

Learning Renormalization Group Flows of Ising Models

Jay Shen ¹ Ying-Jer Kao ² Wanda Hou ³ Yi-Zhuang You ³

¹University of Chicago

²National Taiwan University

³UC San Diego

Introduction

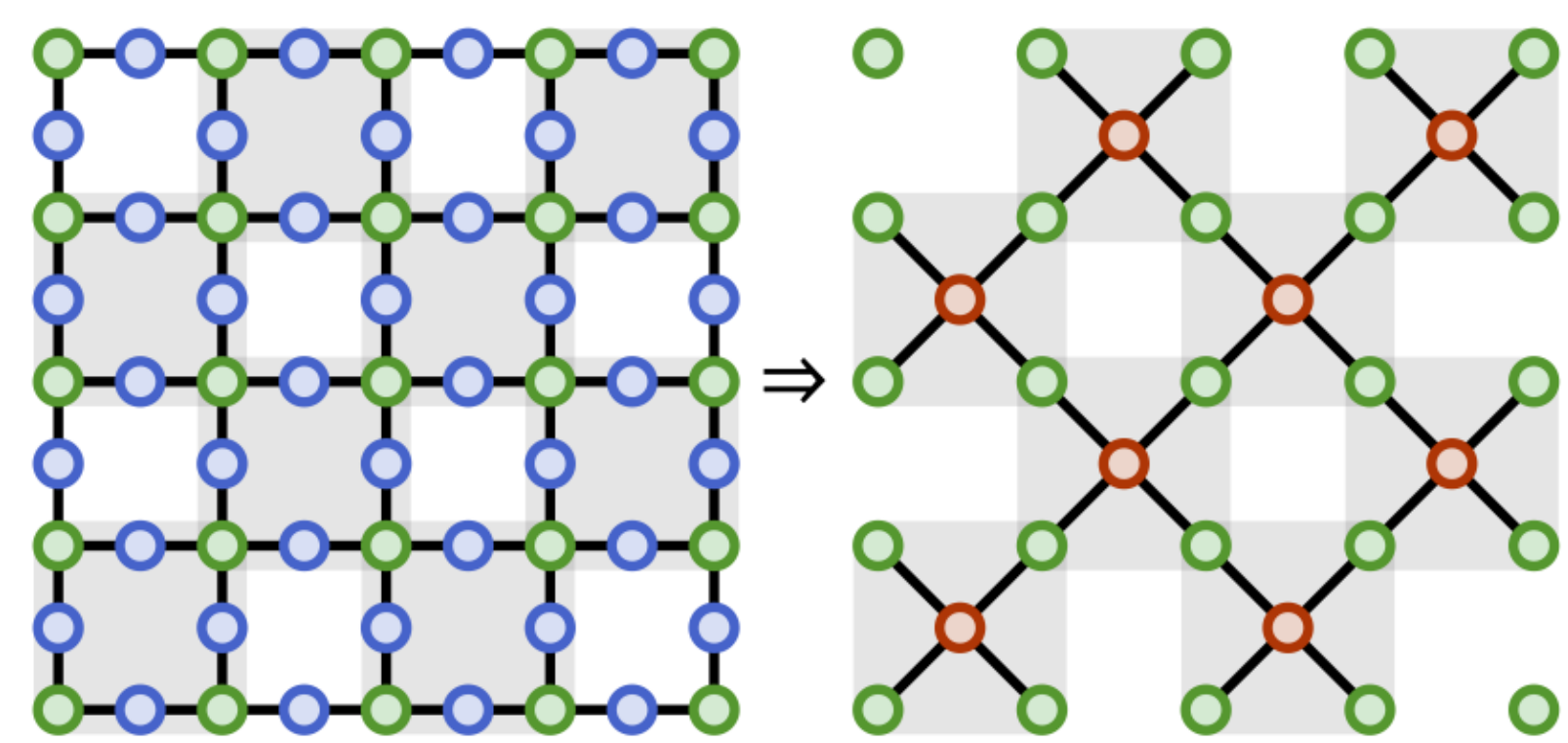
Real space renormalization is a theoretically powerful but practically difficult technique for investigating scale and phase behavior in physical systems. For even the simplest problems, formulating the so-called renormalization group (RG) flow is incredibly difficult but, once accurately described, incredibly useful. Hou et al. proposed the MLRG method which approximates RG flows of Ising models automatically using neural networks. Here, we refine their algorithm, discuss its design space, and demonstrate some successes and limitations.

Theory

The isotropic homogenous Ising model is a graph with N nodes $\sigma_i \in \{-1, 1\}$ and adjacency list \mathcal{E} , with $(i, j) \in \mathcal{E}$ indicating σ_i and σ_j are neighboring nodes on the lattice. The values of neighboring σ_i and σ_j correlated according to coupling constant J .

A renormalization transformation defines a coarse-grained lattice (primed) with $N' < N$ nodes and solves for the coupling constant J' which best matches the fine-grained system's behavior:

$$H = -J \sum_{(i,j) \in \mathcal{E}} \sigma_i \sigma_j \Rightarrow H' = -J' \sum_{(i,j) \in \mathcal{E}'} \sigma_i \sigma_j$$



In principle, there exists a function f , called the renormalization group flow, which maps fine-grained coupling constants J to coarse-grained coupling constants J' by predicting the difference $f(J) = J' - J$:

Rather than deriving this f analytically, we learn an approximation $f_\theta(J)$ using a neural network. The neural network parameters are adjusted to minimize the KL divergence between Ising model Gibbs distributions:

$$\text{Loss} = D_{KL} \left(\frac{1}{Z} e^{H(J)} \left\| \frac{1}{Z'} e^{H'(f_\theta(J)+J)} \right\| \right)$$

The Improved MLRG Algorithm

1. Sample values of J

Hamiltonian Monte Carlo (HMC) [?] is used to refine the zeros (critical points) of f_θ by sampling from the unnormalized probability distribution:

$$\tilde{P}(J) = \exp \left[-\alpha J^2 - \beta \|f_\theta(J)\|^2 \right]$$

For stability, the potential is annealed in during training by increasing β from 0 to some β_T and decreasing α from α_0 to 0.

2. Forward pass to get J'

$$J' = J + f_\theta(J)$$

3. Evaluate the error

We implement the Ising lattices as Restricted Boltzmann Machines (RBMs) with visible and hidden spins, allowing for easy definition of the renormalization transformation and the appropriate KL divergence. We use the common contrastive approximation for the loss:

$$D_{KL} \left(\frac{1}{Z} e^{H(J)} \left\| \frac{1}{Z'} e^{H'(f_\theta(J)+J)} \right\| \right) \approx F_{coarse}(\vec{v}_{fine}) - F_{coarse}(\vec{v}_{coarse})$$

Here, F_{coarse} is the free energy function of the coarse-grained RBM and \vec{v}_{fine} and \vec{v}_{coarse} are visible spins Gibbs-sampled from the fine and coarse-grained RBMs, respectively.

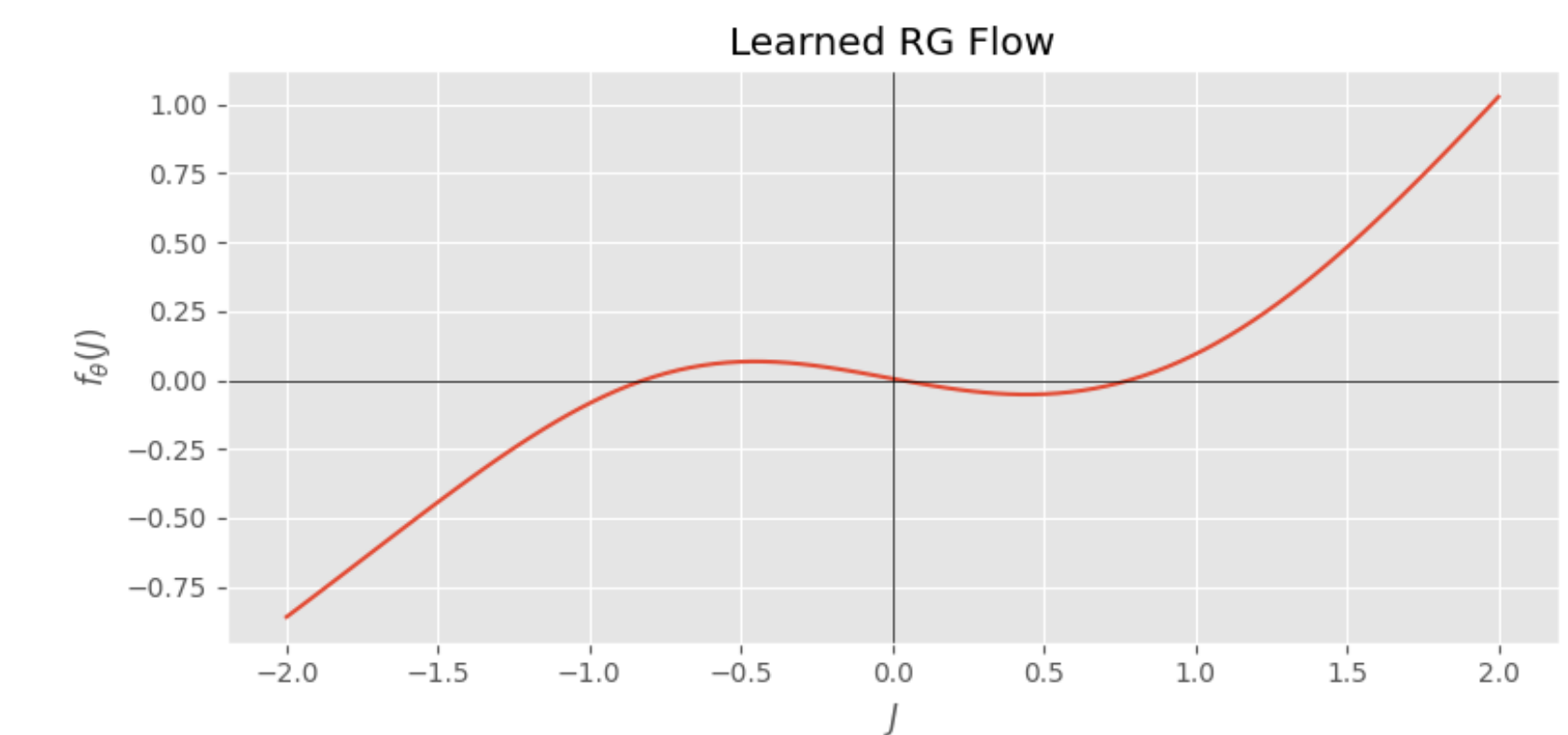
4. Backpropagate error to θ

Both the contrastive and the full KL losses are composed functions of $f_\theta(J)$, so we can backpropagate to adjust the parameters θ .

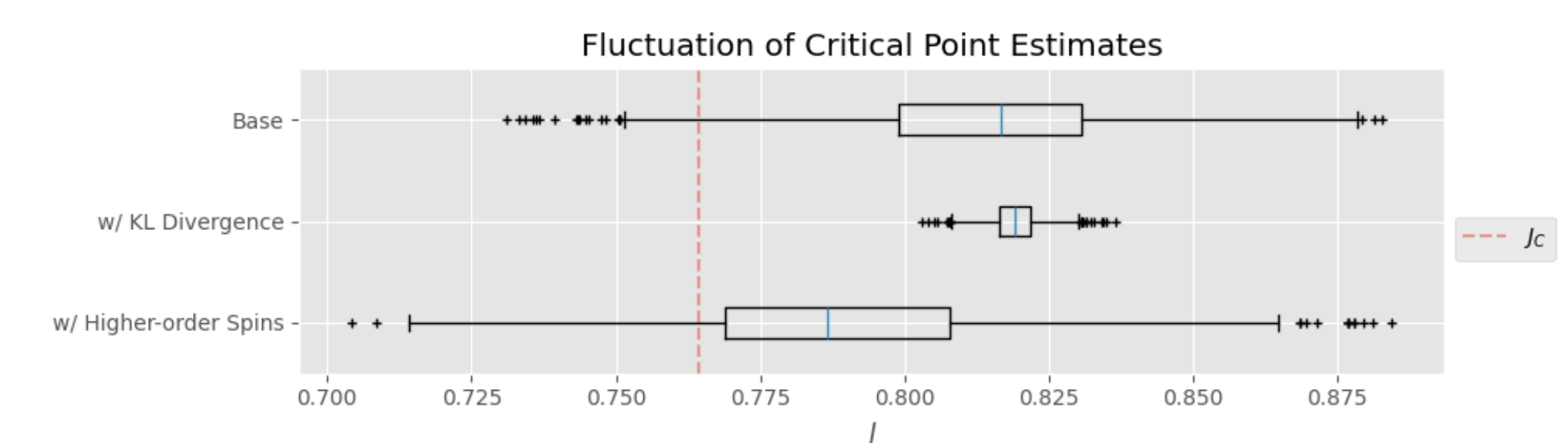
Notable Elements of the Algorithmic Space

- Hou et al. learn the C-theorem monotone $C(J)$, whereas we learn $f(J) = \nabla C(J)$, which is substantially faster.
- Hou et al. use higher-order spins to obtain better critical point estimates. We adopt this approach, and show that it is compatible with our other modifications.
- For small lattices, the KL-divergence is tractable to compute but mitigates only variance not bias.
- For training stability, we find the number of fine-grained Gibbs steps and the number of HMC trajectories should be above some threshold. Tuning hyperparameters is minimally effective.

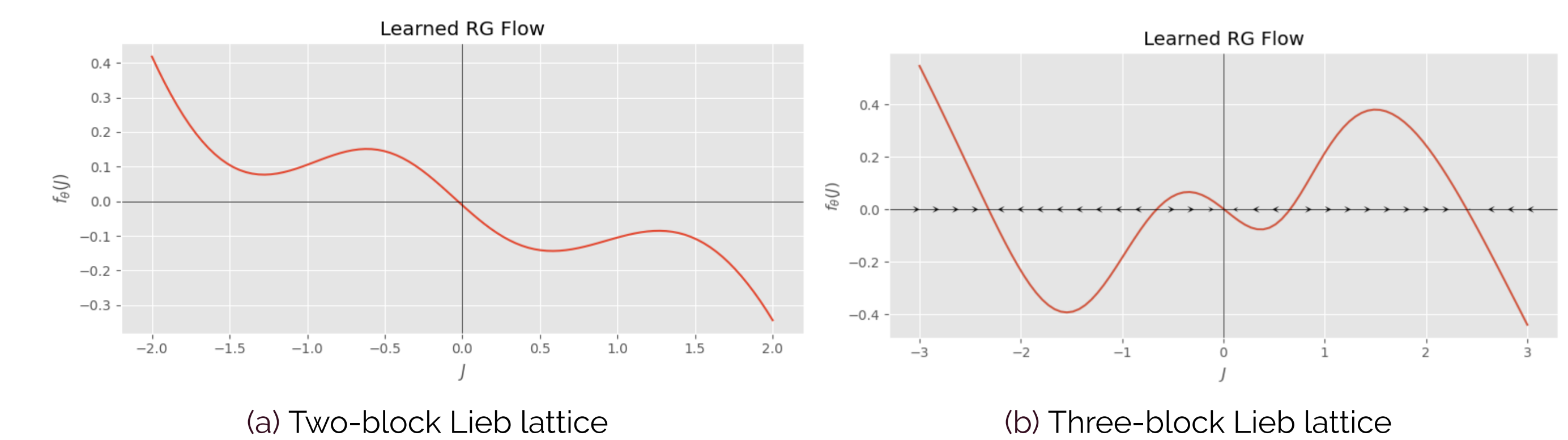
Results



Training to convergence yields a flow with expected properties of odd parity, fixed points at 0 and $\pm\infty$, and a pair of finite critical points.



We obtain estimates close to the actual critical point. Using the KL divergence and higher-order spins can achieve improvements in variance and bias, respectively.



The MLRG approach is limited in that it does not learn the same flows for lattices that should, in principle, be thermodynamically identical.

It is interesting to note that there appears to be an even-odd effect among learned flows—the algorithm finds critical points only for odd lattices.

References

- [1] Simon Duane, A.D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. 1987.
- [2] Geoffrey E. Hinton. Training products of experts by minimizing contrastive divergence. 2002.
- [3] Geoffrey E. Hinton. *A Practical Guide to Training Restricted Boltzmann Machines*. 2012.
- [4] Wanda Hou and Yi-Zhuang You. Machine learning renormalization group for statistical physics, 2023.
- [5] A. B. Zamolodchikov. "Irreversibility" of the flux of the renormalization group in a 2d field theory. 1986.