# Rapidly mixing Markov chains for dismantleable constraint graphs*

Martin Dyer †  Mark Jerrum ‡  Eric Vigoda §

July 25, 2001

**Abstract**

If $G = (V_G, E_G)$ is an input graph, and $H = (V_H, E_H)$ a fixed constraint graph, we study the set $\Omega$ of homomorphisms (or colorings) from $V_G$ to $V_H$, i.e., functions that preserve adjacency. Brightwell and Winkler introduced the notion of *dismantleable* constraint graph to characterize those $H$ whose associated set $\Omega$ of homomorphisms is, for every $G$, connected under single vertex recolorings. Given fugacities $\lambda(c) > 0$ ($c \in V_H$) our focus is on sampling a coloring $\omega \in \Omega$ according to the *Gibbs distribution*, i.e., with probability proportional to $\prod_{v \in V_G} \lambda(\omega(v))$. The *Glauber dynamics* is a Markov chain on $\Omega$ which recolors a single vertex at each step, and leaves invariant the Gibbs distribution. We prove that, for each dismantleable $H$ and degree bound $\Delta$, there exist positive constant fugacities on $V_H$ such that the Glauber dynamics has mixing time $O(n^2)$, for all graphs $G$ whose vertex degrees are bounded by $\Delta$.

*This work was supported by EPSRC Research grant "Sharper Analysis of Randomised Algorithms: A Computational Approach" and in part by the ESPRIT Project RAND-APX. The last author was partially supported by a Koshland Scholar award from the Weizmann Institute of Science.

†School of Computing, University of Leeds, Leeds LS2 9JT, United Kingdom. Email: dyer@comp.leeds.ac.uk

‡Laboratory for Foundations of Computer Science, University of Edinburgh, JCMB, The King's Buildings, Edinburgh EH9 3JZ, United Kingdom. Email: mrj@dcs.ed.ac.uk

§Department of Computer Science, University of Chicago, 1100 E 58th Street, Chicago, IL 60637, USA. Email: vigoda@cs.uchicago.edu

# 1 Introduction

Graph homomorphisms provide a natural generalization of many well-studied combinatorial problems, including independent sets and colorings. Our focus is the computational complexity of randomly generating a homomorphism and computing the number of homomorphisms.

We consider an input graph $G = (V_G, E_G)$ with maximum degree $\Delta$ and a constraint graph $H = (V_H, E_H)$, where the latter may have loops. Let $n = |V_G|$, $h = |V_H|$, and let $x \sim y$ denote adjacency of a pair of vertices in $G$ or $H$. Our interest is in the set $\Omega$ of *$H$-colorings* (or homomorphisms) $\sigma : V_G \to V_H$ where $\sigma(v) \sim \sigma(w)$ for all $v \sim w$. Just two of the numerous combinatorial structures that can be expressed as $H$-colorings are illustrated in Figure 1. It is easily verified that when $H = H_{\mathrm{IS}}$ the set $\Omega$ of $H$-colorings of a graph $G$ is in direct correspondence with the set of all independent sets in $G$; while when $H = H_{\mathrm{3Col}}$, the set $\Omega$ consists of all proper 3-colorings of $G$.


Figure 1: Independent sets (left) and proper 3-colorings (right).

Each color $c \in V_H$ is assigned a *fugacity* $\lambda(c)$. For convenience, we extend $\lambda$ to colorings $\sigma \in \Omega$ by defining $\lambda(\sigma) = \prod_{v \in V_G} \lambda(\sigma(v))$. We are interested in the *Gibbs* (or "hard-core") probability distribution $\pi$ over $\Omega$ given by $\pi(\sigma) = \lambda(\sigma)/Z$ for all $\sigma \in \Omega$, where the normalising factor $Z = \sum_{\sigma \in \Omega} \lambda(\sigma)$ is the *partition function* for $H$-colorings of $G$. The complexity of exactly computing $Z$ has been investigated by Dyer and Greenhill [6] and is well-understood; in fact the problem is #P-complete for every non-trivial $H$ even when we restrict attention to low-degree graphs $G$ and uniform fugacity 1.

The picture appears (at the moment) more vague when we consider approximate computation of $Z$, though Dyer, Goldberg, Greenhill and Jerrum [7] discuss the case where all fugacities are equal, and uncover a part of the picture. In this paper, we examine the closely related problem of sampling colorings from the Gibbs distribution, or at least a close approximation to that distribution. We identify a class of constraint graphs $H$ for which sampling can be done in polynomial time, for appropriately chosen non-zero fugacities. Since approximate counting is efficiently reducible to sampling for this class of constraint graphs, we obtain a polynomial-time approximation algorithm (technically an FPRAS) for this class of constraint graphs with their associated fugacities.

The typical approach to efficient sampling is to set up an ergodic Markov chain $(X_t)$ on $\Omega$ whose stationary distribution is the desired distribution $\pi$. By simulating this Markov chain for sufficiently many steps, samples from a distribution arbitrarily close to $\pi$ may be obtained. There are a number of ways of specifying appropriate transition probabilities for this Markov chain; a basic one is provided by the *Glauber dynamics*, which changes one vertex color at a time, according to the following experiment:

G1. Choose a vertex $v \in V_G$ uniformly at random.

G2. Let $X_{t+1}(w) := X_t(w)$ for all $w \neq v$.

G3. Let $S = \{c \in V_H : c \sim X_t(w) \text{ for all } w \sim v\}$ denote the set of valid colors for $v$.

G4. Choose the color $X_{t+1}(v)$ randomly from $S$ with probability proportional to its fugacity.

In what circumstances might the Glauber dynamics provide an effective solution to the problem of sampling $H$-colorings? Certainly we require the state space $\Omega$ to be connected in some sense. A Markov chain with finite state space $\Omega$ and transition matrix $P : \Omega \times \Omega \to [0, 1]$ is called *ergodic* if the following conditions hold:

- *Irreducibility*: for all states (colorings) $\sigma, \tau \in \Omega$, there exists a time $t = t(\sigma, \tau)$ such that $P^t(\sigma, \tau) > 0$;
- *Aperiodicity*: for all states $\sigma$, $\gcd\{t : P^t(\sigma, \sigma) > 0\} = 1$.

It is a classical theorem from stochastic processes that a finite, ergodic Markov chain has a unique stationary distribution.

In many situations it is easy to verify that a candidate distribution is stationary. Specifically, a distribution $\pi'$ on $\Omega$ that satisfies the so-called detailed balance conditions,

$$\pi'(\sigma)\, P(\sigma, \tau) = \pi'(\tau)\, P(\tau, \sigma) \quad \text{for all } \sigma, \tau \in \Omega, \tag{1}$$

is a stationary distribution. It is straightforward to verify that the Gibbs distribution $\pi$ satisfies the detailed balance conditions for the Glauber dynamics on $H$-colorings. Nevertheless, we still need to check that $\pi$ is the *unique* stationary distribution, i.e., that the Glauber dynamics meets the conditions of irreducibility and aperiodicity.

Since the Glauber dynamics satisfies $P(\sigma, \sigma) > 0$ for all colorings $\sigma$ (it is always valid to recolor a vertex with the same color as before) it is immediate that the Markov chain specified by the Glauber dynamics is aperiodic. The question of irreducibility is more subtle. Brightwell and Winkler [1] characterized the constraint graphs $H$ for which the Glauber dynamics on $H$-colorings is irreducible for every input graph $G$. They call such constraint graphs "dismantleable" and prove a host of equivalent conditions for dismantleability.

For efficient sampling, however, it is not sufficient that the Glauber dynamics converges eventually to the stationary (Gibbs) distribution; we need to know that the rate of convergence is rapid. Our main result is that for every dismantleable $H$ and degree bound $\Delta$ there exists a set of non-zero fugacities such that the Glauber dynamics has polynomial mixing time[1] uniformly over graphs $G$ of maximum degree at most $\Delta$. The term *rapidly mixing* is often applied to a Markov chain, such as this one, whose mixing time is polynomial in some natural measure of input size, in this case the order of $G$.

If $H$ is a complete graph with loops on all vertices then the Glauber dynamics is trivially rapidly mixing for any input graph $G$. Cooper, Dyer and Frieze [3] proved that, for any other $H$, there exists a set of fugacities (actually, one may set all fugacities equal to 1) and a sufficiently large degree $\Delta$ such that the Glauber dynamics on $H$-colorings has exponential mixing time, for some infinite family of $\Delta$-regular input graphs $G$. In contrast, we prove here that, provided $H$ is dismantleable, there is another set of fugacities such that the Glauber dynamics on $H$-colorings has polynomial mixing time, on every bounded degree input graph $G$.

[1]A precise definition of mixing time is given in §2, but roughly it is the time $t$ at which the $t$-step distribution of a Markov chain comes sufficiently close to the stationary distribution $\pi$ in $l_1$ distance.

## 2 Definitions and results

In graph-theoretic terms, a constraint graph $H$ is said to be *dismantleable* if there exists an ordering $<$ of $V_H$ for which the following holds:

> Let $c^*$ be the smallest vertex (color) in the ordering $<$. Then, for all $c \in V_H \setminus \{c^*\}$, there exists $p(c) < c$ such that $c' \leq c$ and $c' \sim c$ entail $c' \sim p(c)$.

It is quite easy to show that the above condition is sufficient to ensure irreducibility. Fix an ordering on colors (i.e., vertices of $H$) for which the dismantleability condition holds, with $c^*$ being the least color in the ordering. Starting from an arbitrary coloring $\sigma \in \Omega$, repeating the following procedure reaches the monochromatic coloring $(c^*)^{V_G}$: let $c$ denote the greatest color appearing in $\sigma$; recolor all occurrences of $c$ by $p(c)$. Note that the steps can be reversed to get from $(c^*)^{V_G}$ back to the original coloring.[2] That dismantleability is a necessary condition is a little trickier, and we refer the reader to [1].

We are interested in the asymptotic distance of the Glauber dynamics from stationarity. The traditional measure of distance in this context is (total) variation distance, defined as:

$$d_{\mathrm{TV}}(P^t(\sigma, \cdot), \pi) = \frac{1}{2} \sum_{\tau \in \Omega} |P^t(\sigma, \tau) - \pi(\tau)|,$$

where $\sigma$ is the initial state (coloring). Our focus is the time to get close to stationarity, known as the mixing time. For an initial state $\sigma \in \Omega$ and $0 \leq \varepsilon < 1$, let

$$T_\sigma(\varepsilon) = \min\{t : d_{\mathrm{TV}}(P^t(\sigma, \cdot), \pi) \leq \varepsilon\}.$$

The *mixing time* is defined as

$$T(\varepsilon) = \max_{\sigma \in \Omega} T_\sigma(\varepsilon).$$

For the purposes of the analysis we do not consider the Glauber dynamics directly, but consider instead a variant dynamics with a more restricted set

[2] Incidentally, the procedure just described provides a demonstration that $(c^*)^{V_G}$ is a valid $H$-coloring of $G$, provided *some* valid $H$-coloring exists. Thus, except in some trivial situations, the vertex $c^*$ must be looped in $H$.

of possible transitions. (It is perhaps counterintuitive that the proof of rapid mixing might be simplified by limiting the available transitions!) In brief, we allow only transitions in which a color $c$ is replaced by its "parent" $p(c)$ or vice versa. If we were simply to restrict the available transitions in this way, while retaining the original fugacities $\lambda(c)$, then we would alter the stationary distribution. To correct for this, we assign to each color $c \in V_H$ a weight $\mu(c)$, which is in general different from $\lambda(c)$. The weight function $\mu$ influences transition probabilities just as $\lambda$ does, but we avoid calling $\mu(c)$ a "fugacity", since the relation of the weights $\mu$ to the stationary distribution of the Markov chain is not quite the same as that of the fugacities $\lambda$ to the Gibbs distribution.

Fix an ordering $<$ and parents $p(\cdot)$ consistent with the definition of dismantleability. Then our new Markov chain $(X_t)$ has transitions $X_t \to X_{t+1}$ defined by the following experiment:

V1. Choose a vertex $v \in V_G$ uniformly at random. Let $c$ denote the current color of $v$.

V2. Let $X_{t+1}(w) := X_t(w)$ for all $w \neq v$.

V3. Let $R = R(c) = \{c' \in V_H : c' = p(c),\, c' = c, \text{ or } c = p(c')\}$ denote the set of "relatives" of color $c$.

V4. Choose the color $c'$ randomly from $R$ with probability proportional to its weight $\mu(c')$. Provided it leads to a valid $H$-coloring, set $X_{t+1}(v) := c'$; otherwise set $X_{t+1}(v) := c$.

As before, the variant dynamics is an ergodic Markov chain provided $H$ is dismantleable. The stationary distribution under the variant dynamics has probabilities $\pi(\sigma)$ proportional to $\prod_{v \in V_G} \mu(\sigma(v))\, \mu(R(\sigma(v)))$, where $\mu(R(c)) = \sum_{c' \in R(c)} \mu(c')$. This may be verified from the detailed balance conditions (1). Suppose $\sigma$ and $\tau$ agree at all vertices except $v$, at which $c = \sigma(v)$ and $c' = \tau(v)$. (Note that $\sigma$ and $\tau$ must be of this form if $P(\sigma, \tau)$ is to be non-zero.) Then we have $P(\sigma, \tau) = \mu(c')/\mu(R(c))$ and $P(\tau, \sigma) = \mu(c)/\mu(R(c'))$. Hence

$$\frac{P(\sigma, \tau)}{P(\tau, \sigma)} = \frac{\mu(c')\, \mu(R(c'))}{\mu(c)\, \mu(R(c))} = \frac{\pi(\tau)}{\pi(\sigma)},$$

as required. Thus the stationary distribution is the Gibbs distribution with fugacities $\lambda(c) = \mu(c)\, \mu(R(c))$. We can now state our main theorem.

**Theorem 1** *For every input graph $G = (V_G, E_G)$ with maximum degree $\Delta$ and dismantleable constraint graph $H = (V_H, E_H)$, there exists a set of weights (depending only on $\Delta$ and $h = |V_H|$ ) such that the variant dynamics (as defined in V1–V4 above) has mixing time $O(n \log n)$. Specifically, $T(\varepsilon) = O\big(n(\log n + \log \varepsilon^{-1})\big)$.*

**Corollary 2** *Under the conditions of Theorem 1, there exists a set of fugacities (depending only on $\Delta$ and $h$) such that the Glauber dynamics (as defined in G1–G4 above) has mixing time $O(n^2)$. Specifically, $T(\varepsilon) = O\big(n(n + \log \varepsilon^{-1})\big)$.*

The corollary follows easily from the Diaconis and Saloff-Coste technique [4] for comparing the associated Dirichlet forms of the Markov chains. Indeed a slightly weaker bound on mixing time, with $n \log n + \log \varepsilon^{-1}$ replacing $n + \log \varepsilon^{-1}$, can be obtained simply by substituting the mixing time bound from Theorem 1 into a general comparison theorem of Randall and Tetali [9, Prop. 4]. However, by applying exactly the same method, but working from first principles, we can avoid the factor $\log n$. The argument is this. Applying Sinclair's [10, Prop. 1(ii)] to the mixing time bound of Theorem 1 with $\varepsilon = n^{-1}$, we discover that the spectral gap of the variant dynamics is $\Omega(n^{-1})$. Now the spectral gaps of the Glauber and variant dynamics are the same to within a constant factor: this can be seen by inspecting the variational characterisation of the second largest eigenvalue. Corollary 2 then follows by plugging the $\Omega(n^{-1})$ bound on spectral gap into [10, Prop. 1(i)].

Since the above comparison argument relies only on each possible transition in the variant dynamics being matched (in general with different probability) in the Markov chain under comparison, Corollary 2 holds for other single-site update rules, e.g., Metropolis. We conjecture that the true mixing time here is also $O\big(n(\log n + \log \varepsilon^{-1})\big)$ but the proof of this (if true) appears rather more complex than our proofs of Theorem 1 and Corollary 2.

As mentioned in the introduction, the existence of an efficient sampling procedure for certain structures usually entails the existence of an efficient approximation algorithm for the partition (or generating) function for those structures. The current situation is no exception.

**Corollary 3** *Under the same conditions as Theorem 1, there is a fully polynomial randomized approximation scheme (FPRAS) for the partition function $Z$ of $H$-colorings of $G$.*

For the definition of FPRAS, and also the reduction from approximate counting to sampling required to establish Corollary 3, see for example Jerrum's survey article [8], specifically §2. The reduction is given there in the context of usual (proper) colorings, but it works almost without change for $H$-colorings, when $H$ is dismantleable.[3] To verify the reduction in the current context it is necessary to check that the ratios $\rho_i$ appearing in that reduction are bounded away from 0. This can be done using a straightforward extension of the argument used to prove ergodicity of the Glauber dynamics.

Finally, note that many constraint graphs $H$ are covered both by Theorem 1 and by the result of Cooper et al. [3]. In other words, there are graphs $H$ — the simplest being $K_2$ with a single loop, which corresponds to independent sets in the input graph $G$ — for which the mixing time of the Glauber dynamics is either polynomial or exponential, depending on the fugacities.

# 3 Coupling

We prove the main theorem via coupling. A coupling of a Markov chain is a joint evolution of two copies of the chain, designed to minimize the time till the copies coalesce. In order to be a valid coupling we need that individually each copy behaves faithfully, and once the pair coalesce they evolve together. To be precise, a (Markovian) coupling is a Markov chain $P'$ on $\Omega \times \Omega$ which satisfies the following conditions:

$$\sum_{\tau' \in \Omega} P'((\sigma,\tau),(\sigma',\tau')) = P(\sigma,\sigma') \qquad \text{for all } \sigma,\sigma',\tau \in \Omega;$$
$$\sum_{\sigma' \in \Omega} P'((\sigma,\tau),(\sigma',\tau')) = P(\tau,\tau') \qquad \text{for all } \sigma,\tau,\tau' \in \Omega;$$

and

$$P'((\sigma,\sigma),(\sigma',\sigma')) = P(\sigma,\sigma') \qquad \text{for all } \sigma,\sigma' \in \Omega.$$

Our goal is to define a coupling which, in expectation, makes progress with respect to an appropriately defined distance metric after every coupled transition. Defining and analyzing a coupling for an arbitrary pair of states

[3] It is in fact possible to reduce approximate counting to sampling without restriction on the graph $H$, but the reduction is then substantially more involved and, crucially in the context of Corollary 3, does not preserve the bound $\Delta$ on vertex degrees [5].

is typically a complicated task; however, the path coupling lemma of Bubley and Dyer simplifies matters. In particular, it suffices to focus on colorings which differ by a single transition of the dynamics, referred to as *adjacent* colorings. We write $\sigma \sim_v \tau$ to indicate that the pair of colorings $\sigma, \tau$ differ only at vertex $v$, and $\sigma \sim \tau$ if $\sigma \sim_v \tau$ for some vertex $v$.

We analyze this set of adjacent colorings with respect to the following distance metric. Each color $c \in V_H$ will be assigned a distance weight $d(c)$. For colorings $\sigma \sim_v \tau$, where without loss of generality $\tau(v) = p(\sigma(v))$, we will assign distance $d(\sigma, \tau) = d(\sigma(v))$. For any $\sigma, \tau \in \Omega$, let $\rho(\sigma, \tau)$ denote the collection of paths $\eta$ such that $\sigma = \eta_0 \sim \eta_1 \sim \cdots \sim \eta_\ell = \tau$, where $\ell(\eta)$ is the length (in terms of number of transitions) of $\eta$. We define the distance between an arbitrary pair of states as the total distance along a shortest path:

$$d(\sigma, \tau) = \min_{\eta \in \rho(\sigma,\tau)} \sum_{0 \leq i < \ell(\eta)} d(\eta_i, \eta_{i+1}).$$

For a pair of colorings $\sigma, \tau$ and a coupling $(\sigma, \tau)$, let $\sigma', \tau'$ denote the resulting pair of colorings after the coupled transition. The specialization of the path coupling lemma to our setting is as follows.

**Lemma 4 (Bubley and Dyer [2])** *If there exists a $\beta < 1$ and a coupling for all $\sigma \sim \tau$ such that*

$$\mathbb{E}[d(\sigma', \tau')] < \beta\, d(\sigma, \tau),$$

*then the mixing time is bounded by*

$$T(\varepsilon) \leq \frac{\log(D/\varepsilon)}{1 - \beta},$$

*where $D = \max_{\sigma, \hat{\sigma} \in \Omega} d(\sigma, \hat{\sigma})$.*

In our application, we shall discover that $\beta = 1 - \Theta(1/n)$ and $D = \Theta(n)$, leading easily to Theorem 1.

# 4 Proof of Theorem 1

Fix an ordering on the colors $c_1 < c_2 < \cdots < c_h$ for which the dismantleability condition holds. Set $d(c_i) = 2(\Delta + 2)^{i+1}$ and let $d_{\max} = d(c_h)$. Now set

$\mu_i = \mu(c_i) = (d_{\max}+1)^{-i}$, for $i = 1, \dots, h$. Observe that the inequality

$$\frac{\sum_{k>i} \mu_k}{\mu_i} < \frac{1}{d_{\max}}$$

holds for all $i$.

Fix a pair of colorings $\sigma \sim_v \tau$ and once again let $\sigma', \tau'$ denote the resulting colorings after our coupled transition. To prove the theorem we need to demonstrate a coupling for which $\mathbb{E}[d(\sigma',\tau')] - d(\sigma,\tau)$ is negative and bounded away from zero.

We couple the two chains so that both chains attempt to modify the same vertex at every step. Therefore, for a vertex $x$, it is well defined to let

$$\mathbb{E}_x[d(\sigma',\tau')] = \mathbb{E}\left[d(\sigma',\tau') \;\middle|\; \text{vertex } x \text{ is selected by the coupled process}\right].$$

Observe that the distance metric does not change when we recolor a vertex sufficiently far from $v$. Specifically, we only need to consider recolorings of $v$ or a neighbor $w$ of $v$. In summary, we have

$$\begin{aligned} n\big(\mathbb{E}[d(\sigma',\tau')] - d(\sigma,\tau)\big) &= \big(\mathbb{E}_v[d(\sigma',\tau')] - d(\sigma,\tau)\big) \\ &\quad + \sum_{w:w\sim v} \big(\mathbb{E}_w[d(\sigma',\tau')] - d(\sigma,\tau)\big). \end{aligned} \tag{2}$$

Of all the colors that are valid for $\sigma'(x)$, let $c^*_\sigma(x)$ denote the one of greatest weight. Note that either $c^*_\sigma(x) = \sigma(x)$ or $c^*_\sigma(x) = p(\sigma(x))$, and that, in the latter case, $c^*_\sigma(x)$ has the greatest weight among the various colors that may be proposed. Either way, if vertex $x$ is selected, then $x$ will acquire color $c^*_\sigma(x)$ unless a color of lower weight than $c^*_\sigma(x)$ is proposed. Thus, conditioned on vertex $x$ being selected, we have

$$\Pr[\sigma'(x) \neq c^*_\sigma(x)] \leq \frac{\sum_{c>c^*_\sigma(x)} \mu(c)}{\mu(c^*_\sigma(x))} < \frac{1}{d_{\max}}. \tag{3}$$

Defining $c^*_\tau(x)$ similarly, the analogous inequality holds with respect to $\tau$ as well.

In the light of inequality (3) and its analogue for $\tau$ we may complete the definition of the coupled process as follows. With probability $1 - 1/d_{\max}$ we couple $\sigma'(x) = c^*_\sigma(x)$ with $\tau'(x) = c^*_\tau(x)$. We call such a coupled move a *type A transition*. With the remaining probability of $1/d_{\max}$ the two chains independently recolor $x$ (using the residual distributions), a *type B transition*. In the latter case, it is clear that $d(\sigma', \sigma) \leq d_{\max}$ and $d(\tau', \tau) \leq d_{\max}$ which implies

$$\mathbb{E}_x[d(\sigma', \tau')|\text{ type B transition}] \leq 2d_{\max} + d(\sigma, \tau). \tag{4}$$

We now proceed to bound the expected change in distance from a type A transition. It is clear that the distance after the type A transition on $x$ is maximized when $c^*_\sigma(x) \neq c^*_\tau(x)$.

Recall that, by convention, $\tau(v) = p(\sigma(v))$, rather than vice versa. Consider the recoloring of a neighbor $w$ of $v$. We begin by proving that $c^*_\sigma(w) \neq c^*_\tau(w)$ implies $\sigma(w) < \sigma(v)$. Suppose $\sigma(w) \geq \sigma(v)$. Vertex $w$ has the same color in both chains, therefore colors $c^*_\sigma(w)$ and $c^*_\tau(w)$ are either $\sigma(w)$ or $p(\sigma(w))$. Since $\sigma(w) \sim \sigma(v)$ and $\sigma(w) \sim \tau(v)$, from the definition of dismantleability and our ordering on $V_H$ we know $p(\sigma(w)) \sim \sigma(v)$ and $p(\sigma(w)) \sim \tau(v)$. Similarly if $u \neq v$ is any other neighbor of $w$, then dismantleability implies $p(\sigma(w)) \sim \sigma(u) = \tau(u)$. Therefore, recoloring vertex $w$ to color $p(\sigma(w))$ is valid in both chains or neither and hence $c^*_\sigma(w) = c^*_\tau(w)$.

Now suppose $c^*_\sigma(w) \neq c^*_\tau(w)$. In addition to the fact $\sigma(w) < \sigma(v)$, it is clear that $c^*_\sigma(w) = \sigma(w)$ and $c^*_\tau(w) = p(\sigma(w))$. Following the type A transition for $w$ we have $\sigma' = \sigma$ and $d(\tau', \tau) = d(\sigma(w)) \leq d(\sigma(v))/(\Delta + 2)$.

In summary, we have proven

$$\mathbb{E}_w[d(\sigma', \tau')|\text{ type A transition}] \leq d(\sigma, \tau) \left[1 + \frac{1}{\Delta + 2}\right]. \tag{5}$$

Combining (4) and (5) — recalling that a type B transition occurs with probability at most $1/d_{\max}$ — yields

$$\mathbb{E}_w[d(\sigma', \tau')] - d(\sigma, \tau) \leq 2 + \frac{d(\sigma, \tau)}{\Delta + 2}. \tag{6}$$

We complete the proof by considering the effect of recoloring $v$ by a type A transition. Since $p(\sigma(v)) = \tau(v)$, it is clear that $p(\sigma(v))$ is a valid color for both $\sigma'(v)$ and $\tau'(v)$. Moreover, $c^*_\sigma(v) = p(\sigma(v))$ and $c^*_\tau(v)$ is either $p(\sigma(v))$ or $p(p(\sigma(v)))$. In other words, $d(\sigma', \tau') > 0$ implies $c^*_\tau(v) = p(p(\sigma(v)))$. In which case we have $d(\sigma', \tau') \leq d(p(\sigma(v))) \leq d(\sigma, \tau)/(\Delta + 2)$. Restating our bound:

$$\mathbb{E}_v[d(\sigma', \tau')|\text{ type A transition}] \leq \frac{d(\sigma, \tau)}{\Delta + 2}.$$

The above inequality together with (4) implies

$$\begin{aligned}\mathbb{E}_v[d(\sigma',\tau')] - d(\sigma,\tau) &\leq 2 + d(\sigma,\tau)\left[\frac{1}{d_{\max}} + \frac{1}{\Delta+2} - 1\right] \\ &\leq 3 - d(\sigma,\tau)\left[1 - \frac{1}{\Delta+2}\right]. \qquad (7)\end{aligned}$$

Let $\Delta(v)$ denote the degree of vertex $v$. Putting inequalities (6) and (7) into (2), and recalling $d(\sigma,\tau) = d(\sigma(v))$, we obtain

$$\begin{aligned}n\big(\mathbb{E}[d(\sigma',\tau')] - d(\sigma,\tau)\big) &\leq \Delta(v)\left[2 + \frac{d(\sigma,\tau)}{\Delta+2}\right] + \left[3 - \frac{(\Delta+1)\,d(\sigma,\tau)}{\Delta+2}\right] \\ &\leq (2\Delta+3) - \frac{d(\sigma(v))}{\Delta+2} \\ &\leq (2\Delta+3) - 2(\Delta+2) \\ &= -1.\end{aligned}$$

Since $\sigma \sim \tau$ implies $d(\sigma,\tau) \leq d_{\max}$, we conclude

$$\mathbb{E}[d(\sigma',\tau')] \leq (1 - 1/nd_{\max})\, d(\sigma,\tau),$$

and hence we may take $\beta = 1 - 1/nd_{\max}$ in Lemma 4. If $\hat{\sigma} = (c_1)^{V_G}$, we know that $d(\sigma,\hat{\sigma}) \leq nhd_{\max}$ for all $\sigma \in \Omega$, by the path of transitions described in §2. Therefore, we may take $D = 2nhd_{\max}$ in Lemma 4. Theorem 1 now follows.

# References


[1] G. R. Brightwell and P. Winkler. Gibbs measures and dismantleable graphs. *J. Combin. Theory Ser. B*, 78(1):141–166, 2000.

[2] R. Bubley and M. Dyer. Path coupling, Dobrushin uniqueness, and approximate counting. In *38th Annual Symposium on Foundations of Computer Science*, pages 223–231, Miami Beach, FL, October 1997. IEEE.

[3] C. Cooper, M. Dyer, and A. Frieze. On Markov chains for randomly $H$-coloring a graph. *J. Algorithms*, 39(1):117–134, 2001.

[4] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *The Annals of Applied Probability*, 3(3):696–730, 1993.

[5] M. Dyer, L.A. Goldberg and M. Jerrum. Counting and Sampling H-colourings. To appear in *Proceedings of the 6th. International Workshop on Randomization and Approximation Techniques in Computer Science*, Springer-Verlag LNCS, September 2002.

[6] M. Dyer and C. Greenhill. The complexity of counting graph homomorphisms. *Random Structures Algorithms*, 17(3-4):260–289, 2000.

[7] M. E. Dyer, L. A. Goldberg, C. S. Greenhill, and M. R. Jerrum. On the relative complexity of approximate counting problems. In *Proceedings of APPROX 2000*, Lecture Notes in Computer Science vol. 1913, pages 108–119, Springer Verlag, 2000.

[8] M. Jerrum. Mathematical foundations of the Markov chain Monte Carlo method. In *Probabilistic Methods for Algorithmic Discrete Mathematics* (M. Habib, C. McDiarmid, J. Ramirez-Alfonsin & B. Reed, eds), Algorithms and Combinatorics vol. 16, Springer-Verlag, 1998, 116–165.

[9] D. Randall and P. Tetali. Analyzing Glauber dynamics by comparison of Markov chains. *J. Math. Phys.*, 41(3):1598–1615, 2000.

[10] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1(4):351–370, 1992.