A very simple algorithm for estimating the number of k-colourings of a low-degree graph.

Mark Jerrum*
Department of Computer Science
University of Edinburgh
United Kingdom

April 22, 1994

Abstract

A fully polynomial randomised approximation scheme is presented for estimating the number of (vertex) k-colourings of a graph of maximum degree Δ, when $k \geq 2\Delta + 1$.

Keywords. Analysis of algorithms, Antiferromagnetic Potts model, Randomised approximation scheme, Rapidly mixing Markov chains, Vertex colouring of graphs.

*The author is a Nuffield Foundation Science Research Fellow, and is supported in part by grant GR/F 90363 of the UK Science and Engineering Research Council, and by Esprit Working Group No. 7097, "RAND." The work described here was partly done while the author was a Visiting Research Fellow at Merton College, Oxford. Address for correspondence: Department of Computer Science, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JZ, United Kingdom; e-mail: `mrj@dcs.ed.ac.uk`.

1 A very simple sampling procedure

Let G be an undirected graph of maximum degree $\Delta = \Delta(G)$ on vertex set $V = \{0, \ldots, n-1\}$, and $C = \{0, \ldots, k-1\}$ be a set of k "colours." Let $X_0 : V \to C$ be a proper colouring of the vertices of G, i.e., one in which every edge has endpoints of different colours. (In this note, vertex colourings will generally be proper, so we often drop the adjective "proper" in what follows.) Such a colouring always exists if $k \geq \Delta + 1$, as can be appreciated by considering a simple sequential colouring algorithm. Indeed Brooks' theorem asserts that a colouring exists when $k \geq \Delta$, provided $\Delta \geq 3$ and G does not contain $K_{\Delta+1}$ as a connected component [2, 3].

Consider the Markov chain (X_t) whose state space $\Omega = \Omega_k(G)$ is the set of all k-colourings of G, and whose transition probabilities from state (colouring) X_t are modelled by the following procedure:

(1) choose a vertex $v \in V$ and a colour $c \in C$ uniformly at random (u.a.r.);

(2) recolour vertex v with colour c; if the resulting colouring X' is proper then let $X_{t+1} = X'$, otherwise let $X_{t+1} = X_t$.

This procedure describes what would be termed, by the statistical physics community, the "Glauber dynamics" of an antiferromagnetic Potts model at zero temperature. The Markov chain (X_t) — which we refer to in the sequel as $M(G, k)$ — is ergodic provided $k \geq \Delta + 2$, in which case the stationary distribution is uniform over Ω. (Precise definitions of various technical terms used in this section will be provided in Section 2.)

We show that $M(G, k)$ is "rapidly mixing," i.e., converges to a close approximation of the stationary distribution in time polynomial in n, provided $k \geq 2\Delta + 1$. This result provides us with a simple and efficient sampling procedure for k-colourings: simulate the Markov chain $M(G, k)$, starting at an arbitrary state, for a sufficiently large (but polynomial) number of steps, and return the current state as result. As a corollary we obtain a so-called fully polynomial randomised approximation scheme (fpras) for the number of k-colourings of a graph in the case $k \geq 2\Delta + 1$.

2 Sampling and approximate counting

For $t \in \mathbb{N}$, let $P^t : \Omega^2 \to [0, 1]$ denote the t-step transition probabilities[1] of the Markov chain $M(G, k)$ defined in Section 1, so that $P^t(x, y) = \Pr(X_t = y \mid X_0 = x)$ for all $x, y \in \Omega$. It is easily verified that $M(G, k)$ is (a) *irreducible*, i.e., for all $x, y \in \Omega$, there is a t such that $P^t(x, y) > 0$, and (b) *aperiodic*, i.e., $\gcd\{t : P^t(x, y) > 0\} = 1$ for all $x, y \in \Omega$. Irreducibility of $M(G, k)$ follows from the observation that any colouring x may be transformed to any other colouring y by sequentially assigning new colours to the vertices V in ascending sequence; before

[1] We drop the superscript t in the case $t = 1$.

assigning a new colour c to vertex v it is necessary to recolour all vertices $u > v$ that have colour c, but there is always at least one "free" colour to allow this to be done, provided $k \geq \Delta + 2$. Aperiodicity follows from the fact that the loop probabilities $P(x, x)$ are non-zero for all $x \in \Omega$.

A finite Markov chain that is irreducible and aperiodic is *ergodic*; i.e., there is a *stationary distribution* $\pi : \Omega \to [0, 1]$ such that $\lim_{t\to\infty} P^t(x, y) = \pi(y)$ for all $x, y \in \Omega$. Computation of the stationary distribution is facilitated by the following observation: if $\pi' : \Omega \to [0, 1]$ is any function satisfying "detailed balance"

$$\pi'(x)P(x, y) = \pi'(y)P(y, x), \quad \text{for all } x, y \in \Omega,$$

and the normalisation condition $\sum_{x\in\Omega} \pi'(x) = 1$, then π' is indeed the stationary distribution. Using this observation, it is easy to verify that the stationary distribution of $M(G, k)$ is uniform.

The efficiency of our approach to sampling k-colourings depends crucially on rate of convergence of $M(G, k)$ to stationarity. There are a number of ways of quantifying "closeness" to stationarity, but they are all essentially equivalent in this application. The *variation distance* at time t with respect to the initial state x is defined to be

$$\delta_x(t) = \max_{S\subseteq\Omega} |P^t(x, S) - \pi(S)| = \frac{1}{2} \sum_{y\in\Omega} |P^t(x, y) - \pi(y)|,$$

where $P^t(x, S) = \sum_{y\in S} P^t(x, y)$, and $\pi(S) = \sum_{x\in S} \pi(x)$. Note that the variation distance provides a uniform bound, over all events $S \subseteq \Omega$, of the difference in probabilities of occurrence of event S under the stationary and t-step distributions. The rate of convergence to stationarity from initial state x may be measured by the function

$$\tau_x(\varepsilon) = \min\{t : \delta_x(t') \leq \varepsilon \text{ for all } t' \geq t\}.$$

Finally, we need to formalise the notion of efficient approximation algorithm. A *randomised approximation scheme* for k-colourings in a graph G is a probabilistic algorithm that takes as input a graph G and an error bound $\varepsilon > 0$, and produces as output a number Y (a random variable) such that[2]

$$\Pr\Big((1 - \varepsilon)\, |\Omega_k(G)| \leq Y \leq (1 + \varepsilon)\, |\Omega_k(G)|\Big) \geq \frac{3}{4}.$$

A randomised approximation scheme is said to be *fully polynomial* [8] if it runs in time polynomial in n (the input length) and ε^{-1}. We shall abbreviate the rather unwieldy phrase "fully polynomial randomised approximation scheme" to *fpras*.

[2] There is no significance in the constant $\frac{3}{4}$ appearing in the definition, beyond its lying strictly between $\frac{1}{2}$ and 1. Any success probability greater than $\frac{1}{2}$ may be boosted to a value arbitrarily close to 1 by making a small number of trials and taking the median of the results [7].

3 An fpras for k-colourings

Our aim is to construct an fpras for the number of k-colourings of a low-degree graph. The key tool is the following result, to the effect that the Markov chain $M(G,k)$ is rapidly mixing.

Lemma 1 *Let G be a graph of maximum degree Δ on n vertices. Assuming $k \geq 2\Delta+1$, the convergence time $\tau(\varepsilon)$ of the Markov chain $M(G,k)$ is bounded above by*

$$\tau_x(\varepsilon) \leq \frac{k}{k-2\Delta}\, n \ln\left(\frac{n}{\varepsilon}\right),$$

regardless of the initial state x.

We defer the proof of Lemma 1 to Section 4, and press on to investigate its consequences. An immediate observation is that we have an polynomial-time almost uniform sampler[3] for k-colourings in a graph, provided $k \geq 2\Delta+1$. There is a close connection between almost uniform sampling and approximate counting, which has been discussed at some length by Jerrum, Valiant, and Vazirani [7]. In the light of this connection, it is not surprising that Lemma 1 leads fairly directly to the main result.

Theorem 2 *There is a fully polynomial randomised approximation scheme for the number of k-colourings in a graph G of maximum degree Δ, under the assumption $k \geq 2\Delta+1$. The time complexity of the approximation scheme is bounded above by*

$$\frac{50k}{k-2\Delta} \times \frac{nm^2}{\varepsilon^2} \ln\left(\frac{4nm}{\varepsilon}\right),$$

where n and m are the numbers of vertices and edges in G, and the time unit is a single simulation step of the Markov chain $M(G,k)$.

Proof The techniques we employ are standard in the area [6]. Recall that $\Omega_k(G)$ is the set of all k-colourings of G. Let m denote the number of edges in G, and let $G = G_m > G_{m-1} > \cdots > G_1 > G_0 = (V, \varnothing)$ be any sequence of graphs in which each graph G_{i-1} is obtained from the previous graph G_i by removing a single edge. We may express the quantity we wish to estimate as a product of ratios:

$$|\Omega_k(G)| = \frac{|\Omega_k(G_m)|}{|\Omega_k(G_{m-1})|} \times \frac{|\Omega_k(G_{m-1})|}{|\Omega_k(G_{m-2})|} \times \cdots \times \frac{|\Omega_k(G_1)|}{|\Omega_k(G_0)|} \times |\Omega_k(G_0)|, \qquad (1)$$

where, it will be observed, $|\Omega_k(G_0)| = k^n$. Our strategy is to estimate the ratio

$$\varrho_i = \frac{|\Omega_k(G_i)|}{|\Omega_k(G_{i-1})|}$$

[3] A precise definition of this phrase is not essential for what follows, and the reader is directed to [7], where the concept goes under the title *almost uniform generator.*

for each i in the range $1 \le i \le m$, and by substituting these quantities into identity (1), obtain an estimate for $|\Omega_k(G)|$.

Suppose that the graphs G_i and G_{i-1} differ in the edge $\{u, v\}$, which is present in G_i but absent from G_{i-1}. Clearly, $\Omega_k(G_i) \subseteq \Omega_k(G_{i-1})$. Any colouring in $\Omega_k(G_{i-1}) \setminus \Omega_k(G_i)$ assigns the same colour to u and v, and may be perturbed to a colouring in $\Omega_k(G_i)$ by recolouring vertex u with one of at least $k - \Delta \ge \Delta + 1$ colours. On the other hand, each colouring in $\Omega_k(G_i)$ can be obtained in at most one way as the result of such a perturbation, and hence

$$\frac{\Delta + 1}{\Delta + 2} \le \varrho_i \le 1. \tag{2}$$

To avoid trivialities, assume $0 < \varepsilon \le 1$, $n \ge 3$, and $\Delta \ge 2$. Let $Z_i \in \{0, 1\}$ denote the random variable obtained by simulating Markov chain $M(G_{i-1}, k)$ from a certain fixed initial state for

$$T = \left\lceil \frac{k}{k - 2\Delta}\, n \ln\left(\frac{4nm}{\varepsilon}\right) \right\rceil$$

steps and returning 1 if the final state is a member of $\Omega_k(G_i)$, and 0 otherwise. Let $\mu_i = \operatorname{Exp} Z_i$ be the expectation of Z_i. By Lemma 1,

$$\varrho_i - \frac{\varepsilon}{4m} \le \mu_i \le \varrho_i + \frac{\varepsilon}{4m}, \tag{3}$$

or, noting inequality (2),

$$\left(1 - \frac{\varepsilon}{3m}\right) \varrho_i \le \mu_i \le \left(1 + \frac{\varepsilon}{3m}\right) \varrho_i, \tag{4}$$

so the mean of a sufficiently large number of independent copies of Z_i will provide a good estimate for ϱ_i. Note that, by inequalities (2) and (3), $\mu_i \ge \frac{1}{2}$.

So let $Z_i^{(1)}, \ldots, Z_i^{(s)}$ be a sequence of $s = \lceil 37\varepsilon^{-2} m \rceil$ independent copies of the random variable Z_i obtained by simulating the Markov chain $M(G_{i-1}, k)$ from the fixed initial state s times, and let $\overline{Z}_i = s^{-1} \sum_{j=1}^{s} Z_i^{(j)}$ be their mean. Since Z_i is a random variable taking values from $\{0, 1\}$, it follows easily that $\mu_i^{-2} \operatorname{Var} Z_i = \mu_i^{-1} - 1 \le 1$, and hence $\mu_i^{-2} \operatorname{Var} \overline{Z}_i \le s^{-1}$. As our estimator for $|\Omega_k(G)|$, we use the random variable $Y = k^n\, \overline{Z}_1 \overline{Z}_2 \ldots \overline{Z}_m$. Note that $\operatorname{Exp} Y = k^n \mu_1 \mu_2 \ldots \mu_m$.

The performance of this estimator is characterised by its variance, which is bounded as follows:

$$\begin{aligned}
\frac{\operatorname{Var} \overline{Z}_1 \overline{Z}_2 \ldots \overline{Z}_m}{(\mu_1 \mu_2 \ldots \mu_m)^2} &= \prod_{i=1}^{m} \left(1 + \frac{\operatorname{Var} \overline{Z}_i}{\mu_i^2}\right) - 1 \\
&\le \left(1 + \frac{1}{s}\right)^m - 1 \\
&\le \exp\left(\frac{\varepsilon^2}{37}\right) - 1 \\
&\le \frac{\varepsilon^2}{36},
\end{aligned}$$

since $\mathrm{e}^{x/37} \le 1 + x/36$ provided $0 \le x \le 1$. Thus, by Chebychev's inequality,

$$\left(1 - \frac{\varepsilon}{3}\right)\mu_1\mu_2\ldots\mu_m \le k^{-n}Y \le \left(1 + \frac{\varepsilon}{3}\right)\mu_1\mu_2\ldots\mu_m$$

with probability at least $\frac{3}{4}$. But from inequality (4), we have

$$\left(1 - \frac{\varepsilon}{2}\right)\varrho_1\varrho_2\ldots\varrho_m \le \mu_1\mu_2\ldots\mu_m \le \left(1 + \frac{\varepsilon}{2}\right)\varrho_1\varrho_2\ldots\varrho_m,$$

which, combined with the previous inequality, implies that the estimator Y satisfies the requirements of an fpras for the number of colourings $|\Omega_k(G)|$.

The computation of the estimator involves ms Markov chain simulations, and each simulation is for T steps, a total of msT steps in total. The constant factor 50 appearing in the statement of the theorem is chosen large enough to absorb the various ceiling functions. $\square$

4 Rapid mixing

This section is devoted to a proof that the Markov chain $M(G,k)$ defined in Section 1 is rapidly mixing.

Proof of Lemma 1 Our strategy is to construct a coupling for $M = M(G,k)$: that is to say, a stochastic process (X_t, Y_t) on $\Omega \times \Omega$ such that each of the processes (X_t) and (Y_t), considered in isolation, is a faithful copy of M. We shall arrange a joint probability space for (X_t) and (Y_t) so that, far from being independent, the two processes tend to *couple*, so that $X_t = Y_t$ for all sufficiently large t. If it can be arranged that coupling occurs rapidly — independently of the initial states X_0 and Y_0 — we may deduce that M is rapidly mixing. The key result we use here is that the variation distance of the distribution of (X_t) from the stationary distribution is bounded above by the probability that (X_t) and (Y_t) have not coupled by time t; see, for example, Aldous [1, Lemma 3.6], or Diaconis [4, Chap. 4, Lemma 5].

The transition $(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$ in the coupling is defined by the following experiment:

(1) select a vertex $v \in V$, u.a.r.;

(2) compute a permutation $g = g(G, X_t, Y_t)$ of C according to a procedure to be explained presently;

(3) choose a colour $c \in C$, u.a.r.;

(4) in the colouring X_t (respectively Y_t), recolour vertex v with colour c (respectively $g(c)$) to obtain a new colouring X' (respectively Y');

(5) if X' (respectively Y') is a proper colouring then let $X_{t+1} = X'$ (respectively $Y_{t+1} = Y'$), otherwise let $X_{t+1} = X_t$ (respectively $Y_{t+1} = Y_t$).

Whatever procedure is used to select the permutation g in step (2), the distribution of $g(c)$ is uniform; thus (X_t) and (Y_t) are both faithful copies of M.

Let $A = A_t \subseteq V$ be the set of vertices on which the colourings X_t and Y_t agree, and $D = D_t \subseteq V$ be the set on which they disagree. Let $d'(v)$ denote the number of edges incident at vertex v that have one endpoint in A and one in D. Observe that

$$\sum_{v \in A} d'(v) = \sum_{v \in D} d'(v) = m', \tag{5}$$

where m' is the number of edges of G that span A and D. The procedure for computing $g = g(G, X_t, Y_t)$ is as follows.

(a) If $v \in D$ then g is the identity.

(b) If $v \in A$ then proceed as follows. Denote by N be the set of neighbours of v in G. Define $C_X \subseteq C$ to be the set of all colours c such that some vertex in N receives c in colouring X_t, but no vertex in N receives c in colouring Y_t. Let C_Y be defined analogously, with the roles of X_t and Y_t interchanged. Observe that $C_X \cap C_Y = \varnothing$ and $|C_X|, |C_Y| \leq d'(v)$. Suppose without loss of generality that $|C_X| \leq |C_Y|$. Choose any subset $C'_Y \subseteq C_Y$ with $|C'_Y| = |C_X|$, and let $C_X = \{c_1, \ldots, c_r\}$ and $C'_Y = \{c'_1, \ldots, c'_r\}$ be arbitrary enumerations of the sets C_X and C'_Y. Finally let g be the permutation $(c_1, c'_1) \cdots (c_r, c'_r)$, which interchanges the colour-sets C_X and C'_Y and leaves all other colours fixed.

It is clear that $|D_{t+1}| - |D_t| \in \{-1, 0, 1\}$. Consider first the probability that $|D_{t+1}| = |D_t| + 1$. For this event to occur, the vertex v selected in line (1) must lie in A, and hence the permutation g is selected by procedure (b) above. If the new colourings X_{t+1} and Y_{t+1} are to disagree at vertex v, then the colour c selected in line (3) must be an element of C_Y. (If $c \in C_X$ then vertex v remains the same colour in both X_{t+1} and Y_{t+1}.) But we have observed that $|C_Y| \leq d'(v)$, and hence

$$\Pr\Big(|D_{t+1}| = |D_t| + 1\Big) \leq \frac{1}{n} \sum_{v \in A} \frac{d'(v)}{k} = \frac{m'}{kn}, \tag{6}$$

where the right-hand equality is by equation (5). Now consider the probability that $|D_{t+1}| = |D_t| - 1$. For this event to occur, the vertex v selected in line (1) must lie in D, and hence the permutation g selected in line (2) is the identity. For the new colourings X_{t+1} and Y_{t+1} to agree at vertex v, it is enough that the colour c selected in line (3) is different from all the colours that X_t and Y_t assign to neighbours of v. The number of colours c that satisfy this condition is at least $k - 2\Delta + d'(v)$, and hence

$$\Pr\Big(|D_{t+1}| = |D_t| - 1\Big) \geq \frac{1}{n} \sum_{v \in D} \frac{k - 2\Delta + d'(v)}{k} = \frac{(k - 2\Delta)}{kn} \times |D| + \frac{m'}{kn}. \tag{7}$$

Define

$$a = \frac{k - 2\Delta}{kn} \quad \text{and} \quad \beta = \beta(m') = \frac{m'}{kn},$$

so that $\Pr\left(|D_{t+1}| = |D_t|+1\right) \leq \beta$ and $\Pr\left(|D_{t+1}| = |D_t|-1\right) \geq a|D_t|+\beta$. Provided $a > 0$, i.e., $k > 2\Delta$, the size of the set D_t tends to decrease with t, and hence, intuitively at least, the event $D_t = \varnothing$ should occur with high probability for some $t \leq T$ with T not too large. Since $D_t = \varnothing$ is precisely the event that coupling has occurred, it only remains to confirm this intuition, and quantify the rate at which D_t converges to the empty set. From equations (6) and (7),

$$\begin{aligned}
\operatorname{Exp}|D_{t+1}| &\leq \beta(|D_t|+1) + (a|D_t|+\beta)(|D_t|-1) + (1 - a|D_t| - 2\beta)|D_t| \\
&= (1-a)|D_t|,
\end{aligned}$$

provided $D_t \neq \varnothing$ and $D_t \neq V$; it may easily be checked by separate arguments that the conclusion $\operatorname{Exp}|D_{t+1}| \leq (1-a)|D_t|$ holds also in these boundary situations. Thus $\operatorname{Exp}|D_t| \leq (1-a)^t|D_0| \leq n(1-a)^t$, and, because $|D_t|$ is an non-negative integer random variable, $\Pr(|D_t| \neq 0) \leq n(1-a)^t \leq n\mathrm{e}^{-at}$. Note that $\Pr(D_t \neq \varnothing) \leq \varepsilon$, provided $t \geq a^{-1}\ln(n\varepsilon^{-1})$, establishing the result. □

Alan Frieze has pointed out to me that the Markov chain $M(G,k)$ also mixes rapidly (i.e., in time polynomial in n) when $k = 2\Delta$, though the exact rate of convergence may be slower in this case.

5 Open questions

For $\Delta \geq 4$, it is NP-complete to determine whether a graph of maximum degree Δ is $(\Delta-1)$-colourable (see Garey, Johnson, and Stockmeyer [5]), and hence there can be no fpras for k-colourings when $k < \Delta$, unless RP = NP. Thus there is a range of k, namely $\Delta \leq k \leq 2\Delta$, in which the existence of an fpras is in question. The Markov chain $M(G,k)$ presented in Section 1 is ergodic provided $k \geq \Delta+2$, and there is no evidence against the view that this condition is also sufficient for $M(G,k)$ to be rapidly mixing. Thus the fpras presented here might conceivably be valid down to $k = \Delta+2$.

When $k = \Delta+1$, the Markov chain $M(G,k)$ is not ergodic, as Lubin and Sokal have observed [9], and designing an fpras for this case would present a distinct challenge. An fpras for the case $k = \Delta$ may be too much to hope for, as the corresponding existence problem is no longer trivial; however, the possibility still cannot be ruled out.

A natural extension to consider is to the k-state antiferromagnetic Potts model at non-zero temperature. Let $\widehat{\Omega} = \widehat{\Omega}_k(G)$ denote the set of all k-colourings of G, including those that are not proper; these are the *configurations* of a k-state Potts system with interaction graph G. For $x \in \widehat{\Omega}$, let the *Hamiltonian* $H(x)$ be the number of edges in G both endpoints of which receive the same colour. Thus the proper colourings x of G are precisely those for which $H(x)=0$. The key problem is to evaluate the *partition function* $Z=\sum_{x\in\widehat{\Omega}}\exp(-\beta H(x))$ of this system, where $\beta\geq 0$ is a parameter known as *inverse temperature*. The significance of Z is that it is the normalising factor in the *Gibbs distribution*, which assigns probability $Z^{-1}\exp(-\beta H(x))$ to each configuration x in the stationary distribution. Observe that the Potts model generalises colouring, in the sense that the number of proper k-colourings of G is the limit of Z as $\beta\to\infty$, i.e., as temperature tends to zero.

The Markov chain presented in Section 1 is easily generalised to allow sampling according to the Gibbs distribution at non-zero temperature. Transition probabilities from state X_t are modelled by the following procedure:

(1) choose a vertex $v\in V$ and a colour $c\in C$ u.a.r.;

(2) recolour vertex v with colour c to obtain a new colouring X', and let $p_{\mathrm{acc}}=\min\Big\{1,\exp\Big(-\beta(H(X')-H(X_t))\Big)\Big\}$;

(3) with probability p_{acc} let $X_{t+1}=X'$, and with probability $1-p_{\mathrm{acc}}$ let $X_{t+1}=X_t$.

The acceptance condition used here is the *Metropolis rule*, familiar in the computer simulation of models in statistical physics, and in combinatorial optimisation by simulated annealing. This generalised Markov chain is ergodic for all β.[4]

Intuitively, evaluation of the partition function Z ought to become easier as $\beta\to 0$, i.e., as temperature increases. However the coupling argument used in the proof of Theorem 2 breaks down; the obstacle to be faced is that the coupling may visit states (X_t,Y_t) such that the event $|D_{t+1}|=|D_t|+1$ occurs with high probability. These pairs (X_t,Y_t) involve non-proper colourings, and do not arise in the zero-temperature limit.

Acknowledgement

I thank Dominic Welsh for stimulating discussions, and for reminding me how little we know about the complexity of approximate counting.

References

[1] David Aldous, Random walks on finite groups and rapidly mixing Markov chains, *Séminaire de Probabilités XVII 1981/82* (A. Dold and B. Eckmann, eds), Springer Lecture Notes in Mathematics **986**, pp. 243–297.

[2] Béla Bollobás, *Extremal Graph Theory*, Academic Press, 1978.

[4] The question of ergodicity arises only in the limit as $\beta\to\infty$.

[3] R. L. Brooks, On colouring the nodes of a network, *Proceedings of the Cambridge Philosophical Society* **37** (1941), pp. 194–197.

[4] Persi Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics, Hayward CA, 1988.

[5] M. R. Garey, D. S. Johnson, and L. Stockmayer, Some simplified NP-omplete graph problems, *Theoretical Computer Science* **1** (1976), pp. 237–267.

[6] Mark Jerrum, The "Markov chain Monte Carlo" method: analytical techniques and applications. To appear in *Proceedings of the IMA Conference on Complex Stochastic Systems and Engineering* (Leeds, September 1993), Oxford University Press.

[7] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani, Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* **43** (1986), pp. 169–188.

[8] Richard M. Karp and Michael Luby, Monte-Carlo algorithms for enumeration and reliability problems, *Proceedings of the 24th IEEE Symposium on Foundations of Computer Science*, Computer Society Press, 1983, pp. 56–64.

[9] Mona Lubin and Alan D. Sokal, Comment on "Antiferromagnetic Potts Model," *Physical Review Letters* **71** (1993), p. 1778.