Partition Function of the Ising Model via Factor Graph Duality

Mehdi Molkaraie
University of Waterloo
Dept. of Statistics & Actuarial Science
Waterloo, Canada N2L 3G1
mmolkaraie@uwaterloo.ca

Hans-Andrea Loeliger
ETH Zurich
Dept. of Information Technology & Electrical Engineering
8092 Zürich, Switzerland
loeliger@isi.ee.ethz.ch

Abstract—The partition function of a factor graph and the partition function of the dual factor graph are related to each other by the normal factor graph duality theorem. We apply this result to the classical problem of computing the partition function of the Ising model. In the one-dimensional case, we thus obtain an alternative derivation of the (well-known) analytical solution. In the two-dimensional case, we find that Monte Carlo methods are much more efficient on the dual graph than on the original graph, especially at low temperature.

I. INTRODUCTION

We consider the problem of computing the partition function of one-dimensional (1D) and finite-size two-dimensional (2D) Ising models. In particular, we propose a method to compute the partition function of finite-size 2D Ising models at low temperature by performing Markov chain Monte Carlo methods on the dual factor graph.

The problem setup is as follows. Let $X_1, X_2, \ldots, X_N$ be random variables, each taking its values in $\mathcal{X} = \{0, 1\}$. (In statistical physics, variables are usually considered as particles which take on two possible states, e.g., spin up $\uparrow$ and spin down $\downarrow$). Let x_i represent a possible realization of X_i and let x stand for a configuration $(x_1, x_2, \ldots, x_N)$.

In 1D Ising models, as shown in Figs. 1 and 2, variables (particles) are considered as a chain of sites on a line. In 2D Ising models, variables are arranged on the sites of a 2D lattice, as depicted in Fig. 3.

We assume that only *adjacent (nearest neighbor)* variables interact with each other. In the absence of an external field, we define the energy of a configuration x as [1]

$$E(x) \triangleq - \sum_{k,\ell \text{ adjacent}} J_{k,\ell}\big([x_k = x_\ell] - [x_k \neq x_\ell]\big) \qquad (1)$$

where the sum runs over all the (unordered) adjacent pairs (k, ℓ) and $[\cdot]$ denotes the Iverson bracket [2, p. 24], which evaluates to one if the condition in the bracket is satisfied and to zero otherwise.

The real coupling parameter $J_{k,\ell}$, controls the strength of the interaction between (x_k, x_ℓ). If $J_{k,\ell} > 0$, the model is known as a ferromagnetic Ising model. The model is called antiferromagnetic if $J_{k,\ell} < 0$, see [1], [3]. If the couplings can be both positive or negative (e.g., chosen at random according

to some distribution) the model is usually known as an Ising spin glass, see [4, Chapter 12].

In thermal equilibrium, the probability of a configuration x, is given by the Boltzmann distribution, defined as [1]

$$p_{\mathrm{B}}(x) \triangleq \frac{e^{-\beta E(x)}}{Z} \qquad (2)$$

where Z is the *partition function* (normalization constant) and $\beta \triangleq \frac{1}{k_{\mathrm{B}}T}$, where temperature is denoted by T, and k_{B} is the Boltzmann constant.

In the rest of this paper, we will assume $\beta = 1$. With this assumption, small values of $|J|$ correspond to models at high temperature, and large values of $|J|$ to models at low temperature.

The Helmholtz free energy is defined as

$$F_{\mathrm{H}} \triangleq -\ln Z, \qquad (3)$$

see [1].

For each adjacent pair (k, ℓ), let

$$\kappa_{k,\ell}(x_k, x_\ell) = e^{J_{k,\ell}([x_k=x_\ell]-[x_k\neq x_\ell])} \qquad (4)$$

and let $f : \mathcal{X}^N \to \mathbb{R}$ be

$$f(x) = \prod_{k,\ell \text{ adjacent}} \kappa_{k,\ell}(x_k, x_\ell) \qquad (5)$$

where the product runs over all the (unordered) adjacent pairs (k, ℓ).

We are interested in computing the partition function

$$Z = \sum_{x \in \mathcal{X}^N} e^{-E(x)} \qquad (6)$$

$$= \sum_{x \in \mathcal{X}^N} \prod_{k,\ell \text{ adjacent}} e^{J_{k,\ell}\big([x_k=x_\ell]-[x_k\neq x_\ell]\big)} \qquad (7)$$

$$= \sum_{x \in \mathcal{X}^N} \prod_{k,\ell \text{ adjacent}} \kappa_{k,\ell}(x_k, x_\ell) \qquad (8)$$

$$= \sum_{x \in \mathcal{X}^N} f(x). \qquad (9)$$

In 1D Ising models, f has a cycle-free factor graph representation and Z, as in (9), can be computed directly by
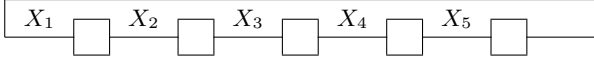
Fig. 1. Factor graph of a 1D Ising model with $N = 5$ and with periodic boundary conditions. The solid boxes represent factors as in (12), and edges represent the variables.
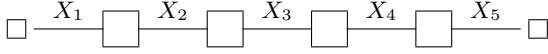

Fig. 2. Factor graph of a 1D Ising model with $N = 5$ and with free boundary conditions. The normal-size boxes represent factors as in (12), and the two small boxes represent constant factors.

sum-product message passing [5], [6], which (in this context) coincides with the transfer matrix method in statistical physics [7], [8, Chapter 5].

In 2D Ising models with constant coupling and in the absence of an external field, the exact value of Z in thermodynamic limits (for $N \to \infty$) was found by Onsager [9].

For finite-size 2D Ising models with arbitrary coupling, estimates of the partition function can be computed by Markov chain Monte Carlo methods [10]–[12]. At high temperatures, the Boltzmann distribution (2) tends to a uniform distribution and Monte Carlo methods generally work very well. At low temperatures, however, variables have long-range interactions; Monte Carlo methods are plagued by slow and erratic convergence, and may break down completely.

In this paper, we consider using the dual factor graph [13]–[15] to compute (or to estimate) the partition function of the Ising model. In the one-dimensional case, we will thus obtain an alternative derivation of the (well-known) analytical solution. In the two-dimensional case, we find that Monte Carlo methods are much more efficient (due to much faster mixing) on the dual graph than on the original graph, especially at low temperature.

The paper is structured as follows. In Section II, we recall the construction of the dual factor graph and the factor graph duality theorem. In Section III, we use this theorem to obtain the partition function of 1D Ising models. In Section IV, we discuss the dual factor graph of 2D Ising models, which we then use for the numerical simulations in Section V.

II. PARTITION FUNCTION VIA FACTOR GRAPH DUALITY

The factorization of a function, as in (5), can be represented by a Forney factor graph. The nodes in a Forney factor graph represent the factors and the edges (or half-edges, which are connected to only one node) represent the variables. The edge (or half-edge) that represents some variable x, is connected to the node representing the factor κ, if and only if x is an argument of κ, see [5], [13].

In the factor graphs that we study in this paper, all variables are binary and there are no half edges. Starting from such a factor graph, we can obtain its dual by replacing each variable x with its dual (frequency) variable $\tilde{x}$, each factor κ with its Fourier transform ν, and each equality constraint with an XOR factor [13], [14].
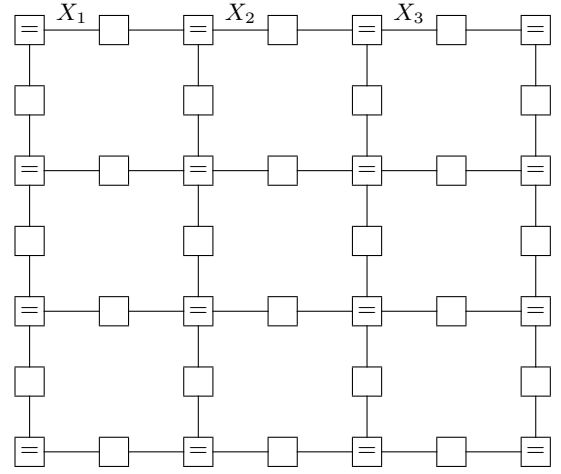

Fig. 3. Factor graph of a 2D Ising model. The unlabeled boxes represent factors as in (4).

For binary variables $\tilde{x}_1, \tilde{x}_2, \ldots, x_k$, the XOR factor is defined as

$$g(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k) \triangleq [\tilde{x}_1 \oplus \tilde{x}_2 \oplus \ldots \oplus \tilde{x}_k = 0] \qquad (10)$$

where $\oplus$ denotes addition modulo 2.

Note that, in general, factors in the dual Forney graph can be negative or even complex-valued [16], [17].

Random variables in the dual domain are denoted by $\tilde{X}$, which also take their values in $\mathcal{X}$. In the dual Forney factor graph, we denote the partition function by Z_d, and the number of edges by E. In such a set-up, according to the factor graph duality theorem [14, Theorem 2],

$$Z_d = |\mathcal{X}^E| Z \qquad (11)$$

Therefore, one alternative method to compute Z, is to first compute the partition function of the dual factor graph Z_d, and then apply (11), see [14], [15]. This procedure is particularly useful, in the cases that computing the partition function can be performed more efficiently in the dual domain. In this paper, we show that this is indeed the case for the Ising model, especially at low temperature.

III. 1D ISING MODELS

We consider a 1D Ising model with size N, with N binary (i.e., $\{0,1\}$-valued) variables, and with *periodic* boundary conditions (i.e., $X_{N+1} = X_1$).

For $1 \leq \ell \leq N$, we have

$$\kappa_\ell(x_\ell, x_{\ell+1}) = \begin{cases} e^{J_\ell}, & \text{if } x_\ell = x_{\ell+1} \\ e^{-J_\ell}, & \text{if } x_\ell \neq x_{\ell+1} \end{cases} \qquad (12)$$

Therefore,

$$Z = \sum_{x \in \mathcal{X}^N} \prod_{\ell=1}^{N} \kappa_\ell(x_\ell, x_{\ell+1}) \qquad (13)$$

The corresponding factor graph of f is shown in Fig. 1. Note that, in order to create periodic boundary conditions, we have simply joined the two ends of the factor graph.

To construct the dual factor graph, each factor (12) is replaced by its 2D discrete Fourier transform (DFT), where the 2D DFT $\nu(\tilde{x}_1, \tilde{x}_2)$ of $\kappa(x_1, x_2)$, is defined as

$$\nu(\tilde{x}_1, \tilde{x}_2) \triangleq \sum_{x_1 \in \mathcal{X}} \sum_{x_2 \in \mathcal{X}} \kappa(x_1, x_2) e^{-i2\pi(x_1 \tilde{x}_1 + x_2 \tilde{x}_2)/|\mathcal{X}|} \quad (14)$$

where $i = \sqrt{-1}$.

Therefore in the dual factor graph, each factor $\nu_\ell(\tilde{x}_\ell, \tilde{x}_{\ell+1})$ has the following form

$$\nu_\ell(\tilde{x}_\ell, \tilde{x}_{\ell+1}) = \begin{cases} 4\cosh J_\ell, & \text{if } \tilde{x}_\ell = \tilde{x}_{\ell+1} = 0 \\ 4\sinh J_\ell, & \text{if } \tilde{x}_\ell = \tilde{x}_{\ell+1} = 1 \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

Computing Z_d is now straightforward since all the factors in (15) are diagonal. We conclude that

$$Z_d = \sum_{\tilde{x} \in \mathcal{X}^N} \prod_{\ell=1}^{N} \nu_\ell(\tilde{x}_\ell, \tilde{x}_{\ell+1}) \quad (16)$$

$$= 4^N \Big(\prod_{\ell=1}^{N} \cosh J_\ell + \prod_{\ell=1}^{N} \sinh J_\ell \Big). \quad (17)$$

The number of edges in the dual factor graph is N. Using the factor graph duality theorem (11), we obtain

$$Z = 2^N \Big(\prod_{\ell=1}^{N} \cosh J_\ell + \prod_{\ell=1}^{N} \sinh J_\ell \Big). \quad (18)$$

Finally, we state (without proof) that computing Z of a 1D Ising model with size N and with *free* boundary conditions is also straightforward. After introducing two constant "dummy" factors at the two ends of the factor graph, as illustrated in Fig. 2, we can directly apply the factor graph duality theorem to compute the partition function as

$$Z = 2 \prod_{\ell=1}^{N-1} (2\cosh J_\ell) \quad (19)$$

see [7, Chapter 2], [8, Chapter 5].

IV. FINITE-SIZE 2D ISING MODELS

We consider a 2D Ising model with size $N = m \times m$, with binary variables (i.e., $\mathcal{X} = \{0, 1\}$), and with factors as in (4). The corresponding Forney factor graph with factors as in (4) is shown in Fig. 3, where the boxes labeled "=" are equality constraints [5].

In the dual Forney factor graph, the equality constraints are replaced by XOR factors (10), and each factor (4) by its 2D DFT, which has the following form

$$\nu_{k,\ell}(\tilde{x}_k, \tilde{x}_\ell) = \begin{cases} 4\cosh J_{k,\ell}, & \text{if } \tilde{x}_k = \tilde{x}_\ell = 0 \\ 4\sinh J_{k,\ell}, & \text{if } \tilde{x}_k = \tilde{x}_\ell = 1 \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

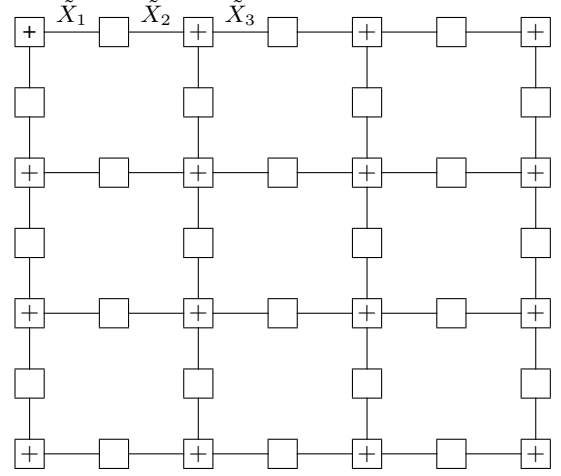The corresponding dual Forney factor graph with factors as in (20) is shown in Fig. 4.

Fig. 4.   The dual Forney factor graph where unlabeled boxes represent factors as in (20) and boxes containing $+$ symbols represent XOR factors as in (10).
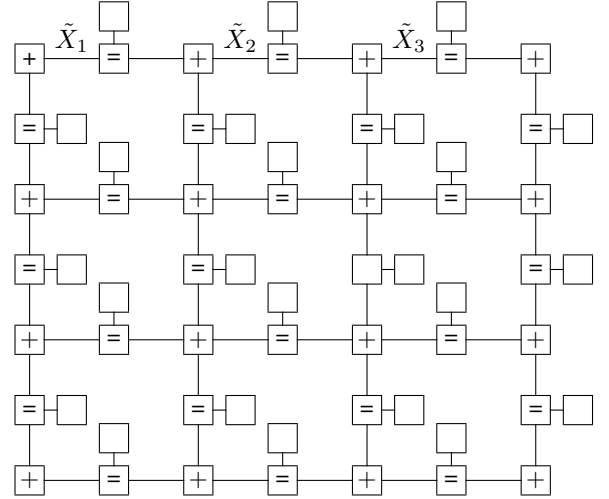

Fig. 5.   The modified dual Forney factor graph where unlabeled boxes represent factors as in (21) and boxes containing $+$ symbols represent XOR factors as in (10).

Since all the factors in (20) are diagonal, it is possible to simplify the dual factor graph in Fig. 4, to construct the modified dual factor graph depicted in Fig. 5, with factors as

$$\nu_k(\tilde{x}_k) = \begin{cases} 4\cosh J_k, & \text{if } \tilde{x}_k = 0 \\ 4\sinh J_k, & \text{if } \tilde{x}_k = 1 \end{cases} \quad (21)$$

The corresponding modified dual Forney factor graph with factors as in (21) is shown in Fig. 5.

We are interested in computing the partition function, as in (8). In our numerical experiments in Section V, we will consider the problem of computing an estimate of the partition function (or equivalently the free energy (3)) of 2D Ising models (with constant or with spatially varying couplings) by Monte Carlo methods as in [12], on the original factor graph with factors as in (4), and on the modified dual factor graph with factors as in (21). Monte Carlo methods on the dual graph may be viewed as simulating the cycles rather than individual variables.
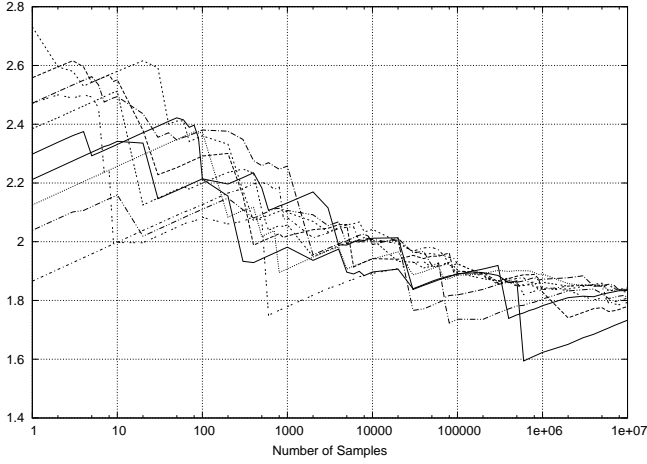
Fig. 6. Estimated free energy per site vs. the number of samples for a 5×5 ferromagnetic Ising model with $J = 0.75$ (low temperature) using Gibbs sampling on the original factor graph. The plot shows 10 different sample paths.
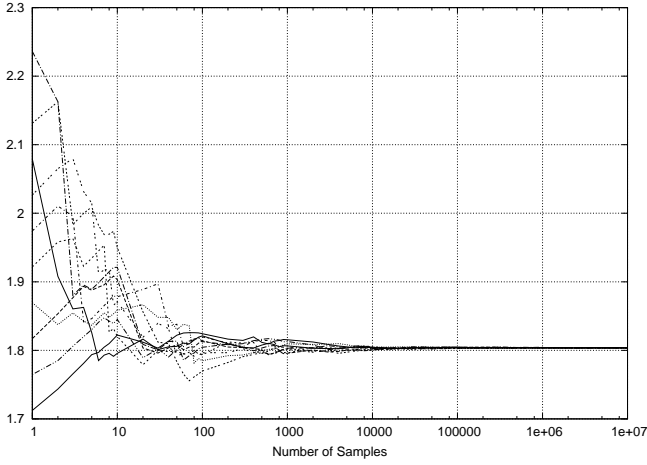

Fig. 8. Estimated free energy per site vs. the number of samples for a 5×5 ferromagnetic Ising model with $J = 1.25$ (low temperature) using uniform sampling on the original factor graph. The plot shows 10 different sample paths.

Fig. 7. Everything as in Fig. 6, but on the (modified) dual factor graph.

Fig. 9. Everything as in Fig. 8, but on the (modified) dual factor graph.

V. NUMERICAL EXPERIMENTS

We apply Monte Carlo methods to compute the free energy (3) per site, i.e., $\frac{1}{N} \log_2 Z$, of 2D Ising models with size $N = m \times m$. Since the value of Z is invariant under the change of sign of J, we will only consider ferromagnetic Ising models.

In Section V-A, we consider 2D ferromagnetic Ising models, in which the coupling parameter J is a positive constant (cf. Section I). For different values of J, we compare the convergence of Gibbs sampling using the Ogata-Tanemura method [12], [18] and uniform sampling [11], [17] in the original factor graph, as in Fig. 3, and in the modified dual factor graph, as in Fig. 5. In Section V-B, we apply uniform sampling on the modified dual factor graph to compute the free energy per site of 2D ferromagnetic Ising models with spatially varying couplings.

Note that in the modified dual factor graph, we can partition the set of random variables $\tilde{X}$, into $\tilde{X}_A$ and $\tilde{X}_B$, with the property that the random variab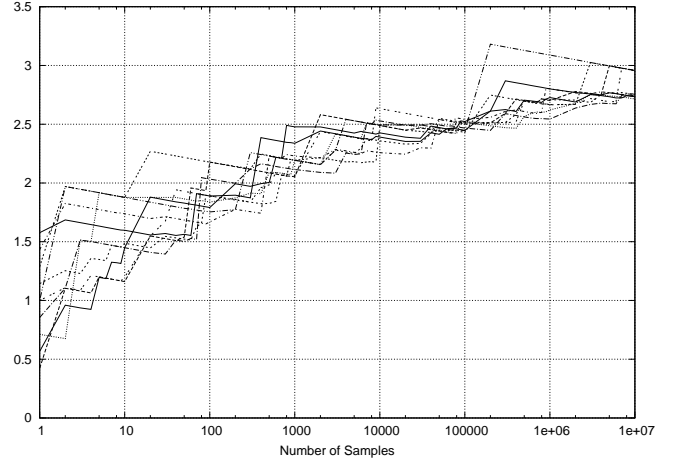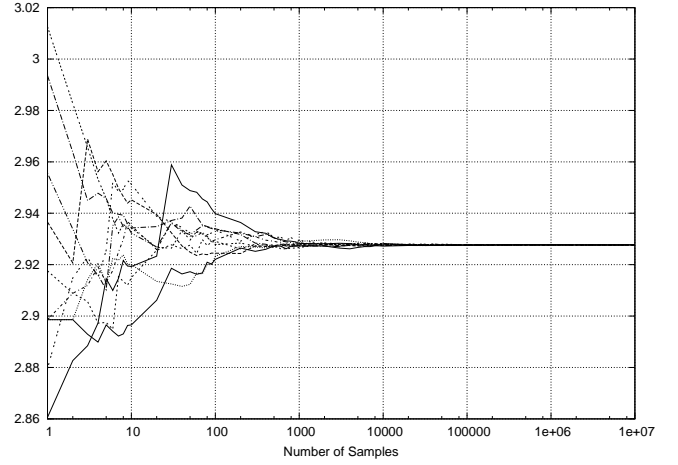les in $\tilde{X}_B$ are linear combinations (involving the XOR factors) of the random variables in $\tilde{X}_A$. Therefore, Monte Carlo methods can be applied directly on $\tilde{X}_A$, then $\tilde{X}_B$ can be updated at each iteration according to the new values of $\tilde{X}_A$. In this case, the size of the state space for Monte Carlo methods only depends on the size of $\tilde{X}_A$.

A. 2D Ising models with constant coupling

We estimate the free energy per site, i.e., $\frac{1}{N} \log_2 Z$, of 2D ferromagnetic Ising models with size $N = 5 \times 5$ at relatively low and at very low temperatures. For $J = 0.75$, Figs. 6 and 7 show simulation results obtained from Gibbs sampling on the original factor graph and on the modified dual factor graph, respectively. From Fig. 7, the estimated $\frac{1}{N} \log_2 Z$ is about 1.802.

Figs. 8 and 9 show simulation results for $J = 1.25$, obtained from uniform sampling on the original factor graph and on the modified dual factor graph, respectively. From Fig. 9, the estimated $\frac{1}{N} \log_2 Z$ is about 2.928.
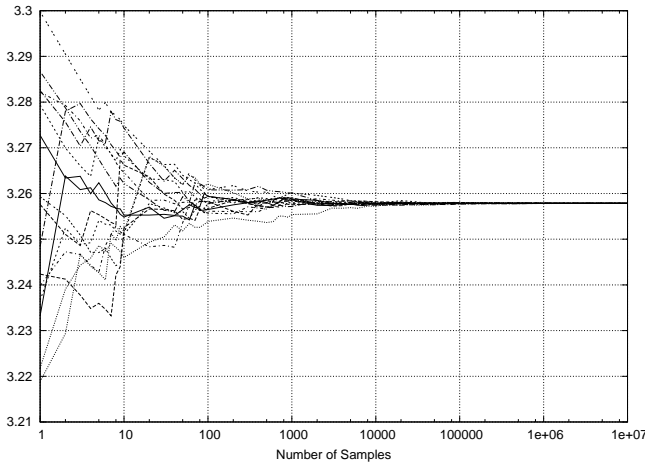
Fig. 10. Estimated free energy per site vs. the number of samples for a 10×10 ferromagnetic Ising model with $J \sim \mathcal{U}[1.0, 1.5]$ using uniform sampling on the modified dual factor graph. The plot shows 15 different sample paths.

Fig. 11. Everything as in Fig. 10, but with $N = 20 \times 20$.

Note that, at low temperature, we observe much faster mixing with Monte Carlo methods on the modified dual factor graph. On the dual factor graph, convergence improves as J increases (i.e., temperature decreases), which is in sharp contrast to convergence on the original factor graph.

B. 2D Ising model with spatially varying couplings

We apply uniform sampling on the modified dual factor graph to estimate $\frac{1}{N} \log_2 Z$ of 2D ferromagnetic Ising models with spatially varying couplings. In our experiments, we set $J_{k,\ell} \sim \mathcal{U}[1.0, 1.5]$ independently for each factor.

For $N = 10 \times 10$, Fig. 10 shows simulation results for one instance of the Ising model, where the estimated $\frac{1}{N} \log_2 Z$ is about 3.258. For one instance of such an Ising model with size $N = 20 \times 20$, simulation results for $\frac{1}{N} \log_2 Z$ are shown in Fig. 11. The estimated free energy per site is about 3.443.

As in our numerical experiments in Section V-A, we observe fast mixing using uniform sampling on the modified dual factor graph. Convergence of Monte Carlo methods improves as J increases (or equivalently as the temperature T decreases).

VI. Conclusion

The dual factor graph theorem [13]–[15] offers new possibilities for computing the partition function of Ising models. In particular, Monte Carlo methods work much better on the dual graph, especially at low temperature. A comparison with the Swensen-Wang algorithm [19] needs to be addressed in future work. Also, the relation between factor graph duality and the Kramers-Wannier duality [20] should be investigated.

References

[1] K. Huang, *Statistical Mechanics*. John Wiley & Sons, 1987.

[2] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 1989.
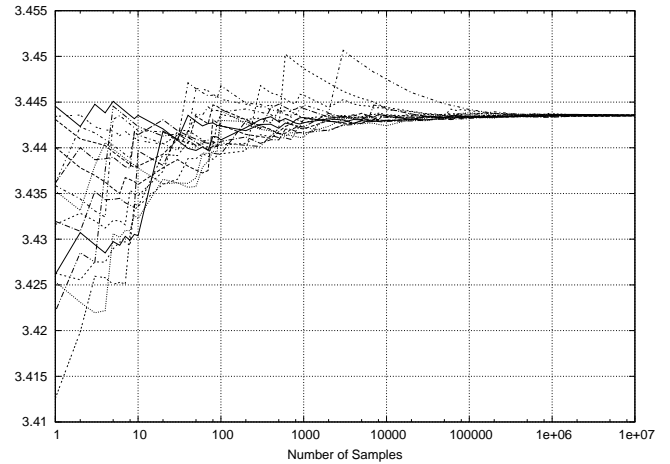
[3] B. A. Cipra, "An introduction to the Ising model," *Amer. Mathematical Monthly,* vol. 94, pp. 937–959, Dec. 1987.

[4] M. Mézard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, 2009.

[5] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Proc. Mag.,* vol. 29, pp. 28–41, Jan. 2004.

[6] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory,* vol. 47, pp. 498–519, Feb. 2001.

[7] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*. Dover Publications, 2007.

[8] J. M. Yeomans, *Statistical Mechanics of Phase Transitions*. Oxford University Press, 1992.

[9] L. Onsager, "Crystal statistics. I. A two-dimensional model with an order-disorder transition," *Phys. Rev.,* vol. 65, pp. 117–149, Feb. 1944.

[10] R. M. Neal, *Probabilistic Inference Using Markov Chain Monte Carlo Methods*. Techn. Report CRG-TR-93-1, Dept. Comp. Science, Univ. of Toronto, Sept. 1993.

[11] D. J. C. MacKay, "Introduction to Monte Carlo methods," in *Learning in Graphical Models,* M. I. Jordan, ed., Kluwer Academic Press, 1998, pp. 175–204.

[12] M. Molkaraie and H.-A. Loeliger, "Monte Carlo algorithms for the partition function and information rates of two-dimensional channels," *IEEE Trans. Inf. Theory,* vol. 59, pp. 495–503, Jan. 2013.

[13] G. D. Forney, Jr., "Codes on graphs: normal realization," *IEEE Trans. Inf. Theory,* vol. 47, pp. 520–548, Feb. 2001.

[14] A. Al-Bashabsheh and Y. Mao, "Normal factor graphs and holographic transformations," *IEEE Trans. Inf. Theory*, vol. 57, pp. 752–763, Feb. 2011.

[15] G. D. Forney, Jr. and P. O. Vontobel, "Partition functions of normal factor graphs," *2011 Information Theory and Applications Workshop*, La Jolla, USA, Feb. 6–11, 2011.

[16] R. N. Bracewell, *The Fourier Transform and its Applications*. McGraw-Hill, 1999.

[17] M. Molkaraie and H.-A. Loeliger, "Extending Monte Carlo methods to factor graphs with negative and complex factors," *Proc. 2012 IEEE Information Theory Workshop,* Lausanne, Switzerland, Sept. 3–7, 2012, pp. 367–371.

[18] Y. Ogata and M. Tanemura, "Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure," *Ann. Inst. Statist. Math.,* vol. 33, pp. 315–338, 1981.

[19] R. H. Swendsen and J. S. Wang, "Nonuniversal critical dynamics in Monte Carlo simulations," *Phys. Rev.,* vol. 58, pp. 86–88, Jan. 1987.

[20] H. A. Kramers and G. H. Wannier, "Statistics of the two-dimensional ferromagnet. Part I," *Phys. Rev.,* vol. 60, pp. 252–262, Aug. 1941.