On approximately counting colourings of small degree graphs*

Russ Bubley,† Martin Dyer, Catherine Greenhill and Mark Jerrum

Abstract

We consider approximate counting of colourings of an n-vertex graph using rapidly mixing Markov chains. It has been shown by Jerrum and by Salas and Sokal that a simple random walk on graph colourings would mix rapidly provided the number of colours, k, exceeded the maximum degree Δ of the graph by a factor of at least 2. Lack of improvements on this bound led to the conjecture that $k \geq 2\Delta$ was a natural barrier. We disprove this conjecture in the simplest case of 5-colouring graphs of maximum degree 3. Our proof involves a computer-assisted proof technique to establish rapid mixing of a new "heat bath" Markov chain on colourings using the method of path coupling. We outline an extension to 7-colourings of triangle-free 4-regular graphs. Since rapid mixing implies approximate counting in polynomial time, we show in contrast that exact counting is unlikely to be possible (in polynomial time). We give a general proof that the problem of exactly counting the number of proper k-colourings of graphs with maximum degree Δ is $\#P$-complete whenever $k \geq 3$ and $\Delta \geq 3$.

1 Introduction

The problem of properly colouring the vertices of an n-vertex graph with some given number of colours, k, has been widely studied [19]. Of equal theoretical interest has been the problem of *counting the number* of proper k-colourings of a graph, the values of the so-called *chromatic polynomial* [19, p. 247]. Unfortunately this is a $\#P$-hard counting problem even for graphs with a fixed bound Δ on the vertex degrees. (See section 6 for a proof of this fact in the context of this paper.) Consequently, exact counting is unlikely to be possible, and attention turns to *approximate* counting, to some given proportional error ϵ. It is well known [15] that *randomized* approximate counting in this sense is equivalent to *approximate uniform generation* of a colouring, and it is this problem we address here, using the *Monte Carlo Markov chain* (MCMC) method [14]. The use of the MCMC method originated in statistical physics (see for example [5, 10, 17]), but rigorous analysis of the method has been a more recent development, with notable contributions from the computer science community [6, 13]. This analysis involves proving that a Markov chain converges quickly to its stationary distribution, using one of a small number of available methods.

In this context, Jerrum [12] and Salas and Sokal [18] independently proved that a simple random walk on the k-colourings of an n-vertex graph would mix rapidly,

*Research supported by the ESPRIT working group RAND2. An earlier version of this paper appeared in the 9th Annual ACM/SIAM Symposium on Discrete Algorithms, San Francisco, January 1998

†Supported in part by an EPSRC Studentship.

provided the number of colours k exceeded the maximum degree Δ of the graph by a factor of more than 2. These proofs were based on entirely different techniques, *coupling* and *Dobrushin uniqueness* respectively. The two results had different merits. Jerrum's result was subsequently extended (but with an $\Omega^*(n^2)$[1] increase in running time) to the case $k = 2\Delta$, whereas Salas and Sokal's was extended to the closely related *Potts model* [3]. However the similarity of these bounds suggested a natural conjecture among the research community in this area, that 2Δ might actually be a barrier for rapid mixing of the underlying Markov chain on colourings. (Note that the chain is known to converge eventually for $k \geq \Delta + 2$.) Although the extension of Jerrum's result naturally weakened the "2Δ conjecture" to $k \geq 2\Delta$, there is still no general refutation of this conjecture. Here we examine only specific cases.

In a recent paper, Bubley and Dyer [3], using a new technique called *path coupling*, showed how a coupling proof could be extended to the Potts model (and beyond). Subsequently, Dyer and Greenhill [9] used this technique to analyse a more rapidly mixing chain on colourings. This reduces the running time for the case $k = 2\Delta$ by an $\Omega^*(n^2)$ factor, but still does not beat the "2Δ barrier".

Here we show that the 2Δ bound cannot be uniformly true, by showing that it can be beaten in the simplest case of 5-colouring graphs of maximum degree 3. To establish our result we analyse a new Markov chain on colourings, from the so-called *heat bath* family, using the technique of path coupling. The proof idea is somewhat novel in this area: we establish the existence of certain required couplings by solving a large number of large linear programs (in fact *transportation problems*). The analysis therefore cannot be carried out "by hand" and requires the use of linear programming software to solve the many sub-problems. Thus our analysis is a "computer proof" of a mathematical theorem, in the spirit of [2, 4].

We also extend our results to 7-colourings of *triangle-free* 4-regular graphs, again beating the 2Δ barrier. (Observe that this includes the physically relevant case of planar grids.) However, our methods are subject to combinatorial explosion in terms of Δ, and we have currently not succeeded in pushing them further.

Using results presented in [7] it is possible to deduce that the simple Jerrum/Salas–Sokal Markov chain of colourings is rapidly mixing for the same values of k and the same families of graphs as our new chains: that is, when $k = 5$ and the graphs have maximum degree 3 or when $k = 7$ and the graphs are triangle-free and 4-regular. Hence 2Δ is no barrier even for this very simple and natural Markov chain. (Note however that the mixing rate which is established in this way, while still polynomial, is much larger than the mixing rate of our chains).

The plan of the paper is as follows. In section 1.1 of the paper we review the path coupling method, and in section 1.2 we briefly review recent work on approximately counting colourings. In section 2 we give a rigorous mathematical definition of a *heat bath Markov chain*. (As far as we are aware, this is the first general definition of this concept.) In section 3 we show that the analysis of a heat bath Markov chain can be reduced (using path coupling) to the task of solving a set of related transportation problems. This leads to a computational method for analysing heat bath Markov chains. In section 4 this approach is applied to a Markov chain on 5-colourings of graphs of maximum degree 3. A table needed for the computation is provided. In section 5, we outline the extension to 7-colourings of triangle-free 4-regular graphs. Finally, in

[1] $\Omega^*(\cdot)$ is the notation which hides factors of $\log n$.

section 6 we prove that the problem, referred to above, of exactly counting the number of proper k-colourings of graphs with maximum degree Δ is $\#P$-complete when $k \geq 3$ and $\Delta \geq 3$.

1.1 Path coupling.

Let Ω be a finite set and let $\mathcal{M}$ be a Markov chain with state space Ω, transition matrix P and unique stationary distribution π. If the initial state of the Markov chain is x then the distribution of the chain at time t is given by $P_x^t(y) = P^t(x, y)$. The *total variation distance* of the Markov chain from π at time t, with initial state x, is defined by

$$\mathrm{d}_{\mathrm{TV}}(P_x^t, \pi) = \frac{1}{2} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)|.$$

A Markov chain is only useful in terms of almost uniform sampling or approximate counting if its total variation distance tends to zero relatively quickly from some (easily obtainable) initial state. Let $\tau_x(\varepsilon)$ denote the least value T such that $\mathrm{d}_{\mathrm{TV}}(P_x^t, \pi) \leq \varepsilon$ for all $t \geq T$. The *mixing rate*[2] of $\mathcal{M}$, denoted by $\tau(\varepsilon)$, is defined by $\tau(\varepsilon) = \max\{\tau_x(\varepsilon) : x \in \Omega\}$. A Markov chain is said to be *rapidly mixing* if the mixing rate is bounded above by some polynomial in n and $\log(\varepsilon^{-1})$, where n is a measure of the size of the elements of Ω. All logarithms are to the base e.

There are relatively few methods available to prove that a Markov chain is rapidly mixing. One such method is *coupling*. A *coupling* for $\mathcal{M}$ is a stochastic process (X_t, Y_t) on $\Omega \times \Omega$ such that each of (X_t), (Y_t), considered independently, is a faithful copy of $\mathcal{M}$. The Coupling Lemma (see for example, Aldous [1]) states that the total variation distance of $\mathcal{M}$ at time t is bounded above by $\mathrm{Prob}[X_t \neq Y_t]$, the probability that the process has not coupled. The difficulty in applying this result lies in obtaining an upper bound for this probability. In the *path coupling* method, introduced by Bubley and Dyer [3], one must only define and analyse a coupling on a subset S of $\Omega \times \Omega$. Choosing the set S carefully can considerably simplify the arguments involved in proving rapid mixing of Markov chains by coupling. The path coupling method is described in the next lemma, taken from [9]. We use the term *path* to refer to a sequence of elements of Ω, which need not be a path of possible transitions in the Markov chain.

Lemma 1.1 *Let δ be an integer valued metric defined on $\Omega \times \Omega$ which takes values in $\{0, \ldots, D\}$. Let S be a subset of $\Omega \times \Omega$ such that for all $(X_t, Y_t) \in \Omega \times \Omega$ there exists a path*

$$X_t = Z_0, Z_1, \ldots, Z_r = Y_t$$

between X_t and Y_t where $(Z_l, Z_{l+1}) \in S$ for $0 \leq l < r$ and $\sum_{l=0}^{r-1} \delta(Z_l, Z_{l+1}) = \delta(X_t, Y_t)$. Define a coupling $(X, Y) \mapsto (X', Y')$ of the Markov chain $\mathcal{M}$ on all pairs $(X, Y) \in S$. Suppose that there exists $\beta < 1$ such that

$$\mathbf{E}\left[\delta(X', Y')\right] \leq \beta\, \delta(X, Y)$$

for all $(X, Y) \in S$. Then the mixing rate $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies

$$\tau(\varepsilon) \leq \frac{\log(D\varepsilon^{-1})}{1 - \beta}.$$

[2] Elsewhere, the *mixing rate* is sometimes defined to be $\tau_x(\varepsilon)$ for some fixed x.

Remark 1.1 The set S is often taken to be

$$S = \{(X, Y) \in \Omega \times \Omega : \delta(X, Y) = 1\}.$$

Here a coupling need only be defined for pairs at distance 1 apart.

Remark 1.2 Notice that Lemma 1.1 does not assume that the Markov chain $\mathcal{M}$ is reversible.

1.2 Colourings of graphs.

Let $G = (V, E)$ be a graph on n vertices and let Δ be the maximum degree of G. Let k be a positive integer and let C be a set of size k. A map from V to C is called a *k-colouring.* A vertex v is said to be *properly coloured* in the colouring X if v is coloured differently from all of its neighbours. A colouring X is called *proper* if every vertex is properly coloured in X. A necessary and sufficient condition for the existence of a proper k-colouring of all graphs with maximum degree Δ is $k \geq \Delta + 1$. Denote by $\Omega_k(G)$ the set of all proper k-colourings of G. The colour assigned to a vertex v in the colouring X is denoted by $X(v)$.

Jerrum [12] and Salas and Sokal [18] independently defined a Markov chain with state space $\Omega_k(G)$ which is irreducible for $k \geq \Delta + 2$ and rapidly mixing for $k \geq 2\Delta$. One version of this chain has the following transition procedure: from current state X, choose a vertex v uniformly at random and choose a colour c uniformly at random from the set of those colours which properly colour v in X. Then recolour v with c to give the new state. A new Markov chain of colourings, denoted by $\mathcal{M}_1(\Omega_k(G))$, was introduced in [9]. This new chain is irreducible for $k \geq \Delta + 1$ and is also rapidly mixing for $k \geq 2\Delta$. Moreover the new chain is provably faster than the Jerrum chain when G is Δ-regular or (after a slight adjustment to the chain) when $2\Delta \leq k \leq 3\Delta - 1$. When $k = 2\Delta$ the new chain is $\Omega^*(n^2)$ times faster than the Jerrum chain (see [9]). The transitions of $\mathcal{M}_1(\Omega_k(G))$ are defined by the following procedure: choose an edge $\{v, w\}$ uniformly at random from E and choose an ordered pair of colours $(c(v), c(w))$ uniformly at random from the set of all those such that both v and w are properly coloured when v is recoloured $c(v)$ and w is recoloured $c(w)$. The resulting colouring is the new state.

2 Definition of a heat bath Markov chain

The notation is adapted from that used in [3]. Let V and C be finite sets and let Ω be a subset of C^V, the set of functions from V to C. Let L be a subset of the power set of V. We shall refer to the elements of L as *lines.* For $X \in \Omega$, $\ell \in L$ and $c \in C^\ell$ let $X_{\ell\to c}$ denote the element of C^V defined by

$$X_{\ell\to c}(u) = \begin{cases} c(u) & \text{if } u \in \ell, \\ X(u) & \text{otherwise.} \end{cases}$$

If $X_{\ell\to c} \in \Omega$ then we say that c is *acceptable* at ℓ in X. Finally let $\mathcal{S}_X(\ell)$ be the set of all elements of C^ℓ which are acceptable at ℓ in X. Let π be a distribution on Ω and, for all $\ell \in L$ and all $X \in \Omega$ let $\pi^*_X(\ell)$ be defined by

$$\pi^*_X(\ell) = \sum_{c \in \mathcal{S}_X(\ell)} \pi(X_{\ell\to c}).$$

The set L of lines and the distribution π can be used to define a Markov chain $\mathcal{M}(L, \pi)$ with state space Ω. The transition procedure from the current state X is as follows:

(i) choose ℓ uniformly at random from L,

(ii) choose $c \in \mathcal{S}_X(\ell)$ with probability $\pi(X_{\ell \to c})/\pi^*_X(\ell)$ and move to $X_{\ell \to c}$.

Clearly the Markov chain $\mathcal{M}(L, \pi)$ is aperiodic.

The definition of a heat bath Markov chain can now be stated.

Definition 2.1 *A Markov chain $\mathcal{M}$ with state space $\Omega \subseteq C^V$ is said to be a* heat bath *Markov chain if there exists a set of lines L and a distribution π on Ω such that $\mathcal{M} = \mathcal{M}(L, \pi)$.*

Suppose that $\mathcal{M}(L, \pi)$ satisfies Definition 2.1. The transition matrix P of $\mathcal{M}(L, \pi)$ has entries

$$P(X, Y) = \pi(Y) \sum_{\ell \supseteq D} (|L||\pi^*_X(\ell)|)^{-1}$$

where $D = \{v \in V : X(v) \neq Y(v)\}$. If $\ell \supseteq D$ then $\mathcal{S}_X(\ell) = \mathcal{S}_Y(\ell)$ and $\pi^*_X(\ell) = \pi^*_Y(\ell)$. Therefore P is reversible with respect to π. If $\mathcal{M}(L, \pi)$ is also irreducible then it is an ergodic chain with stationary distribution π. Say that the set L of lines is *sufficient* if $\mathcal{M}(L, \pi)$ is irreducible.

In the special case that π is the uniform distribution on Ω we will write $\mathcal{M}(L)$ instead of $\mathcal{M}(L, \pi)$. Here the transition procedure of $\mathcal{M}(L)$ involves choosing a line $\ell \in L$ uniformly at random and a colour map $c \in \mathcal{S}_X(\ell)$ uniformly at random and making a transition to $X_{\ell \to c}$. All the heat bath Markov chains considered in this paper involve the uniform distribution. Two examples are the the two Markov chains of k-colourings of a graph described in section 1.2. Both fit Definition 2.1, where π is the uniform distribution on $\Omega_k(G)$. For the Jerrum/Salas–Sokal chain the set L of lines is just the set of singleton vertices. In the case of the chain $\mathcal{M}_1(\Omega_k(G))$ the set of lines L is the edge set E. In both cases the set L of lines is sufficient, so the corresponding Markov chain is ergodic.

Let $H(X, Y)$ denote the *Hamming distance* between X and Y, defined by

$$H(X, Y) = |\{v \in V : X(v) \neq Y(v)\}|.$$

Suppose we want to analyse the mixing rate of $\mathcal{M}(L)$ using path coupling on pairs at Hamming distance 1 apart. Let X and Y be two elements of Ω which differ just at $v \in V$. A coupling $(X, Y) \mapsto (X', Y')$ must be defined for every $\ell \in L$. Call a coupling at ℓ *optimal* if it minimises the expected value of $H(X', Y')$. Note that the *optimal* coupling defined here is not the same as the *maximal* coupling which is referred to elsewhere in the literature. Refer to $\mathbf{E}[H(X', Y') - 1]$ as the *cost* of the coupling. If $\mathcal{S}_X(\ell) = \mathcal{S}_Y(\ell)$ then an optimal coupling at ℓ is obtained by choosing the same element $c \in \mathcal{S}_X(\ell)$ in both X and Y. Here $H(X', Y') = 0$ if v lies on the line ℓ and $H(X', Y') = 1$ otherwise. Suppose now that $\mathcal{S}_X(\ell) \neq \mathcal{S}_Y(\ell)$. It follows that $v \notin \ell$. Given $c_X \in \mathcal{S}_X(\ell)$ and $c_Y \in \mathcal{S}_Y(\ell)$, let $P(c_X, c_Y)$ denote the probability that the pair (c_X, c_Y) is chosen by the coupling and let

$$h(c_X, c_Y) = |\{u \in \ell : c_X(u) \neq c_Y(u)\}|,$$

the number of elements of ℓ which are assigned different values by c_X and c_Y. Then the expected value of $H(X', Y') - 1$ is given by

$$\sum P(c_X, c_Y) h(c_X, c_Y),$$

where the sum is over all $(c_X, c_Y) \in \mathcal{S}_X(\ell) \times \mathcal{S}_Y(\ell)$.

3 The related transportation problems

In this section it will be shown that an optimal coupling for a heat bath Markov chain is equivalent to the optimal solutions of a related set of *transportation problems*. For more information on transportation problems see, for example [16]. Let m and n be positive integers and let K be an $m \times n$ matrix of nonnegative integers (the *cost matrix*). Let a be a vector of m positive integers and let b be a vector of n positive integers such that $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j = N$. An $m \times n$ matrix Z of nonnegative numbers is a solution of the transportation problem defined by a, b, and K if $\sum_{j=1}^{n} Z_{i,j} = a_i$ for $1 \leq i \leq m$ and $\sum_{i=1}^{m} Z_{i,j} = b_j$ for $1 \leq j \leq n$. The *cost* of this solution is measured by $\sum_{i=1}^{m} \sum_{j=1}^{n} K_{i,j} Z_{i,j}$. The elements of a are called *row sums* and the elements of b are called *column sums*. An *optimal solution* of this transportation problem is a solution which minimises the cost. Every entry of an optimal solution is a nonnegative integer. Efficient algorithms exist for solving transportation problems.

Returning to the problem of defining an optimal coupling for the heat bath Markov chain $\mathcal{M}(L)$ at the line ℓ, let $N_X = |\mathcal{S}_X(\ell)|$ and $N_Y = |\mathcal{S}_Y(\ell)|$. Let a be the N_X-dimensional vector with each entry equal to N_Y and let b be the N_Y-dimensional vector with each entry equal to N_X. Let K be the matrix with N_X rows and N_Y columns, corresponding to the elements of $\mathcal{S}_X(\ell)$ and $\mathcal{S}_Y(\ell)$ respectively, such that the (c_X, c_Y)th entry of K is $h(c_X, c_Y)$. A coupling at ℓ defines an $N_X \times N_Y$ matrix Z with (c_X, c_Y)th entry given by

$$N_X N_Y \operatorname{Prob}\left[X' = X_{\ell \to c_X}, Y' = Y_{\ell \to c_Y}\right].$$

The matrix Z is a solution of the transportation problem defined by a, b, and K. An optimal solution of this transportation problem corresponds to a optimal coupling at ℓ. The cost of an optimal solution equals $N_X N_Y$ times the cost of an optimal coupling. Therefore one can attempt to prove that $\mathcal{M}(L)$ is rapidly mixing by adapting the path coupling method as follows. Here 'nonisomorphic' means 'nonisomorphic in some appropriate sense'.

(i) compile a complete list of nonisomorphic pairs (X, Y) such that $H(X, Y) = 1$,

(ii) for each such pair, calculate the contribution of all lines ℓ such that $\mathcal{S}_X(\ell) = \mathcal{S}_Y(\ell)$,

(iii) for all other lines ℓ, solve the corresponding transportation problem to find the cost of the optimal coupling,

(iv) combine all this information to determine whether the expected value of $H(X', Y') - 1$ is negative in all cases. If this is the case then let the maximum of these values be $\beta - 1$. The mixing time $\tau(\varepsilon)$ of $\mathcal{M}(L)$ satisfies

$$\tau(\varepsilon) \leq \frac{\log(n\varepsilon^{-1})}{1 - \beta},$$

by Lemma 1.1. If $1 - \beta$ is only polynomially small then $\mathcal{M}(L)$ is rapidly mixing.

Remark 3.1 In order to prove that the Markov chain is rapidly mixing it may not be necessary to find an optimal solution in each case. Instead, it may be sufficient to find a solution with a low enough cost in each case.

Remark 3.2 Certainly the list of pairs (X, Y) used in the above procedure must contain at least one element from each isomorphism class in order that the calculations are conclusive. If the list is a transversal then no unnecessary calculations are performed. In many cases however the amount of effort required to find a transversal of the isomorphism classes is prohibitive and ruling out obviously isomorphic pairs will suffice. Moreover, in most cases one need only consider a certain 'neighbourhood' around the line ℓ, rather than the entire maps X, Y.

4 Five-colouring degree-three graphs

Let G be a graph with maximum degree 3 and let $\Omega_k(G)$ denote the set of proper k-colourings of G. The two known Markov chains on k-colourings described in section 1.2 are rapidly mixing for $k \geq 6$. In this section a Markov chain on $\Omega_5(G)$ is defined and a computational proof that the chain is rapidly mixing is given following the method described in section 3.

Let $C = \{1, \dots, 5\}$ be the colour set and let $\mathcal{M}$ be the following Markov chain on $\Omega_5(G)$. If w is a vertex with degree d then let the neighbours of w be denoted by $u_1(w), \dots, u_d(w)$. Denote by $\tilde{w}$ the set defined by

$$\tilde{w} = \{w, u_1(w), \dots, u_d(w)\}.$$

Define the set of lines, L, by

$$L = \{\tilde{w} : w \in V\}.$$

Then $\mathcal{S}_X(\tilde{w})$ is

$$\mathcal{S}_X(\tilde{w}) = \left\{c \in C^{\tilde{w}} : X_{\tilde{w}\to c} \in \Omega_5(G)\right\}. \tag{1}$$

The Markov chain $\mathcal{M} = \mathcal{M}(L)$ is a heat bath Markov chain in the sense of Definition 2.1, where π is the uniform distribution on $\Omega_5(G)$. The transitions of $\mathcal{M}$ from current state X follow the pattern described in section 2: a line $\tilde{w} \in L$ is chosen uniformly at random; a corresponding colour mapping $c \in \mathcal{S}_X(\tilde{w})$ is chosen uniformly at random and a transition is made to $X_{\tilde{w}\to c}$. The chain $\mathcal{M}$ is irreducible: for example it can perform all the moves of the Jerrum/Salas–Sokal chain. Therefore the set of lines L is sufficient and the chain $\mathcal{M}$ is ergodic with stationary distribution the uniform distribution on $\Omega_5(G)$.

The mixing rate of $\mathcal{M}$ will now be analysed using path coupling on pairs at Hamming distance 1 apart. Let $X, Y \in \Omega_5(G)$ and let $d = H(X, Y)$. It is not always possible to form a path

$$X = Z_0, Z_1, \dots, Z_d = Y$$

of length d between $X, Y \in \Omega_5(G)$ where $H(Z_l, Z_{l+1}) = 1$ for $0 \leq l < d$ and $Z_l \in \Omega_5(G)$ for $0 \leq l \leq d$. However we can always form such a path where $Z_l \in C^V$ for $0 \leq l \leq d$. If $X \in \Omega_5(G)$ then the definition of the set $\mathcal{S}_X(\tilde{w})$ given in (1) is equivalent to the following definition:

$$\mathcal{S}_X(\tilde{w}) = \left\{c \in C^{\tilde{w}} : \text{all elements of } \tilde{w} \text{ are properly coloured in } X_{\tilde{w}\to c}\right\}.$$

The latter definition makes sense for all $X \in C^V$. Using the latter definition, we can extend the chain $\mathcal{M}$ to act on the state space C^V. It is easy to see that the extended chain has the stationary distribution:

$$\pi(X) = \begin{cases} |\Omega_5(G)|^{-1} & \text{if } X \in \Omega_5(G), \\ 0 & \text{otherwise.} \end{cases}$$

The extended chain is no longer reversible. However, by Remark 1.2 we may apply Lemma 1.1 to the extended chain. If the extended chain is rapidly mixing then the original chain is also rapidly mixing with the same upper bound on the mixing rate. This follows since mixing time is defined as the maximum over all initial states, and the original chain only involves a subset of these.

Now suppose that X and Y are two colourings which differ only at v. Without loss of generality suppose that $X(v) = 1$ and $Y(v) = 2$. Denote the degree of v by d_v. Let w be a vertex for which $\tilde{w}$ corresponds to the line chosen in the transition procedure. If $w = v$ or $\{w, v\} \in E$ then $\mathcal{S}_X(\tilde{w}) = \mathcal{S}_Y(\tilde{w})$. Therefore the same choice of c may be made in X and Y. It follows that these vertices contribute $-(d_v + 1)/n$ to the expected value of $H(X', Y') - 1$. The only other lines which may make nonzero contributions to the expected value of $H(X', Y') - 1$ are those corresponding to vertices w which satisfy the following criteria:

(i) $w \neq v$ and $\{w, v\} \notin E$,

(ii) there exists $u \in V$ such that $\{w, u\} \in E$ and $\{u, v\} \in E$.

Call such a vertex a *critical* vertex. Define the *multiplicity* $\mu(w)$ of a critical vertex w to be the number of neighbours of w which are neighbours of v. Now w may have up to three neighbours. If w has degree 2 or 3 then label its second neighbour by z_1: if w has degree 3 then label its third neighbour by z_2. If z_i is present then it may have up to three neighbours, for $i = 1, 2$. Let the second and third neighbours of z_i be labelled $z_{i,1}$, $z_{i,2}$ if they are present, for $i = 1, 2$. Finally u may have a third neighbour which we denote by z_3 if it is present. There are fifteen nonisomorphic configurations on these vertices, depending on the degree of w and on whether the edges

$$\{v, z_1\}, \{v, z_2\}, \{z_1, z_2\}, \{z_1, u\}$$

are present or absent in E. The details are given in Table 2, where the second column holds the degree of w and the third column holds a string of four characters, each equal to 'Y' or 'N'. The characters refer to the four edges listed above in that order and 'Y' indicates that the edge is present in E while 'N' indicates that it is absent. The fourth column of Table 2 holds $\mu_i(w)$, the multiplicity of w in the ith configuration. The last two columns of Table 2 are explained below. Note that any identification or edges between (without loss of generality) the vertices $z_{1,1}, z_{2,1}, z_3$ can be ignored, as this situation can be modelled by colouring the affected vertices with the same (respectively different) colours. The only vertices which affect the sets $\mathcal{S}_X(\tilde{w})$, $\mathcal{S}_Y(\tilde{w})$ are the vertices

$$v, z_3, z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}.$$

Call these vertices the *rim vertices* and call an assignation of colours to these vertices a *rim colouring*. Following Remark 3.2, it suffices to consider a complete list of nonisomorphic rim colourings for each configuration. Let λ_i be the number of nonisomorphic rim colourings of the ith configuration, $1 \le i \le 15$. These values are listed in the fifth column of Table 2, while a complete list of the nonisomorphic rim colourings for each configuration can be found in Table 1. Each rim colouring is shown as a list of six characters representing the colours of the six rim vertices in the order given above. If a rim vertex equals v then its colour is given as 0. Similarly, if a rim vertex equals one of u, z_1, or z_2, then its colour is given as ρ. If a rim vertex is absent then its colour is given as ω. The first element of each rim colouring is always 0 but it is included so that any symmetry in the configuration is apparent.

Suppose that in a given configuration the vertex z_1 is present and is not joined by an edge to v, u (or z_2 if present). Then it suffices to assume that z_1 has degree 3, as follows. If z_1 has degree 2 then its third neighbour $z_{1,2}$ may be added and coloured with the same colour as $z_{1,1}$ without affecting the sets $\mathcal{S}_X(\tilde{w})$, $\mathcal{S}_Y(\tilde{w})$. If z_1 has degree 1 then the sets $\mathcal{S}_X(\tilde{w})$, $\mathcal{S}_Y(\tilde{w})$ respectively contain exactly four times as many elements as those obtained in the case that vertex z_1 is not present. Let a, b, and K be the row sums, column sums and cost matrix obtained when z_1 is absent and let a', b', and K' be those obtained when z_1 is present. Then K' has the block form

$$K' = \begin{bmatrix} K & * & * & * \\ * & K & * & * \\ * & * & K & * \\ * & * & * & K \end{bmatrix}.$$

and the vector a' (respectively b') consists of four copies of the vector $4a$ (respectively $4b$) concatenated together. It is not hard to see that the cost of an optimal solution of the transportation problem defined by a', b' and K' is bounded above by $16k$, where k is the cost of an optimal solution of the transportation problem defined by a, b and K. Therefore the cost of a optimal coupling when z_1 is present is bounded above by the cost of a optimal coupling when z_1 is absent. The same argument holds with the roles of z_1 and z_2 reversed and reduces the number of rim colourings for certain configurations.

For each rim colouring of a given configuration the sets $\mathcal{S}_X(\tilde{w})$, $\mathcal{S}_Y(\tilde{w})$ can be formed. From these the matrix K of costs can be obtained and given as input to an algorithm for the transportation problem. The cost of an optimal solution corresponds to the value of $N_X N_Y \mathbf{E}\left[H(X',Y')-1\right]$ for an optimal coupling of X and Y at w. Let $M_i(w)$ be the greatest cost of an optimal coupling over all possible rim colourings for the ith configuration, $1 \le i \le 15$. These values are listed in the sixth column of Table 2. These results lead to the following theorem.

Theorem 4.1 *The Markov chain $\mathcal{M}$ is rapidly mixing with mixing rate*

$$\tau(\varepsilon) \le \frac{161}{144} n \log(n\varepsilon^{-1}).$$

Proof. The maximum contribution of a critical vertex in the ith configuration is denoted by $M_i(w)$. These values were calculated using an algorithm for the transportation problem and are listed in Table 2. Denote by d_v the degree of the vertex v. If $d_v = 0$ then the expected value of $H(X',Y')-1$ is $-1/n$. If $d_v > 0$ then the expected value of $H(X',Y')-1$ in the ith configuration is given by

$$\mathbf{E}\left[H(X',Y')-1\right] \le -\frac{(d_v+1)}{n} + \sum_{w} \frac{M_i(w)}{n}$$

Config 1:	$01\omega\omega\omega\omega$	$03\omega\omega\omega\omega$	$0\omega\omega\omega\omega\omega$				
Config 2:	$0111\omega\omega$	$0122\omega\omega$	$0134\omega\omega$	$0313\omega\omega$	$0334\omega\omega$	$0\omega11\omega\omega$	$0\omega33\omega\omega$
	$0112\omega\omega$	$0123\omega\omega$	$0311\omega\omega$	$0314\omega\omega$	$0344\omega\omega$	$0\omega12\omega\omega$	$0\omega34\omega\omega$
	$0113\omega\omega$	$0133\omega\omega$	$0312\omega\omega$	$0333\omega\omega$	$0345\omega\omega$	$0\omega13\omega\omega$	
Config 3:	$0\rho1\rho\omega\omega$	$0\rho3\rho\omega\omega$	$0\rho\omega\rho\omega\omega$				
Config 4:	$0101\omega\omega$	$0104\omega\omega$	$0304\omega\omega$	$0\omega03\omega\omega$			
	$0103\omega\omega$	$0303\omega\omega$	$0\omega01\omega\omega$	$0\omega0\omega\omega\omega$			
Config 5:	$0\rho0\rho\omega\omega$						
Config 6:	011111	011323	013333	031145	031414	033444	$0\omega1233$
	011112	011324	013334	031212	031415	033445	$0\omega1234$
	011113	011333	013344	031213	031424	033455	$0\omega1313$
	011122	011334	013345	031214	031425	034444	$0\omega1314$
	011123	011344	013434	031233	031433	034445	$0\omega1323$
	011133	011345	013435	031234	031434	034455	$0\omega1324$
	011134	012222	031111	031244	031435	034545	$0\omega1333$
	011212	012223	031112	031245	031444	$0\omega1111$	$0\omega1334$
	011213	012233	031113	031313	031445	$0\omega1112$	$0\omega1344$
	011222	012234	031114	031314	031455	$0\omega1113$	$0\omega1345$
	011223	012323	031122	031323	033333	$0\omega1122$	$0\omega3333$
	011233	012324	031123	031324	033334	$0\omega1123$	$0\omega3334$
	011234	012333	031124	031333	033344	$0\omega1133$	$0\omega3344$
	011313	012334	031133	031334	033345	$0\omega1134$	$0\omega3345$
	011314	012344	031134	031344	033434	$0\omega1212$	$0\omega3433$
	011322	012345	031144	031345	033435	$0\omega1213$	$0\omega3435$
Config 7:	010111	010213	010324	030333	030434	$0\omega0134$	$0\omega0\omega11$
	010112	010233	010333	030334	030435	$0\omega0311$	$0\omega0\omega12$
	010113	010234	010334	030344	030455	$0\omega0312$	$0\omega0\omega13$
	010122	010311	010344	030345	$0\omega0111$	$0\omega0313$	$0\omega0\omega33$
	010123	010312	010345	030411	$0\omega0112$	$0\omega0314$	$0\omega0\omega34$
	010133	010313	030311	030412	$0\omega0113$	$0\omega0333$	
	010134	010314	030312	030413	$0\omega0122$	$0\omega0334$	
	010211	010322	030313	030415	$0\omega0123$	$0\omega0344$	
	010212	010323	030314	030433	$0\omega0133$	$0\omega0345$	
Config 8:	010101	010203	030303	$0\omega0101$	$0\omega0303$	$0\omega0\omega03$	
	010102	010303	030304	$0\omega0102$	$0\omega0304$	$0\omega0\omega0\omega$	
	010103	010304	030405	$0\omega0103$	$0\omega0\omega01$		
Config 9:	$011\rho1\rho$	$013\rho3\rho$	$031\rho4\rho$	$01\omega\rho1\rho$	$03\omega\rho4\rho$	$0\omega1\rho3\rho$	$0\omega\omega\rho\omega\rho$
	$011\rho2\rho$	$013\rho4\rho$	$033\rho3\rho$	$01\omega\rho2\rho$	$01\omega\rho\omega\rho$	$0\omega3\rho3\rho$	
	$011\rho3\rho$	$031\rho1\rho$	$033\rho4\rho$	$01\omega\rho3\rho$	$03\omega\rho\omega\rho$	$0\omega3\rho4\rho$	
	$012\rho2\rho$	$031\rho2\rho$	$034\rho4\rho$	$03\omega\rho1\rho$	$0\omega1\rho1\rho$	$0\omega\omega\rho1\rho$	
	$012\rho3\rho$	$031\rho3\rho$	$034\rho5\rho$	$03\omega\rho3\rho$	$0\omega1\rho2\rho$	$0\omega\omega\rho3\rho$	
Config 10:	$010\rho1\rho$	$010\rho3\rho$	$030\rho3\rho$	$010\rho\omega\rho$	$0\omega0\rho1\rho$	$0\omega0\rho\omega\rho$	
	$010\rho2\rho$	$030\rho1\rho$	$030\rho4\rho$	$030\rho\omega\rho$	$0\omega0\rho3\rho$		
Config 11:	$010\rho0\rho$	$030\rho0\rho$	$0\omega0\rho0\rho$				
Config 12:	$0\rho1\rho11$	$0\rho1\rho22$	$0\rho1\rho34$	$0\rho3\rho13$	$0\rho3\rho34$	$0\rho\omega\rho11$	$0\rho\omega\rho33$
	$0\rho1\rho12$	$0\rho1\rho23$	$0\rho3\rho11$	$0\rho3\rho14$	$0\rho3\rho44$	$0\rho\omega\rho12$	$0\rho\omega\rho34$
	$0\rho1\rho13$	$0\rho1\rho33$	$0\rho3\rho12$	$0\rho3\rho33$	$0\rho3\rho45$	$0\rho\omega\rho13$	
Config 13:	$0\rho0\rho11$	$0\rho0\rho12$	$0\rho0\rho13$	$0\rho0\rho33$	$0\rho0\rho34$		
Config 14:	$0\rho\rho\rho1\rho$	$0\rho\rho\rho3\rho$	$0\rho\rho\rho\omega\rho$				
Config 15:	$0\rho\rho\rho0\rho$						

Table 1: The rim colourings for Configurations 1–15.

for $1 \leq i \leq 15$, where the sum is over all critical w. By inspection of Table 2 it is clear

i	d_w	edges	$\mu_i(w)$	λ_i	$M_i(w)$
1	1	NNNN	1	3	5/12
2	2	NNNN	1	20	13/28
3	2	NNNY	1	3	73/156
4	2	YNNN	2	8	25/29
5	2	YNNY	2	1	2/3
6	3	NNNN	1	112	381/748
7	3	YNNN	2	59	2230/2343
8	3	YYNN	3	17	94/71
9	3	NNYN	1	31	527/1081
10	3	YNYN	2	11	1888/2115
11	3	YYYN	3	3	10/9
12	3	NNNY	1	20	250/483
13	3	YNNY	2	5	67/90
14	3	NNYY	1	3	25/57
15	3	NYYY	2	1	3/4

Table 2: The 15 configurations.

that

$$M_i(w) < \frac{2\mu_i(w)}{3} \leq \frac{(d_v+1)\mu_i(w)}{2d_v}$$

for $1 \leq i \leq 15$. Therefore the expected value of $H(X', Y') - 1$ is less than

$$\frac{(d_v+1)}{2d_v n}\left(\sum_w \mu_i(w) - 2d_v\right) \leq 0$$

for $1 \leq i \leq 15$. To compute an upper bound for the mixing rate, let

$$T = \max\left\{M_i(w)/\mu_i(w) : 1 \leq i \leq 15\right\}.$$

By inspection of Table 2 one can see that $T = 250/483$, corresponding to Configuration 12. Then

$$\mathbf{E}\left[H(X', Y') - 1\right] \leq -\frac{4}{n} + \frac{6T}{n} = -\frac{144}{161n}.$$

Therefore, by Lemma 1.1, the mixing rate of $\mathcal{M}$ is as stated.

Remark 4.1 In order to prove that $\mathcal{M}$ is rapidly mixing it suffices to establish the inequality $T < 2/3$, as shown in the proof of Theorem 4.1. To show this, the exact value of $M_i(w)$ was calculated for $1 \leq i \leq 15$. As mentioned in Remark 3.1, much calculation can be saved by halting the algorithm for the transportation problem in each case as soon as a feasible solution with low enough cost has been found. In two separate calculations using different heuristics, the authors established the two bounds $T \leq 1174/1767$ and $T \leq 515/966$ using this method. The first gives rise to the upper bound

$$\tau(\varepsilon) \leq \frac{589}{8} n \log(n\varepsilon^{-1})$$

on the mixing rate of $\mathcal{M}$, which is more than 65 times greater than the bound given in Theorem 4.1. This illustrates that the increased efficiency which results in this way may incur a significant loss of tightness in the bound on the mixing rate. However, the second calculation gives rise to the bound

$$\tau(\varepsilon) \le \frac{161}{129} n \log(n\varepsilon^{-1}),$$

which is very close to the bound provided by Theorem 4.1.

5 Further applications

In principle it is not difficult to extend the result of section 4 to graphs of larger (bounded) degree—the algorithm for generating all the non-isomorphic configurations readily extends to this case. However there is a problem that prevents us from doing this: the number of rim colourings suffers from combinatorial explosion, as does the size of the transportation problems to be solved. This may be offset in some fashion by restricting attention to smaller classes of graphs. The high-degree of symmetry inherent in lattice graphs makes these ideal candidates.

It would be interesting to see just how general these classes of graphs can be made before the computation becomes intractable. Determining the cut-off point is a largely subjective matter and dependent on available resources. One additional computation that we discovered to be tractable is the proof of rapid mixing of $\Omega_7(G)$, where G is a 4-regular, triangle-free graph. It should be clear where the simplifications arise from these two additional restrictions over the case considered in section 4: the regularity condition ensures that we need not consider the cases where some of the rim vertices are absent, and demanding that the graph is triangle-free means that we do not need to consider the cases where some of the rim vertices are adjacent to the critical vertex that they rim.

To further save on the amount of computational time needed, we did not actually find the optimum solutions of the transportation problems. Instead we made use of Remark 3.1, since it transpired that using the matrix minimum heuristic was sufficient.

We will borrow some of the notation of section 4 for the remainder of this section.

In order to ensure that the chain was rapidly mixing, we needed to show that we had

$$\mathbf{E}\left[H(X', Y') - 1\right] \le -\frac{5}{n} + \sum_{w} \frac{M_i(w)}{n}$$

for all critical vertices w. (The 5 arises from the fact that there are 5 choices of vertex, v and all of its neighbours, whose choice would ensure that $H(X', Y') = 0$.)

It is sufficent therefore to show that for each rim colouring, Q, we have

$$\text{cost}(Q) \times \frac{12}{\mu(w)} - 5 \le 0 \quad \text{i.e. } \text{cost}(Q) \le \frac{5}{12}\mu(w) \tag{2}$$

where $\text{cost}(Q)$ is the cost of the optimal solution to the transportation problem associated with Q.

It was necessary to consider 42574 non-isomorphic rim colourings. The worst cases for the different values of $\mu(w)$ are listed in Table 3. Each rim colouring is shown as a list of 12 numbers in four groups of three. Each group of three numbers corresponds to a neighbour z of w and represents the colours of the neighbours of z other than w. The notation for the colours is as explained in section 4 (although here the set of colours is

Rim colouring	$\mu(w)$	Proven bound
0 3 4 5 5 6 5 5 6 5 6 7	1	62/189
0 3 4 5 5 6 5 6 6 5 6 7	1	62/189
0 3 4 0 5 6 3 3 7 4 4 7	2	950/1489
0 3 4 0 5 6 3 3 7 4 7 7	2	950/1489
0 3 4 0 5 6 3 3 7 5 5 7	2	950/1489
0 3 4 0 5 6 3 3 7 5 7 7	2	950/1489
0 3 4 0 5 6 3 7 7 4 7 7	2	950/1489
0 3 4 0 5 6 3 7 7 5 7 7	2	950/1489
0 3 4 0 3 5 0 6 7 3 4 5	3	914/969
0 3 4 0 3 5 0 6 7 3 4 6	3	914/969
0 3 4 0 3 5 0 6 7 3 6 7	3	914/969
0 3 4 0 3 4 0 3 5 0 6 7	4	1212/997
0 3 4 0 3 5 0 3 6 0 4 7	4	1212/997

Table 3: The worst configurations for 7-colouring 4-regular triangle-free graphs.

$\{1, \ldots, 7\}$, with seven elements). By inspection, all of these worst case rim colourings satisfy equation (2), from which the rapid mixing result follows.

The mixing rate may be bounded by considering the largest value of the proven bound over $\mu(w)$ in Table 3. This is 62/189. Thus $\mathbf{E}[H(X', Y') - 1] \le -67/63n$. It follows then from Lemma 1.1 that the Markov chain $\mathcal{M}$ is rapidly mixing with mixing rate bounded above by

$$\frac{63}{67} n \log(n\varepsilon^{-1}).$$

Remark 5.1 A method is outlined in [7] whereby, under certain conditions, the rapid mixing of a given Markov chain can be used to deduce the rapid mixing of a second Markov chain with the same state space and stationary distribution as the first. Using this and the results of section 4 it is possible to prove that the simple Jerrum/Salas–Sokal chain is rapidly mixing when $k = 5$ for graphs of maximum degree 3. Similarly, using the results of this section it is possible to prove that the simple chain is rapidly mixing when $k = 7$ for triangle-free 4-regular graphs. The mixing rate of the Jerrum/Salas–Sokal chain is bounded above by $O(n^8 \log(n))$ in both cases. This demonstrates that the 2Δ barrier can be broken even using the simple Jerrum/Salas–Sokal chain for these values of k and families of graphs.

6 A $\#P$-completeness proof

In this section we present a proof that the problem of counting the number of k-colourings of graphs with maximum degree Δ is $\#P$-complete for fixed k, Δ such that $k \ge 3$ and $\Delta \ge 3$. First we must establish a preliminary result involving the *chain* graph on $r + 1$ vertices $u_1, \ldots, u_{r+1}$ with edges $\{u_i, u_{i+1}\}$ for $1 \le i \le r$.

Lemma 6.1 *Let C_r denote the chain graph with $r + 1$ vertices $u_1, \ldots, u_{r+1}$ and r edges $\{u_i, u_{i+1}\}$ for $1 \le i \le r$. Let γ, γ' be two fixed, distinct colours and let σ_r, δ_r be defined* *by*

$$\begin{aligned}\sigma_r &= |\{X \in \Omega_k(C_r) : X(u_1) = \gamma,\ \ X(u_{r+1}) = \gamma\}|, \\ \delta_r &= |\{X \in \Omega_k(C_r) : X(u_1) = \gamma,\ \ X(u_{r+1}) = \gamma'\}|\end{aligned}$$

for $r \geq 1$. *Then*

$$\sigma_r = \frac{(k-1)^r - (-1)^{r-1}(k-1)}{k} \qquad \text{and} \qquad \delta_r = \frac{(k-1)^r + (-1)^{r-1}}{k}$$

for $r \geq 1$.

Proof. This result follows easily by induction, using the following observations. When $r = 1$ the equations give the correct values $\sigma_1 = 0$ and $\delta_1 = 1$. If $r > 1$ then

$$\sigma_r = (k-1)\delta_{r-1}, \qquad \text{and} \qquad \delta_r = \sigma_{r-1} + (k-2)\delta_{r-1}$$

giving the inductive step. □

We now show that it is $\#P$-complete to count the number of proper k-colourings of arbitrary graphs of maximum degree Δ when $k \geq 3$ and $\Delta \geq 3$.

Theorem 6.1 *Denote by* $\#kColMaxDeg\Delta$ *the problem of computing* $|\Omega_k(G)|$ *for an arbitrary graph* G *of maximum degree* Δ. *If* $k \geq 3$ *and* $\Delta \geq 3$ *then* $\#kColMaxDeg\Delta$ *is* $\#P$*-complete.*

Proof. Let $\#kCol$ denote the problem of computing the number of proper k-colourings for an arbitrary graph. If $k \geq 3$ then, by an immediate corollary of Vertigan's thesis (see [11]), the problem $\#kCol$ is $\#P$-complete.

Let Π be the problem which takes as an instance a graph G with maximum degree Δ and bipartition of the edge set $E = M \cup B$, and asks how many k-colourings of G are there which make every edge in M monochromatic and every edge in B bichromatic. There is an easy reduction from $\#kCol$ to Π, which we now describe. Given an instance H of $\#kCol$, define an instance G of Π as follows. Let the vertex set of G equal the vertex set of H initially and let B equal the edge set of H, $M = \emptyset$. Let Δ' be the maximum degree of H. Suppose first that $\Delta' > \Delta$. Then replace each vertex v with degree greater than Δ by a chain of new vertices, each of degree at most Δ, joined together by edges in M. Ensure that each neighbour of v in H is connected (by an edge in B) to exactly one of these new vertices and that at least one of the new vertices has degree Δ. Next suppose that $\Delta' < \Delta$. In this case choose a vertex v of degree Δ' in H and make v a vertex of degree Δ in G, as follows. Add $\Delta - \Delta'$ vertices w to the vertex set of G, and make each new vertex a neighbour of v by adding the edge $\{v, w\}$ to M. The number of k-colourings of G which make all edges in M monochromatic and all edges in B bichromatic is given by $|\Omega_k(H)|$. This shows that $\#kCol$ is polynomial-time reducible to Π, so Π is $\#P$-complete.

We now give a polynomial-time reduction of Π to $\#kColMaxDeg\Delta$. Let H be an instance of Π with edge bipartition $E = M \cup B$. Let $m = |M|$ and let C_r be the chain graph with $r + 1$ vertices and r edges, as in Lemma 6.1. For $r \geq 1$, let G_r be the r-stretch of H with respect to M. That is, G_r is obtained from H by replacing each edge $\{v, w\}$ in M by a copy of the chain C_r, with the identifications $u_1 = v$, $u_{r+1} = w$.

Let N_s be the number of k-colourings of H where s edges in M are monochromatic and the remaining $m - s$ edges in M (together with all edges in B) are bichromatic. Then

$$|\Omega_k(G_r)| = \sum_{s=0}^{m} N_s \, {\sigma_r}^s \, {\delta_r}^{m-s} = {\delta_r}^m \sum_{s=0}^{m} N_s \, {x_r}^s, \tag{3}$$

where $x_r = \sigma_r / \delta_r$. We may express the equations (3), for $1 \le r \le m + 1$, as a matrix equation

$$\begin{bmatrix} {\delta_1}^{-m} |\Omega_k(G_1)| \\ {\delta_2}^{-m} |\Omega_k(G_2)| \\ \vdots \\ {\delta_{m+1}}^{-m} |\Omega_k(G_{m+1})| \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cdots & {x_1}^m \\ 1 & x_2 & \cdots & {x_2}^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m+1} & \cdots & {x_{m+1}}^m \end{bmatrix} \begin{bmatrix} N_0 \\ N_1 \\ \vdots \\ N_m \end{bmatrix}. \tag{4}$$

Now

$$x_r = 1 - \frac{k}{(k-1)(1-k)^{r-1} + 1},$$

using Lemma 6.1. It follows that the values x_r are distinct. If we knew the values $|\Omega_k(G_r)|$, for $1 \le r \le m + 1$, then the values $N_0, N_1, \ldots, N_m$ may be obtained in polynomial time by inverting the Vandermonde matrix in (4). Since N_m is the answer which satisfies Π, this completes the reduction. Therefore $\#kColMaxDeg\Delta$ is $\#P$-complete.
□

By Theorem 6.1, the counting problem associated with the Jerrum/Salas–Sokal chain [12, 18] and the Dyer–Greenhill chain [9] is $\#P$-complete. Also, by Theorem 6.1 with $\Delta = 3$ and $k = 5$, the counting problem associated with the Markov chain of section 4 is $\#P$-complete.

Further $\#P$-completeness results for graph colourings can be found in [8]. For example, we show that the problem of counting k-colourings in graphs with maximum degree Δ remains $\#P$-complete when restricted to triangle-free graphs, so long as $k \ge 3$ and $\Delta \ge 3$. We also show that the problem of counting proper k-colourings in Δ-regular graphs is $\#P$-complete, for $k \ge 3$ and $\Delta \ge 3$. Finally, we show that it is $\#P$-complete to count k-colourings of triangle-free 4-regular graphs whenever $k \ge 3$. Thus in particular, the counting problem associated with the Markov chain of section 5 is $\#P$-complete.

References

[1] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, in A. Dold and B. Eckmann, eds., *Séminaire de Probabilités XVII 1981/1982*, vol. 986 of *Springer-Verlag Lecture Notes in Mathematics*, Springer-Verlag, New York, 1983 pp. 243–297.

[2] K. Appel and W. Haken, *Every Planar Map is Four Colorable*, vol. 98 of *Contemporary Mathematics*, American Mathematical Society, Providence, RI, (1989).

[3] R. Bubley and M. Dyer, *Path coupling: A technique for proving rapid mixing in Markov chains*, in *38th Annual Symposium on Foundations of Computer Science*, IEEE, Los Alimitos, 1997 pp. 223–231.

[4] E. G. Coffman, D. S. Johnson, P. W. Shor, and R. R. Weber, *Markov chains, computer proofs and average-case analysis of best fit bin packing*, in *25th Annual Symposium on the Theory of Computing*, ACM, New York, 1993 pp. 412–421.

[5] M. Creutz, *A Monte Carlo study of quantised SU(2) gauge theory*, Physical Review D, 21 (1980), pp. 2308–2315.

[6] M. Dyer, A. Frieze, and R. Kannan, *A random polynomial-time algorithm for approximating the volume of convex bodies*, Journal of the ACM, 38 (Jan. 1991) 1, pp. 1–17.

[7] M. Dyer and C. Greenhill, *On Markov chains for independent sets*, (1997), (Preprint).

[8] M. Dyer and C. Greenhill, *Some #P-completeness proofs for colourings and independent sets*, Tech. Rep. 97.47, School of Computer Studies, University of Leeds, (1997).

[9] M. Dyer and C. Greenhill, *A more rapidly mixing Markov chain for graph colourings*, Random Structures and Algorithms, 13 (1998), pp. 285–317.

[10] K. Fredenhagen and M. Marcu, *A modified heat bath method suitable for Monte Carlo simulations on vector and parallel processing*, Physics Letters B, 193 (1987) 4, pp. 486–488.

[11] F. Jaeger, D. L. Vertigan, and D. Welsh, *On the computational complexity of the Jones and Tutte polynomials*, Mathematical Proceedings of the Cambridge Philosophical Society, 108 (1990), pp. 35–53.

[12] M. Jerrum, *A very simple algorithm for estimating the number of k-colorings of a low-degree graph*, Random Structures and Algorithms, 7 (1995), pp. 157–165.

[13] M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM Journal on Computing, 18 (Dec. 1989) 6, pp. 1149–1178.

[14] M. Jerrum and A. Sinclair, *The Markov chain Monte Carlo method: an approach to approximate counting and integration*, in D. Hochbaum, ed., *Approximation Algorithms for NP-Hard Problems*, PWS Publishing, Boston, 1996 pp. 482–520.

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

[16] B. Krekó, *Linear Programming*, Sir Isaac Pitman and Sons Ltd, London, (1968).

[17] R. Y. Rubenstein, *Simulation and the Monte Carlo Method*, Wiley, New York, (1981).

[18] J. Salas and A. D. Sokal, *Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem*, Journal of Statistical Physics, 86 (Feb. 1997) 3–4, pp. 551–579.

[19] B. Toft, *Colouring, stable sets and perfect graphs*, in R. L. Graham, M. Grötschel, and L. Lovász, eds., *Handbook of Combinatorics*, Elsevier, Amsterdam, 1995 pp. 233–288.