables $n_{\hat{\mu}} \in \{0, 1\}$ on each of the links x, x', y, y' and the $\delta_{i,j}$ is satisfied if the sum of the $n_{\hat{\mu}}$ are even. By contracting these tensors together in the pattern of the lattice, we recreate the closed loop paths sampled by the worm algorithm. In terms of these tensors the partition function can be written as shown in Eq. 3.2. The most important aspect of this reformulation is that it can be coarse-grained efficiently. Under this transformation, the partition function becomes

$$Z = 2^V (\cosh(\beta))^{2V} \text{Tr} \prod_i T_{xx'yy'}^{(i)} \longrightarrow 2^V (\cosh(\beta))^{2V} \text{Tr} \prod_{2i} T_{XX'YY'}^{'(2i)}, \tag{3.2}$$
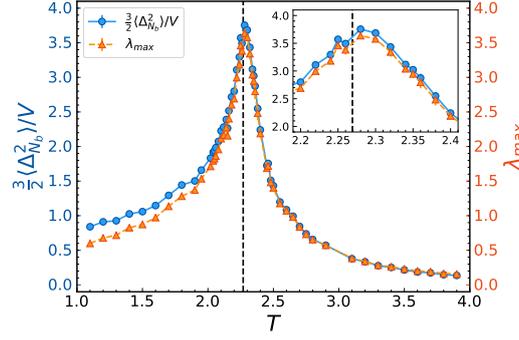
Figure 2: λ_{max} and $\frac{3}{2}\left\langle \Delta_{N_b}^2 \right\rangle /V$ (per unit volume) vs. T, illustrating the relation between the eigenvalue corresponding to the first principal component and the logarithmic divergence of the specific heat. The inset shows a qualitative agreement near the critical temperature. *Figure reproduced with permission from [1], @2018 APS.*

where $2i$ denotes the sites of the coarser lattice with twice the lattice spacing of the original lattice. This can't be repeated indefinitely, and we are forced to truncate by projecting the product states onto a smaller number of states that optimizes the closeness to the exact answer [3]. Using this blocking procedure, the scale factor is $b = 2$, with eigenvalue in the relevant direction $\lambda = b^{1/\nu} = 2$ since $\nu = 1$. Starting with an initial lattice of size $L = 64$, after ℓ iterations we are left with an effective size of the coarse-grained lattice $L_{\text{eff}} = L/b^{\ell}$.
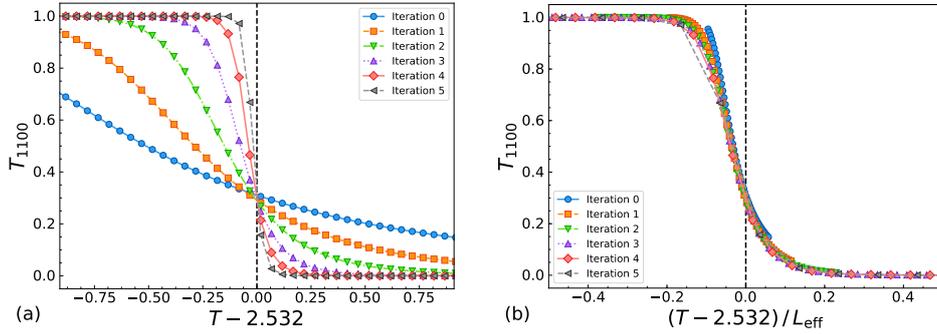

Figure 3: (a) T_{1100} vs. $T - T_c^{(2s)}$ under successive iterations of the blocking transformation, ; (b) T_{1100} vs. $(T - T_c^{(2s)})/L_{\text{eff}}$ illustrating the data collapse, where $T_c^{(2s)}$ is the critical temperature of the two state projection. *Figure reproduced with permission from [1], @2018 APS.*

We can compute $\langle N_b \rangle$ using this tensor method and compare it to our result from the worm algorithm. Using $N_b = \sum_{\ell} n_{\ell}$ for the sum of bond numbers at every link, we have for $\langle N_b \rangle$

$$\langle N_b \rangle = \frac{1}{Z} \sum_{\{n\}} \left(\sum_l n_l \right) \left(\prod_l \tanh^{n_l}(\beta) \right) \times \left(\prod_i \delta_{n_x+n_{x'}+n_y+n_{y'},0 \bmod 2}^{(i)} \right). \tag{3.3}$$

This can be viewed as $\langle N_b \rangle = \sum_{\ell} \langle n_{\ell} \rangle$. By translation and 90° rotational invariance, all $\langle n_{\ell} \rangle$ are equal. Consequently, it is enough to calculate $\langle n_{\ell} \rangle$ for one particular link (call it $\langle n \rangle$) and multiply it by $2V$: $\langle N_b \rangle = 2V\langle n \rangle$. The procedure for calculating $\langle n \rangle$ is detailed in [1]. The results of this calculation are shown in Fig 4.
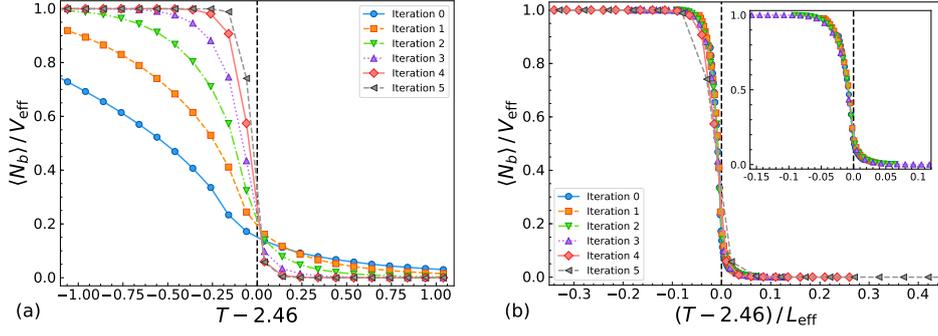
3

Figure 4: (a) $\langle N_b \rangle$ vs $(T - 2.46)$ under successive blocking steps calculated using 2-state HOTRG. (b) $\langle N_b \rangle$ vs $(T - 2.46)/L_{eff}$ Note that the value of 2.46 was determined qualitatively by choosing the value which gave the best resulting data collapse. *Figure reproduced with permission from [1], @2018 APS.*

4. Image coarse-graining

As in the TRG coarse-graining procedure, the image is first divided up into blocks of 2×2 squares, each of which is then replaced by a single site with new link variables determined by the sum of the external link variables in a given direction. Explicitly, if a given block has exactly one external link in a given direction, the blocked site retains this link in the blocked configuration, otherwise it is neglected. We denote this particular blocking procedure $1 + 1 \rightarrow 0$. An example of this blocking procedure carried out on a sample configuration can be found in [1]. This choice of blocking procedure is particularly useful since it maintains the closed loop restriction and can be iterated efficiently.

The first observation of this iterated blocking procedure is that it preserves the location of the peak of the fluctuations $\langle \Delta_{N_b}^2 \rangle$, and can be stabilized for the first few iterations by dividing by $V_{\text{eff}} \ln(L_{\text{eff}})$, as seen in Fig 5. Next, we consider $\langle N_b \rangle$ for successive iterations. The results are
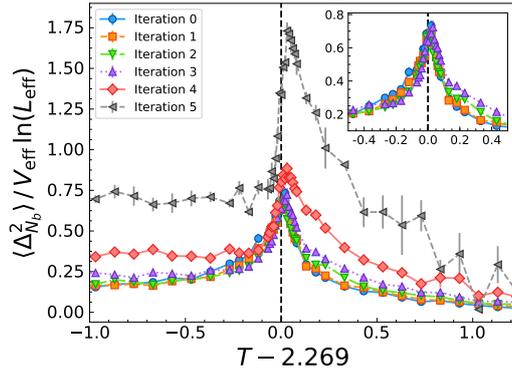

Figure 5: $\langle \Delta_{N_b}^2 \rangle$ vs. T under iterated blocking steps. The results are scaled by $1/V_{eff} \log(L_{eff})$ in order to demonstrate the data collapse near the critical temperature. This collapse is especially apparent in the inset, which shows the results under the first three blocking steps, with $L_{\text{eff}} = 64, 32, 16$, and 8. *Figure reproduced with permission from [1], @2018 APS.*

shown in Fig 6. The behavior of the results for both high and low temperatures can be understood by considering the different types of plaquettes that are possible [1]. An approximate data collapse
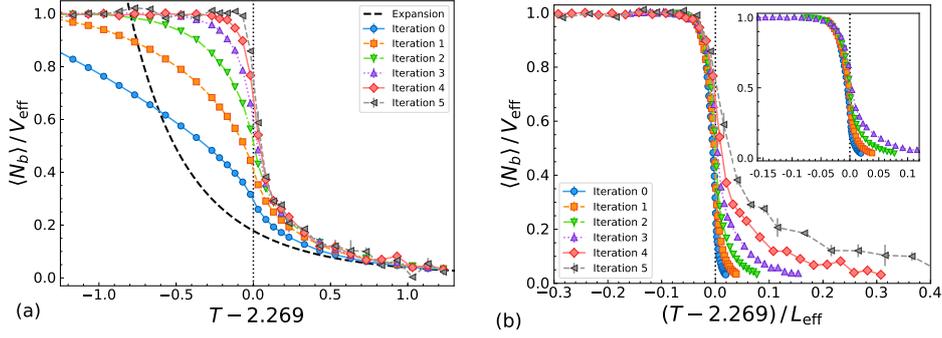
Figure 6: (a) $\langle N_b \rangle$ vs. T under iterated blocking steps The dashed black line illustrates the high temperature expansion, showing that the dominant configurations are those consisting of small, isolated plaquettes. (b) $\langle N_b \rangle$ vs. $(T - 2.269)/L_{eff}$. *Figure reproduced with permission from [1], @2018 APS.*

can be obtained by rescaling the temperature axis with respect to the critical value as in Fig 3. After this rescaling by a factor 2 at each iteration we observe a reasonable collapse on the low-temperature side, but since the unrescaled curves merge on the high temperature side, the rescaling then splits them and there is no collapse on that side.

5. Possible applications: CIFAR-10

For a possible application of these ideas, we consider the CIFAR-10 image set [7] consisting of $60,000$ 32×32 color images in 10 classes. First, each of these images were converted to a grayscale with pixel values in the range $[0, 1]$. Next, a grayscale cutoff value was chosen so that all pixels with values below the cutoff would become black, and pixels above the cutoff would become white, resulting in images consisting entirely of black and white pixels. Finally, each of these images were converted to "worm-like" images by drawing the boundaries separating black and white collections of pixels. An example of these preprocessing steps are illustrated in Fig. 7. This process was carried out on a mini-batch consisting of 500 randomly selected images from the CIFAR-10 image set. For each image in our mini-batch, we calculated $\langle N_b \rangle$ and $\langle \Delta_{N_b}^2 \rangle$ over a range of grayscale cutoff values in $[0, 1]$ in steps of 0.02. Each of these images were then iteratively blocked using the $(1 + 1 \rightarrow 0)$ blocking procedure described in Sec. 4, calculating $\langle N_b \rangle$ and $\langle \Delta_{N_b}^2 \rangle$ for each successive blocking step, as shown in Fig. 8. Immediately we see that there is no identifiable low temperature phase, and that for cutoff values near both 0 and 1, we obtain images which are mostly empty, similar to the high temperature configurations obtained from the worm algorithm. This suggests that there is no such notion of criticality (as characterized by the abrupt transition from a low to high temperature phase) like we found for the two-dimensional Ising model.

References

[1] S. Foreman, J. Giedt, Y. Meurice and J. Unmuth-Yockey, *Examples of renormalization group transformations for image sets*, *ArXiv e-prints, Phys. Rev. E (in press)* (2018) [1807.10250].

[2] Y. Liu, Y. Meurice, M. P. Qin, J. Unmuth-Yockey, T. Xiang, Z. Y. Xie et al., *Exact blocking formulas for spin and gauge models*, *Phys. Rev.* **D88** (2013) 056005 [1307.6543].
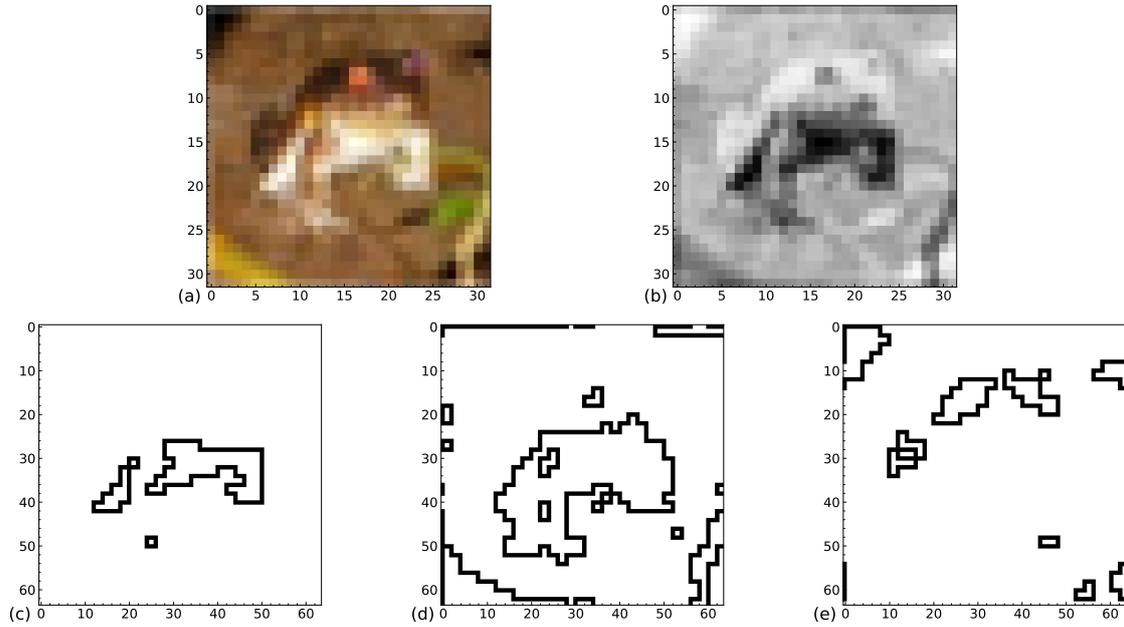
Figure 7: Example of preprocessing steps for converting CIFAR-10 images to 'worm-like' images, illustrating the resulting image for different values of the grayscale cutoff. (a) Original image from CIFAR-10 dataset. (b) Image converted to grayscale. (c) Resulting image from cutoff values of 0.25, (d) 0.5, and (e) 0.75. *Figure reproduced with permission from [1], @2018 APS.*
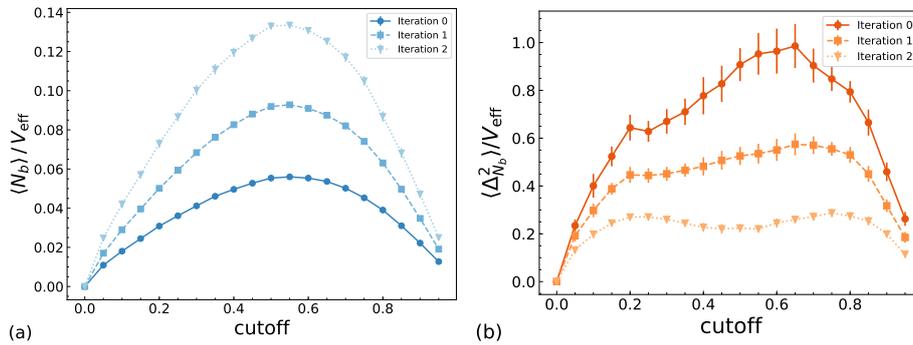

Figure 8: $\langle N_b \rangle$ and $\langle \Delta_{N_b}^2 \rangle$ vs. grayscale cutoff value for 500 randomly chosen images from the CIFAR-10 dataset. *Figure reproduced with permission from [1], @2018 APS.*

[3] Y. Meurice, *Accurate exponents from approximate tensor renormalizations*, *Phys. Rev.* **B87** (2013) 064422 [1211.3675].

[4] Z. Y. Xie, J. Chen, M. P. Qin, J. W. Zhu, L. P. Yang and T. Xiang, *Coarse-graining renormalization by higher-order singular value decomposition*, *Phys. Rev. B* **86** (2012) 045139.

[5] R. Savit, *Duality in field theory and statistical systems*, *Rev. Mod. Phys.* **52** (1980) 453.

[6] N. Prokof'ev and B. Svistunov, *Worm algorithms for classical statistical models*, *Phys. Rev. Lett.* **87** (2001) 160601.

[7] A. Krizhevsky and G. Hinton, *Learning multiple layers of features from tiny images*, *Master's thesis, Department of Computer Science, University of Toronto* (2009) .

PoS(LATTICE2018)245