Journal of Computational Mathematics
Vol.xx, No.x, 2023, 1–9.
http://www.global-sci.org/jcm
doi:10.4208/jcm.xxx

# ANNEALED IMPORTANCE SAMPLING FOR ISING MODELS WITH MIXED BOUNDARY CONDITIONS*

Lexing Ying[1)]
*Department of Mathematics, Stanford University, Stanford, CA 94305, USA*
*Email: lexing@stanford.edu*

**Abstract**

This note introduces a method for sampling Ising models with mixed boundary conditions. As an application of annealed importance sampling and the Swendsen-Wang algorithm, the method adopts a sequence of intermediate distributions that keeps the temperature fixed but turns on the boundary condition gradually. The numerical results show that the variance of the sample weights is relatively small.

*Mathematics subject classification:* 82B20, 82B80.
*Key words:* Ising model, Annealed importance sampling, Swendsen-Wang algorithm.

## 1. Introduction

This note is concerned with the Monte Carlo sampling of Ising models [7, 12] with mixed boundary conditions. Consider a graph $G = (V, E)$ with the vertex set $V$ and the edge set $E$. We assume that $V = I \cup B$, where $I$ is the subset of interior vertices and $B$ the subset of boundary vertices. Throughout the note, we use $i, j$ to denote the vertices in $I$ and $b$ for the vertices in $B$. In addition, $ij \in E$ denotes an edge between two interior vertices $i$ and $j$, while $ib \in E$ denotes an edge between an interior vertex $i$ and a boundary vertex $b$. The boundary condition is specified by $f = (f_b)_{b\in B}$ with $f_b = \pm 1$.

A spin configuration $s = (s_i)_{i\in I}$ over the interior vertex set $I$ is an assignment of $\pm 1$ value to each vertex $i \in I$. The energy of the spin configuration $s$ is given by the Hamiltonian function $H(s)$ defined via

$$H(s) = -\sum_{ij\in E} s_i s_j - \sum_{ib\in E} s_i f_b.$$

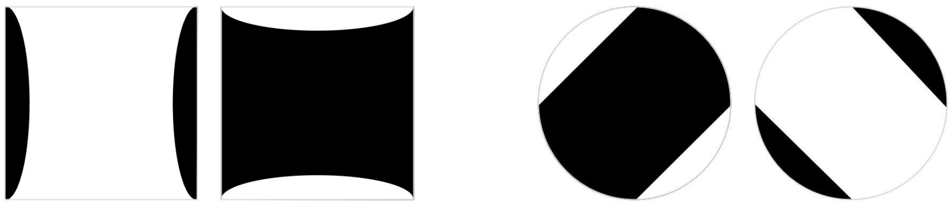
At an inverse temperature $\beta > 0$, the configuration probability of $s = (s_i)_{i\in I}$ is given by the Gibbs or Boltzmann distribution

$$p_{\mathrm{I}}(s) = \frac{e^{-\beta H(s)}}{Z_\beta} \sim \exp\left(\beta \sum_{ij\in E} s_i s_j + \beta \sum_{ib\in E} s_i f_b\right), \tag{1.1}$$

where $Z_\beta = \sum_s e^{-\beta H(s)}$ is the renormalization constant or the partition function. More detailed discussions about the Ising models can be found for example in [3, 11].

A key feature of this Ising model is that, for certain mixed boundary conditions, the distribution (1.1) exhibits macroscopically different profiles below the critical temperature. Fig. 1.1 showcases two such examples. On the left, the square Ising lattice has the +1 condition on

* Received July 26, 2022 / Revised version received November 2, 2022 / Accepted November 28, 2022 / Published online xx xx, 2023 /
1) Corresponding author



(a) (b)

Fig. 1.1. Ising models with mixed boundary conditions. (a) A square model. (b) A model support on a disk. In each case, a mixed boundary condition is specified and the model exhibits two dominant profiles on the macroscopic scale.

the vertical sides but the $-1$ condition on the horizontal sides. The two dominant macroscopic profiles are a $-1$ cluster linking two horizontal sides and a $+1$ cluster linking two vertical sides, shown in Fig. 1.1(a). On the right, a triangular Ising lattice supported on a disk has the $+1$ condition on two disjoint arcs and the $-1$ condition on the other two. Its two dominant profiles are given in Fig. 1.1(b). Notice that in each case, the two dominant profiles have comparable probability. Hence, it is important for any sampling algorithm to transition between these macroscopically different profiles efficiently.

One of the most well-known methods for sampling Ising models is the Swendsen-Wang algorithm [13], which will be briefly reviewed in Section 2. For Ising models with free boundary condition for example, the Swendsen-Wang algorithm exhibits rapid mixing for all temperatures. However, for the mixed boundary conditions shown in Fig. 1.1, the Swendsen-Wang algorithm experiences slow convergence under the critical temperatures, i.e., $T < T_c$ or equivalently $\beta > \beta_c$. The reason is that, for such a boundary condition, the energy barrier between the two dominant profiles is much higher than the typical energy fluctuations. In other words, the Swendsen-Wang algorithm needs to break a macroscopic number of edges between aligned adjacent spins in order to transition from one dominant profile to the other. However, breaking so many edges simultaneously is an event with exponentially small probability when the mixed boundary condition is specified.

Annealed importance sampling is a method proposed by Neal [10], designed for sampling distributions with multiple modes. The main idea is to

(1) introduce an easily-to-sample initial distribution,

(2) design a sequence of (typically temperature-dependent) intermediate distributions that interpolates between the initial and the target distributions,

(3) generate sample paths that connects the simple initial distribution and the hard target distribution,

(4) compute a path-dependent scalar to weight the samples at the target distribution.

Annealed importance sampling has been widely applied in Bayesian statistics and data assimilation for sampling and estimating partition functions.

In this note, we address the problem of sampling (1.1) by combining the Swendsen-Wang algorithm with annealed importance sampling. The main novelty of our approach is that, instead of adjusting the temperature, we freeze the temperature and adjust the mixed boundary condition.

**Related works.** Alexander and Yoshida [1, 2] studied the spectral gap of the 2D Ising models with mixed boundary conditions. In [14], the double flip move is introduced for models with mixed boundary conditions that enjoy exact or approximate symmetry. When combined with the Swendsen-Wang algorithm, it can accelerate the mixing of these Ising model under the critical temperature significantly. However, it only applies to problem with exact or approximate symmetries, but not more general settings.

Recently Gheissari and Lubetzky [6] studied the effect of the boundary condition for the 2D Potts models at the critical temperature. Chatterjee and Diaconis [4] showed that most of the deterministic jumps can accelerate the Markov chain mixing when the equilibrium distribution is uniform.

The rest of the note is organized as follows. Sections 2 and 3 review the Swendsen-Wang algorithm and annealed importance sampling, respectively. Section 4 describes the algorithm and provides several numerical examples. Section 5 discusses some future directions.

## 2. Swendsen-Wang Algorithm

In this section, we briefly review the Swendsen-Wang algorithm. First, notice that

$$p_{\mathrm{I}}(s) \sim \exp\left(\beta \sum_{ij\in E} s_i s_j + \beta \sum_{ib\in E} f_b s_i\right) = \exp\left(\beta \sum_{ij\in E} s_i s_j + \beta \sum_{i\in I}\left(\sum_{ib\in E} f_b\right) s_i\right).$$

Therefore, we can interpret $h_i \equiv \sum_{ib\in E} f_b$ as an external field and view the mixed boundary condition problem as a special case of the model with external field $h = (h_i)_{i\in I}$

$$p_{\mathrm{I}}(s) \sim \exp\left(\beta \sum_{ij\in E} s_i s_j + \beta \sum_{i\in I} h_i s_i\right). \tag{2.1}$$

This viewpoint simplifies the discussion and the description of the Swendsen-Wang algorithm is given below under this setting.

The Swendsen-Wang algorithm is a Markov Chain Monte Carlo method for sampling $p_{\mathrm{I}}(s)$. In each iteration, it generates a new configuration $(t_i)_{i\in I}$ based on the current configuration $(s_i)_{i\in I}$ with two substeps:

- Generate an edge configuration $w = (w_{ij})_{ij\in E}$. If the spin values $s_i$ and $s_j$ are different, set $w_{ij} = 0$. If $s_i$ and $s_j$ are the same, $w_{ij}$ is sampled from the Bernoulli distribution $\mathrm{Ber}(1 - e^{-2\beta})$, i.e., equal to 1 with probability $1 - e^{-2\beta}$ and 0 with probability $e^{-2\beta}$.

- Regards all edges $ij \in E$ with $w_{ij} = 1$ as linked. Compute the connected components. For each connected component $\gamma$, define $h_\gamma = \sum_{i\in\gamma} h_i$. Set the spins $(t_i)_{i\in\gamma}$ of the new configuration $t$ to 1 with probability $e^{\beta h_\gamma}/(e^{\beta h_\gamma} + e^{-\beta h_\gamma})$ and to 0 with probability $e^{-\beta h_\gamma}/(e^{\beta h_\gamma} + e^{-\beta h_\gamma})$.

Associated with (2.1), two other probability distributions are important for analyzing the Swendsen-Wang algorithm [5]. The first one is the joint vertex-edge distribution

$$p_{\mathrm{IE}}(s, w) \sim \prod_{ij\in E} \left(\left(1 - e^{-2\beta}\right)\delta_{s_i=s_j}\delta_{w_{ij}=1} + e^{-2\beta}\delta_{w_{ij}=0}\right) \cdot e^{\beta \sum_{i\in I} h_i s_i}. \tag{2.2}$$

The second one is the edge distribution

$$p_{\mathrm{E}}(w) \sim \prod_{w_{ij}=1} \left(1 - e^{-2\beta}\right) \prod_{w_{ij}=0} e^{-2\beta} \cdot \prod_{\gamma \in \mathcal{C}_w} \left(e^{-\beta h_\gamma} + e^{\beta h_\gamma}\right), \tag{2.3}$$

where $\mathcal{C}_w$ is the set of the connected components induced by $w$.

Summing $p_{\mathrm{IE}}(s,w)$ over $s$ or $w$ gives the following two identities:

$$\sum_s p_{\mathrm{IE}}(s,w) = p_{\mathrm{E}}(w), \quad \sum_w p_{\mathrm{IE}}(s,w) = p_{\mathrm{I}}(s) \tag{2.4}$$

(see for example [5]). A direct consequence of (2.4) is that the Swendsen-Wang algorithm can be viewed as a data augmentation method [8]: The first substep samples the edge configuration $w$ conditioned on the spin configuration $s$, while the second substep samples a new spin configuration conditioned on the edge configuration $w$.

Eq. (2.4) also imply that Swendsen-Wang algorithm satisfies the detailed balance. To see this, let us fix two spin configurations $s$ and $t$ and consider the transition between them. Since such a transition in the Swendsen-Wang move happens via an edge configuration $w$, it is sufficient to show

$$p_{\mathrm{I}}(s) P_w(s,t) = p_{\mathrm{I}}(t) P_w(t,s)$$

for any compatible edge configuration $w$. Here $P_w(s,t)$ is the transition probability from $s$ to $t$ via $w$ and similarly for $P_w(t,s)$. Since the transition probabilities from $w$ to the spin configurations $s$ and $t$ are proportional to $e^{\beta \sum_i h_i s_i}$ and $e^{\beta \sum_i h_i t_i}$ respectively, it reduces to showing

$$p_{\mathrm{I}}(s) P(s,w) e^{\beta \sum_i h_i t_i} = p_{\mathrm{I}}(t) P(t,w) e^{\beta \sum_i h_i s_i}, \tag{2.5}$$

where $P(s,w)$ is the probability of obtaining the edge configuration $w$ from $s$ and similarly for $P(t,w)$.

Using (2.1), this is equivalent to showing

$$\begin{aligned} & e^{\beta \sum_{ij \in E} s_i s_j + \beta \sum_{i \in I} h_i s_i} P(s,w) e^{\beta \sum_{i \in I} h_i t_i} \\ = & e^{\beta \sum_{ij \in E} t_i t_j + \beta \sum_{i \in I} h_i t_i} P(t,w) e^{\beta \sum_{i \in I} h_i s_i}. \end{aligned} \tag{2.6}$$

The next observation is that

$$e^{\beta \sum_{ij \in E} s_i s_j} P(s,w) = e^{\beta \sum_{ij \in E} t_i t_j} P(t,w), \tag{2.7}$$

i.e., this quantity is independent of the spin configuration. To see this, notice first that if an edge $ij \in E$ has configuration $w_{ij} = 1$ then $s_i = s_j$. Second, if $ij \in E$ has configuration $w_{ij} = 0$, then $s_i$ and $s_j$ can either be the same or different. In the former case, the contribution to $e^{\beta \sum_{ij \in E} s_i s_j} P(s,w)$ from $ij$ is $e^{2\beta} \cdot e^{-2\beta} = 1$ up to a uniform normalization constant. In the latter case, the contribution is also $1 \cdot 1 = 1$ up to the same uniform constant. After canceling the two terms of (2.7) in (2.6), proving (2.5) is equivalent to

$$e^{\beta \sum_{i \in I} h_i s_i} \cdot e^{\beta \sum_{i \in I} h_i t_i} = e^{\beta \sum_{i \in I} h_i t_i} \cdot e^{\beta \sum_{i \in I} h_i s_i},$$

which is trivial.

The Swendsen-Wang algorithm however does not encourage transitions between the macroscopic profiles shown for example in Fig. 1.1. With certain mixed boundary conditions, such a transition requires breaking a macroscopic number of edges between aligned adjacent spins, which has an exponentially small probability.

## 3. Annealed Importance Sampling

Given a target distribution $p(s)$ that is hard to sample directly, annealed importance sampling (AIS), proposed by Neal [10], introduces a sequence of distributions

$$p_0(\cdot), \dots, p_L(\cdot) \equiv p(\cdot),$$

where $p_0(\cdot)$ is easy to sample and each $p_l(\cdot)$ is associated with a detailed balance Markov Chain $T_l(s,t)$, i.e.,

$$p_l(s)T_l(s,t) = p_l(t)T_l(t,s). \tag{3.1}$$

Though the detailed balance condition can be relaxed, it simplifies the description.

Given this sequence of intermediate distributions, AIS proceeds as follows:

1. Sample a configuration $s_{1/2}$ from $p_0(\cdot)$.
2. For $l = 1, \dots, L-1$, take one step (or a few steps) of $T_i(\cdot,\cdot)$ (associated with the distributions $p_l(\cdot)$) from $s_{l-1/2}$. Let $s_{l+1/2}$ be the resulting configuration.
3. Set $s := s_{L-1/2}$.
4. Compute the weight
$$w := \frac{p_1(s_{1/2})}{p_0(s_{1/2})} \cdots \frac{p_L(s_{L-1/2})}{p_{L-1}(s_{L-1/2})}.$$

The claim is that the configuration $s$ with weight $w$ samples the target distribution $p_L(\cdot)$. To see this, consider the path $(s_{1/2}, \dots, s_{L-1/2})$. This path is generated with probability

$$p_0(s_{1/2})T_1(s_{1/2}, s_{3/2}) \cdots T_{L-1}(s_{L-3/2}, s_{L-1/2}).$$

Multiplying this with $w$ and using the detailed balance (3.1) of $T_l$ gives

$$\begin{aligned} & p_0(s_{1/2})T_1(s_{1/2}, s_{3/2}) \cdots T_{L-1}(s_{L-3/2}, s_{L-1/2}) \cdot \frac{p_1(s_{1/2})}{p_0(s_{1/2})} \cdots \frac{p_L(s_{L-1/2})}{p_{L-1}(s_{L-1/2})} \\ = {} & p_L(s_{L-1/2})T_{L-1}(s_{L-1/2}, s_{L-3/2}) \cdots T_1(s_{3/2}, s_{1/2}), \end{aligned}$$

which is the probability of going backward, i.e., starting from a sample $s_{L-1/2}$ of $p_L(\cdot) \equiv p(\cdot)$. Taking the margin of the last slot $s_{L-1/2}$ proves that $s := s_{L-1/2}$ with weight $w$ samples the distribution $p_L(\cdot) \equiv p(\cdot)$.

## 4. Algorithm and Results

Our proposal is to combine AIS with the Swendsen-Wang algorithm in order for sampling Ising models with mixed boundary conditions. The key ingredients are:

- Set the initial $p_0(\cdot)$ to be
$$p_0(s) \sim \exp\left(\beta \sum_{ij\in E} s_i s_j\right).$$
This initial distribution has no external field and hence can be sampled efficiently with the Swendsen-Wang algorithm.

- Choose a monotone sequence of $(\theta_l)_{0\le l\le L}$ with $\theta_0 = 0$ and $\theta_L = 1$ and set at level $l$

$$p_l(s) \sim \exp\left(\beta \sum_{ij\in E} s_i s_j + \beta \sum_{i\in I} (\theta_l h_i) s_i\right),$$

  i.e., the distribution with external field $\theta_h h$. The Markov transition matrix $T_l(\cdot,\cdot)$ is implemented with the Swendsen-Wang algorithm associated with $p_l(\cdot)$. As proven in Section 2, $T_l(\cdot,\cdot)$ satisfies the detailed balance.

Below we demonstrate the performance of the proposed method with several examples. In each example, $K = 500$ samples $(s^{(k)}, w^{(k)})_{1\le k\le K}$ are generated. For each sample $s^{(k)}$, the initial choice $s^{(k)}_{1/2}$ is obtained by running 100 iterations of the Swendsen-Wang algorithm at $p_0(\cdot)$. In our implementation, the monotone sequence $(\theta_l)_{0\le l\le L}$ is chosen to be an equally spaced sequence with $L = 400$, i.e., $\theta_l = l/L$. Although the equally-spaced sequence is not necessarily the ideal choice in terms of variance minimization, it work reasonably well numerically.

In order to monitor the variance of the algorithm, we record the weight history within each level $l$

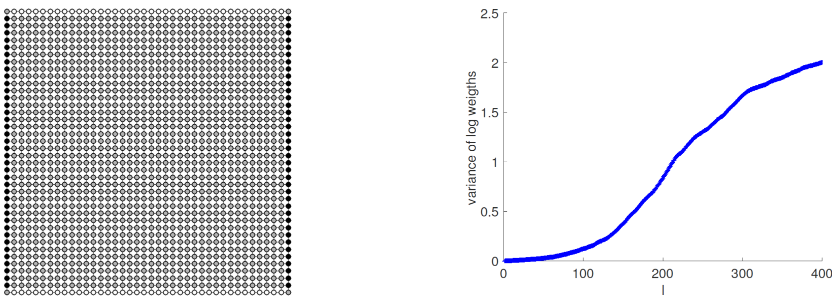
$$w_l^{(k)} := \frac{p_1\left(s^{(k)}_{1/2}\right)}{p_0\left(s^{(k)}_{1/2}\right)} \cdots \frac{p_l\left(s^{(k)}_{l-1/2}\right)}{p_{l-1}\left(s^{(k)}_{l-1/2}\right)}$$

for $l = 1,\dots,L$. These weights are then normalized at each level $l$

$$\tilde{w}_l^{(k)} = \frac{w_l^{(k)}}{K^{-1}\sum_{g=1}^{K} w_l^{(g)}}.$$

Following the practice of [10], we shall plot the empirical variance of the logarithm of the normalized weights $\mathrm{Var}[\{\log \tilde{w}_l^{(k)}\}]$ as a function of level $l = 1,\dots,L$. On the other hand, the sample efficiency, a quantity between 0 and 1, is defined as $(1 + \mathrm{Var}[\{\tilde{w}_L^{(k)}\}])^{-1}$ using the empirical variance of the normalized weights.

**Example 4.1.** The Ising model is a square lattice, as shown in Fig. 4.1(a). The mixed boundary condition is $+1$ on the two vertical sides and $-1$ on the two horizontal sides. The experiments are performed for the problem size $n_1 = n_2 = 40$ at the inverse temperature $\beta = 0.5$. Fig. 4.2(b) plots the empirical variance of the logarithm of the normalized weights,

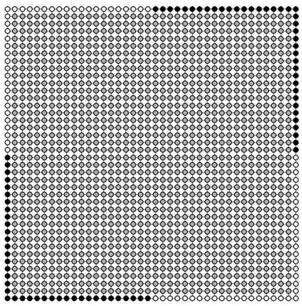



(a) (b)

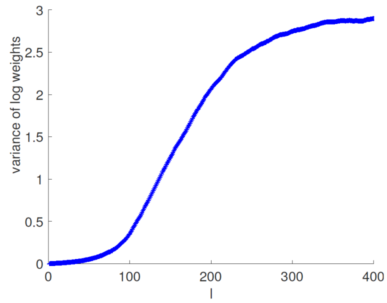
Fig. 4.1. (a) The lattice along with the external field. (b) The empirical variance of the logarithm of the normalized weights as a function of level $l$.

$\mathrm{Var}[\{\log \tilde{w}_l^{(k)}\}]$, as a function of the level $l$. The sample efficiency $(1 + \mathrm{Var}[\{\tilde{w}_L^{(k)}\}])^{-1}$ is 0.26, which translates to $L(1 + \mathrm{Var}[\{\tilde{w}_L^{(k)}\}]) \approx 1530$ Swendsen-Wang iterations per effective sample.

**Example 4.2.** The Ising lattice is again a square as shown in Fig. 4.2(a). The mixed boundary condition is $+1$ in the first and third quadrants but $-1$ in the second and fourth quadrants. The experiments are performed for the problem size $n_1 = n_2 = 40$ at the inverse temperature $\beta = 0.5$. Fig. 4.2(b) plots $\mathrm{Var}[\{\log \tilde{w}_l^{(k)}\}]$ as a function of the level $l$. The sample efficiency $(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}])^{-1}$ is 0.09, which translates to about $L(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}]) \approx 4470$ Swendsen-Wang iterations per effective sample.



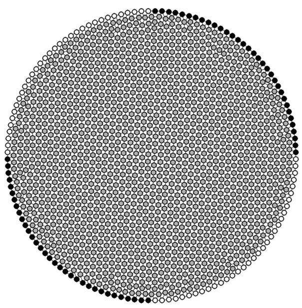
(a)



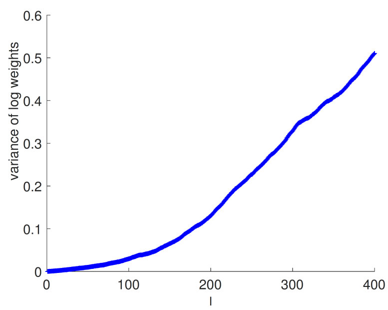

(b)

Fig. 4.2. (a) The lattice along with the external field. (b) The empirical variance of the logarithm of the normalized weights as a function of level $l$.

**Example 4.3.** The Ising model is a random quasi-uniform triangular lattice supported on the unit disk, as shown in Fig. 4.3(a). The lattice does not have rotation and reflection symmetry due to the random triangulation. The mixed boundary condition is $+1$ in the first and third quadrants but $-1$ in the second and fourth quadrants. The experiments are performed with a fine triangulation with mesh size $h = 0.05$ at the inverse temperature $\beta = 0.3$. Fig. 4.3(b) plots $\mathrm{Var}[\{\log \tilde{w}_l^{(k)}\}]$ as a function of the level $l$, which remains quite small. The sample efficiency $(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}])^{-1}$ is 0.67, which translates to about $L(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}]) \approx 600$ Swendsen-Wang iterations per effective sample.



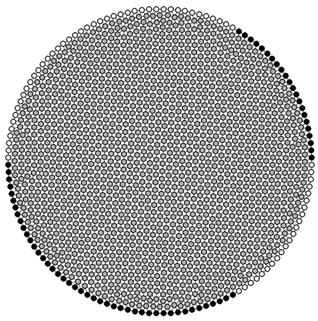
(a)



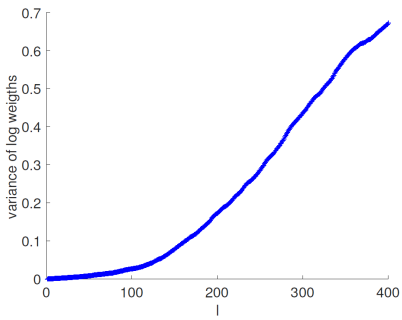

(b)

Fig. 4.3. (a) The lattice along with the external field. (b) The empirical variance of the logarithm of the normalized weights as a function of level $l$.

**Example 4.4.** The Ising model is again a random quasi-uniform triangular lattice supported on the unit disk, as shown in Fig. 4.4(a). The mixed boundary condition is $+1$ on the two arcs with angle in $[0, \pi/3]$ and $[\pi, 5\pi/3]$ but $-1$ on the remaining two arcs. The experiments are performed with a fine triangulation with mesh size $h = 0.05$ at the inverse temperature $\beta = 0.3$. Fig. 4.4(b) plots $\mathrm{Var}[\{\log \tilde{w}_l^{(k)}\}]$ as a function of the level $l$, which remains quite small. The sample efficiency $(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}])^{-1}$ is 0.55, which translates to about $L(1+\mathrm{Var}[\{\tilde{w}_L^{(k)}\}]) \approx 730$ Swendsen-Wang iterations per effective sample.



(a)




(b)

Fig. 4.4. (a) The lattice along with the external field. (b) The empirical variance of the logarithm of the normalized weights as a function of level $l$.

## 5. Discussions

This note introduces a method for sampling Ising models with mixed boundary conditions. As an application of annealed importance sampling and the Swendsen-Wang algorithm, the method adopts a sequence of intermediate distributions that fixes the temperature but turns on the boundary condition gradually. The numerical results show that the variance of the sample weights remain to be relatively small.

There are many unanswered questions. First, the sequence of $(\theta_l)_{0\le l\le L}$ that controls the intermediate distributions is empirically specified to be equally-spaced. Two immediate questions are (1) what the optimal $(\theta_l)_{0\le l\le L}$ sequence is and (2) whether there is an efficient algorithm for computing it.

Second, historically annealed importance sampling is introduced following the work of tempered transition [9]. We have implemented the current idea within the framework of tempered distribution. However, the preliminary results show that it is less effective compared to annealed importance sampling. A more thorough study in this direction will be useful.

Finally, annealed importance sampling (AIS) is a rather general framework. For a specific application, the key to efficiency is the choice of the distribution $p_0(\cdot)$: it should be easy-to-sample, while at the same time sufficiently close to the target distribution $p(\cdot)$. However, since the target distribution is hard-to-sample, these two objectives often compete with each other. Since there are many other hard-to-sample models in statistical mechanics, a potentially fruitful direction of research is to apply AIS with cleverly chosen $p_0(\cdot)$ to these models.

**Acknowledgments.** This work is partially supported by NSF Grant No. DMS 2011699.

## References

[1] K.S. Alexander, The spectral gap of the 2-D stochastic Ising model with nearly single-spin boundary conditions, *J. Stat. Phys.*, **104**:1 (2001), 59–87.

[2] K.S. Alexander and N. Yoshida, The spectral gap of the 2-D stochastic Ising model with mixed boundary conditions, *J. Stat. Phys.*, **104**:1 (2001), 89–109.

[3] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Elsevier, 2016.

[4] S. Chatterjee and P. Diaconis, Speeding up Markov chains with deterministic jumps, *Probab. Theory Related Fields*, **178**:3 (2020), 1193–1214.

[5] R.G. Edwards and A.D. Sokal, Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm, *Phys. Rev. D*, **38**:6 (1988), 2009.

[6] R. Gheissari and E. Lubetzky, The effect of boundary conditions on mixing of 2D Potts models at discontinuous phase transitions, *Electron. J. Probab.*, **23** (2018), 1–30.

[7] E. Ising, *Beitrag zur Theorie des Ferro-und Paramagnetismus*, PhD Thesis, University of Hamburg, 1924.

[8] J.S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer, 2001.

[9] R.M. Neal, Sampling from multimodal distributions using tempered transitions, *Stat. Comput.*, **6**:4 (1996), 353–366.

[10] R.M. Neal, Annealed importance sampling, *Stat. Comput.*, **11**:2 (2001), 125–139.

[11] H. Nishimori and G. Ortiz, *Elements of Phase Transitions and Critical Phenomena*, Oxford University Press, 2010.

[12] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, *Phys. Rev.*, **65**:3-4 (1944), 117–149.

[13] R.H. Swendsen and J.S. Wang, Nonuniversal critical dynamics in Monte Carlo simulations, *Phys. Rev. Lett.*, **58**:2 (1987), 86–88.

[14] L. Ying, Double Flip Acceleration for Ising Models with Mixed Boundary Conditions, *Preprint*, (2022).