# Very rapidly mixing Markov Chains for $2\Delta$-colourings and for independent sets in a graph with maximum degree 4.

Michael Molloy*
Department of Computer Science
University of Toronto
Toronto, Canada

October 3, 2000

**Abstract**

We introduce a new technique for analyzing the mixing rate of Markov chains. We use it to prove that the Glauber dynamics on $2\Delta$-colourings of a graph with maximum degree $\Delta$ mixes in $O(n \log n)$ time. We prove the same mixing rate for the Insert/Delete/Drag chain of [5] on independent sets of graphs with maximum degree 4.

In this paper, we introduce a simple idea for analyzing the mixing rate of Markov chains, and give two applications: (i) to the Glauber dynamics on graph colourings, and (ii) to Dyer and Greenhill's Insert/Delete/Drag chain on independent sets[5].

In a basic coupling argument, one shows that for any two initial states $X, Y$ of a particular Markov chain, one can couple a walk starting with $X$ to a walk starting with $Y$ such that with high probability these two walks coincide within a short period of time. The most popular way to do this is to show that the expected change in the "distance" between $X, Y$ after one step is at most $t \leq 0$. If $t$ is strictly less than 0, then with high probability

*This work is supported by an NSERC Research Grant.

the distance drops to 0, i.e. the walks coincide, within $O(-\frac{1}{t}\log D)$ where $D$ is the maximum possible distance between $X$ and $Y$ (here, $D$ is non-negative and integer-valued). If, say, $t = -\theta(\frac{1}{n})$ and $D = O(n)$, where $n$ is some measure of the size of the problem (eg. the number of vertices in the underlying graph), this implies that the mixing rate of the Markov chain is $O(n \log n)$, which in many cases is known to be the best possible order of magnitude. If $t = 0$ then as long as the distance can be shown to change frequently enough, then one can usually show that the mixing rate is bounded by a (typically larger) polynomial. (See [7] or [11] for a more thorough description.)

Our simple idea is this: Consider a chain for which in the worse case the expected change in $\mathrm{dist}(X, Y)$ after 1 step is $t$. We can sometimes show that with positive probability, within a few steps of our coupled walks we will move to a pair of states for which the expected change in distance after the next step is at most $t^{'} < t$. This implies that, in the worse case, the overall expected change in distance after a few steps is at most $t^{*}$ for some $t^{'} < t^{*} < t$. So if $t > 0$ then so long as $t^{*} \leq 0$ we still have rapid mixing. Or if $t = 0$ then if $t^{*} = -\theta(\frac{1}{n})$ then we show that the mixing rate is only $O(n \log n)$.

In this paper we will be using the idea of *path coupling*, introduced by Bubley and Dyer [2]. Using this method, one only needs to directly analyse worse case bounds for a subset of the pairs of states - typically pairs at distance 1. Path coupling then infers the same worse case bound on all pairs.

Our first application of our idea is to a Markov chain whose states are the set of $C$-colourings of a graph $G$ with maximum degree $\Delta$. The simplest commonly analysed such chain is commonly referred to in the statistical physics community as the *Glauber dynamics*. In each step, we choose a vertex $v \in G$ and a colour $c \in \{1, ..., C\}$ uniformly at random. If $c$ does not appear on a neighbour of $v$, then we recolour $v$ with $c$. It is easy to prove that this chain is ergodic and that its stationary distribution is the uniform distribution.

Standard (and easy) bounds on the Coupon Collectors Problem show that it takes at least $\theta(n \log n)$ time for the Glauber dynamics to meet every vertex at least once, and that this fact can be used to show that the mixing rate is at least $\theta(n \log n)$ (see, eg., [4]). Jerrum [10] and, independently, Salas and Sokal [13] showed that when $C > 2\Delta$ this chain mixes in $O(n \log n)$ time, and that when $C = 2\Delta$ it mixes in polytime. Until recently, the best known bound for the case $C = 2\Delta$ was $O(n^3)$ time. The best way to prove this is to apply path coupling and to show that for $C = 2\Delta$, the worse-case expected change in the distance between two coupled states is 0. This makes this example a good candidate for our idea, and we use it to show:

**Theorem 1** *For any fixed $\Delta \geq 2$, the Glauber dynamics on $2\Delta$-colourings of a graph with maximum degree $\Delta$ and with $n$ vertices mixes in $O(n \log n)$ time.*

In [4], Dyer et al prove that for triangle-free graphs, the Glauber dyamics mixes in time $O(n \log n)$ for $C \geq (2-\epsilon)\Delta$, where $\epsilon$ is a small positive constant. In [8] we show how to combine the ideas from this paper and from [4] to remove the "triangle-free" condition from their result. It is worth noting that Dyer and Greenhill [6] showed that a different chain with a uniform stationary distribution mixes in $O(n \log n)$ time for $C = 2\Delta$, and more recently Vigoda [14] showed that a different such chain mixes in time $O(n \log n)$ for $C \geq \frac{11}{6}\Delta$. Using comparison techniques, Vigoda showed that his result implies that the Glauber dynamics mixes in time $O(n^2 \log n)$ for the same values of $C$.

With a little more care, it is not difficult to strengthen the proof of this paper to prove the same mixing rate when $\Delta$ grows with $n$. For the sake of simplicity, we do not provide the details here, as both of these results are implied by the work in [8].

Our next example has to do with sampling independent sets in a graph. A very natural chain to do this is the so-called Insert/Delete chain: At each step we choose a vertex $v$ uniformly at random. If $v$ is in our independent set, $I$, then we delete it with probability $\frac{1}{2}$; if $v \notin I$ and if $v$ does not have a neighbour in $I$, then we insert $v$ with probability $\frac{1}{2}$. A straightforward path-coupling analysis shows that this chain mixes in polytime for (the relatively uninteresting case) $\Delta \leq 2$. Luby and Vigoda [12] and, independently, Dyer and Greenhill [5], introduced variants on this chain and showed that it mixes in polytime for $\Delta \leq 4$.

The variant introduced by Dyer and Greenhill is referred to as the Insert/Delete/Drag chain (IDD) and it runs like Insert/Delete with the following operation: If $v \notin I$ and if $v$ has exactly one neighbour $w \in I$ then with probability $\frac{1}{8}$ we delete $w$ and insert $v$. Again, it is straightforward to prove that this chain is ergodic and has the uniform stationary distribution. (The chain introduced by Luby and Vigoda is very similar.)

The best bound on the running time of this, or any other chain, for graphs with maximum degree 4 is $O(n^2 \log n)$, where $n$ is the number of vertices in the graph. The best way to prove this is through path coupling, whereby one proves that the worse-case expected change in the distance between two coupled states is 0. Again, our idea can be applied to prove:

**Theorem 2** *The IDD chain on the independent sets of a graph with maximum degree 4 and with $n$ vertices mixes in $O(n \log n)$ time.*

At this point it is not known whether any rapidly mixing Markov chain samples independent sets uniformly at random for graphs with maximum degree 5. Dyer, Frieze and Jerrum [9] showed that there is no such chain for graphs on $n$ vertices and with maximum degree $\Delta \geq 6$, which changes the independent set by at most $\epsilon n$ vertices, for some small constant $\epsilon > 0$. Furthermore, there is no such chain for $\Delta \geq 25$ unless $RP = NP$.

There is a simple variation of IDD whose uniform distribution is a hard-core distribution where each independent set $I$ has probability of being selected proportional to $\lambda^{|I|}$. Dyer and Greenhill [5] showed that as long as $\Delta \leq \frac{2}{\lambda} + 2$, this chain mixes in polytime (Luby and Vigoda[12] obtain a similar bound on their chain). The proof of Theorem 2 can be easily modified to prove rapid mixing so long as $\Delta \leq \frac{2}{\lambda} + 2 + \frac{\epsilon}{\Delta^6}$ for some small constant $\epsilon > 0$. With more care, the argument can be improved to yield rapid mixing so long as $\Delta \leq \frac{2}{\lambda} + 2 + \epsilon$.

Throughout this paper, we will use $N_v$ to denote the open neighbourhood of $v$; i.e., the set of vertices $u \neq v$ such that $u$ is adjacent to $v$. For a set of vertices $S$, we use $N(S)$ to denote the neighbourhood of $S$, i.e. $(\cup_{v \in S} N_v) - S$, and we use $\deg_S(v)$ to denote the degree in $S$ of $v$, i.e. $|N_v \cap S|$.

# 1 Colourings

In this section, we prove Theorem 1. We will apply path coupling, introduced in [2], using $H(X, Y)$, the Hamming distance between two colourings $X, Y$, i.e. the number of vertices on which they differ. We will also use what is essentially the *path coupling with stopping time* approach introduced by Dyer et al in [4]. This technique is similar in spirit to *delayed path coupling* introduced by Czumaj et al in [3].

First, we will specify what we mean by the "mixing time" of a Markov chain. We use $\pi$ to represent the stationary distribution of the chain, i.e. $\pi(x)$ is the limiting distribution of the probability that the chain is in the state $x$. For any initial state $x_0$ and integer $t > 0$ we use $P^t(x_0)$ to denote the distribution of the chain at step $t$, i.e. $P^t(x_0, x)$ is the probability that, starting at $x_0$, our chain is in the state $x$ at step $t$. The *total variation distance* between $\pi$ and $P^t(x_0)$ is

$$d_{\mathrm{TV}}(P^t(x_0), \pi) = \frac{1}{2} \sum_{x \in \Omega} |P^t(x_0, x) - \pi(x)|,$$

where $\Omega$ is the state set of the chain. We say that a Markov chain is *rapidly mixing* if there is some $T$ which is a polynomial function of $n, \log(1/\epsilon)$ such that for every initial state $x_0$, $d_{\mathrm{TV}}(P^t(x_0), \pi) \leq \epsilon$ for every $t \geq T$. (Of course, $n$ is some measure of the size of the states of the chain; in this paper, it is the number of vertices in the underlying graph.) The polynomial $T$ is the *mixing time*. Often we omit $\epsilon$ from the running time, and so we state the running time as a function of $n$, where it is understood that multiplying by a polynomial in $\log(1/\epsilon)$ will yield a suitable function $T$.

A coupling of a Markov chain is a stochastic process $(X_t, Y_t)$ on $\Omega \times \Omega$ such that each of the marginal processes $(X_t), (Y_t)$ is a faithful copy of the chain. A coupling can be thought of as a pair of processes which somehow share the random choices which drive them. Concrete examples of coupling will appear later in this paper. When we design a coupling, we usually do so in a manner which makes it likely that the two processes will soon coincide, and once they coincide they remain equal. This yields a bound on the mixing time of the chain, using the following fact, which (essentially) first appeared in Aldous[1]:

**Fact 3** *Suppose that for all initial states $X_0 = x_0, Y_0 = y_0$, we have $\mathbf{Pr}(X_t \neq Y_t) \leq p$ then for any initial state $X_0 = x_0$, we have:*

$$d_{\mathrm{TV}}(P^t(x_0), \pi) \leq p.$$

We will explain path coupling soon. In order to do so, we will have to expand the state space of the Glauber dynamics to include non-proper colourings, i.e. colourings in which two adjacent vertices might have the same colour. Note that if the Glauber dynamics begins with a proper colouring, it will never generate a non-proper colouring, so we are only introducing the possibility of non-proper colourings to the analysis, not to the range of chains which begin with a proper colouring.

Consider any graph $G$ of maximum degree $\Delta$, and any two (not neccessarily proper) $2\Delta$-colourings $X, Y$ of $G$ which differ on exactly one vertex $w$. Consider running a single iteration of the Glauber dynamics on $X, Y$ under the following coupling:

Pick a uniformly random pair $(v, c)$ for $X$. If $v$ is a neighbour of $w$ and if $c = X(w)$ then choose the pair $(v, Y(w))$ for $Y$. Similarly, if $v$ is a neighbour of $w$ and if $c = Y(w)$ then choose the pair $(v, X(w))$ for $Y$. Otherwise, choose $(v, c)$ for $Y$.

*Path coupling* is a technique which can be used to extend this coupling (and its analysis) to the case where $X, Y$ differ in more than one vertex. More precisely, setting $t = H(X, Y)$, we take an arbitrary sequence of (not neccessarily proper) $2\Delta$-colourings $X = X^0, X^1, ..., X^t = Y$ such that each pair $(X^j, X^{j+1})$ differs on exactly one vertex and couple each pair $(X^j, X^{j+1})$ in the same manner as we coupled $(X, Y)$ above. Note that, given $X, Y$, it is easy to construct such a sequence; for example, if $X, Y$ differ on $v_1, ..., v_t$ then we can obtain $X^j$ from $X^{j-1}$ by changing the colour of $v_j$ from $X(v_j)$ to $Y(v_j)$. Note also that in order to do so, it is important to allow the chain to include some non-proper colourings.

**Claim 4** *For any two (not neccessarily proper) $2\Delta$-colourings $X, Y$ the expected change in $H(X, Y)$ after one step of the Glauber dynamics, under the coupling described above, is at most 0.*

The proof appears in several other places, but we include it here because it is short and the ideas in it will be used later in this paper.

**Proof** First we consider the case where $X, Y$ differ on precisely one vertex $w$. Suppose that $X(w) = R$ and $Y(w) = G$. If the dynamic chooses for $X$ the vertex $w$ and any of the at least $\Delta$ colours not appearing on $N(w)$ then $H(X, Y)$ will decrease to 0. If the dynamic chooses for $X$ any neighbour $u$ of $w$ and the colour $G$ then it will pick $u, R$ for $Y$ and possibly $H(X, Y)$ will increase by 1. No other choice can change $H(X, Y)$. So there are at least $\Delta$ possible choices which will decrease $H$ and at most $\Delta$ choices which will increase $H$. Since each choice is equally likely, the expected change in $H$ is at most 0.

For general $X, Y$ we consider the sequence of colourings $X = X^0, X^1, ..., X^t = Y$ which is being used to induce the coupling on $(X, Y)$. Since each pair $(X^j, X^{j+1})$ differs on exactly one vertex, the expected change in $H(X^j, Y^j)$ is at most 0. The expected change in $H(X, Y)$ is at most the sum of the expected changes in $H(X^j, Y^j)$, and so is also at most 0. $\square$

As we described earlier, Claim 3 will yield a polynomial bound on the mixing time (in fact, it yields $O(n^3 \log n)$), but to obtain $O(n \log n)$ we must bound the expected change in $H(X, Y)$ away from 0. We set $(X_0, Y_0) = (X, Y)$, and we set $(X_i, Y_i)$ to be equal to $(X, Y)$ after the $i$th step of the Glauber dynamics. We will couple $X$ and $Y$ using a particular coupling rule $\mathcal{C}$. It will be more convenient to delay the formal definition of $\mathcal{C}$ until later. Our main lemma is the following:

**Lemma 5** *There exists a constant $K$, such that for any initial unequal pair $(X, Y)$, coupled by $\mathcal{C}$,* $\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - H(X, Y)) \leq -\frac{1}{5000\Delta^6} \times H(X, Y)$.

Theorem 1 follows from Lemma 5 via the "delayed path coupling lemma" from Czumaj et al[3]:

**Lemma 6** *Let $S \subseteq \Omega \times \Omega$ be such that the transitive closure of $S$ is the whole of $\Omega \times \Omega$. Let $\rho$ be an integer-valued metric on $\Omega \times \Omega$ which takes values in $\{0, \ldots, D\}$. Given $(X_0, Y_0) \in S$, let $(X_0, Y_0)$, $(X_1, Y_1), \ldots, (X_t, Y_t)$ be the $t$-step evolution of a (possibly non-Markovian) coupling starting from $(X_0, Y_0)$. Suppose that there exists a constant $\gamma$ such that $0 < \gamma < 1$ and*

$$\mathbf{Exp}\left[\rho(X_t, Y_t)\right] \leq \gamma \rho(X_0, Y_0) \tag{1}$$

*for all $(X_0, Y_0) \in S$. Then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies*

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

For completeness, and because the implication is very short, we show directly that Lemma 5 implies Theorem 1, instead of going through Lemma 6.

**Proof of Theorem 1:** Consider any pair of colourings $X, Y$, and set $L = 5000\Delta^6 \log(n/\epsilon) \times Kn$. Since $H(X, Y) \leq n$, repeated applications of Lemma 5 imply that

$$\mathbf{Exp}(H(X_L, Y_L)) \leq n \times (1 - \tfrac{1}{5000\Delta^6})^{5000\Delta^6 \log(n/\epsilon)} < \epsilon.$$

The theorem now follows from Markov's Inequality, i.e. the fact that since $H(X_L, Y_L)$ is a non-negative integer, $\mathbf{Pr}(H(X_L, Y_L) \neq 0) \leq \mathbf{Exp}(H(X_L, Y_L)) < \epsilon$, and from Fact 3. $\square$

To prove Lemma 5, we need a few definitions. We are mostly interested in steps of the process where $H(X, Y)$ changes. It will simplify the analysis to focus on a (somewhat larger set of) steps where the dynamics chooses a vertex which *might* (depending on the choice of colour) cause $H(X, Y)$ to increase. We will label the time steps where such a vertex is chosen as $i_1, i_2, \ldots$; specifically, we define $i_j$ to be the $j$th index $i$ for which the dynamics chooses for $X$ either a vertex on which $X, Y$ disagree, or a neighbour of such a vertex. $\mathcal{C}$ will be defined so that $H(X, Y)$ doesn't change on any other steps. If $X_{i_j} = Y_{i_j}$ then $i_{j'} = \infty$ for any $j' > j$. Of course, the values of $i_1, i_2, \ldots$ are randomly determined by the process.

We define $Z_0 = 1$ and $Z_j = H(X_{i_j}, Y_{i_j})$ (if $X_{i_j} = Y_{i_j}$ then $Z_{j'} = 0$ for any $j' > j$). Thus, the sequence $Z_0, Z_1, \ldots$, includes every value that $H(X, Y)$ takes on. We define $E_j$ to be the conditional expected value of $Z_j - Z_{j-1}$ conditioned on the pair $X_{i_{j-1}}, Y_{i_{j-1}}$. Thus, $E_j$ is a function of $X_{i_{j-1}}, Y_{i_{j-1}}$. Claim 4 implies that $E_j$ is always at most 0, for any $X_{i_{j-1}}, Y_{i_{j-1}}$. Our main step is to show that with non-negligible probability, at least one $E_j$ is less than 0. In particular, we define $A_j$ to be the event that $E_j \leq -\frac{1}{12\Delta^2}$. When we discuss $A_j$, the underlying random choices are those made throughout the entire process. Thus, $\mathbf{Pr}(A_j)$ is a function of the initial colourings $X, Y$ and the graph $G$. The same is true of any other probabilities discussed in this paper, although, of course, when they are conditional on some property of the process, then that property also contributes to the conditional value of the probability. Since the value of $E_j$ is determined entirely by $X_{i_{j-1}}, Y_{i_{j-1}}$, $\mathbf{Pr}(A_j)$ is the probability that the process generates some $X_{i_{j-1}}, Y_{i_{j-1}}$ such that $E_j \leq -\frac{1}{12\Delta^2}$.

**Lemma 7** *If $(X, Y)$ initially differ on exactly one vertex $w$, then either $A_1$ holds or $\mathbf{Pr}(A_i) > \frac{1}{128\Delta^4}$ for some $i \in \{2, 3\}$.*

**Proof** Initially, $X(w) = R, Y(w) = G$. Until step $i_1$, $\mathcal{C}$ is defined as in Claim 4; i.e. if we choose $(v, c)$ for $X$ then we also choose $(v, c)$ for $Y$, unless $v$ is a neighbour of $w$ and $c \in \{R, G\}$. In this latter case, we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$.

We consider 2 cases for the neighbourhood of $w$.

*Case 1: $w$ has two adjacent neighbours $a, b$.* Much of the focus of our argument will be on the case where step $i_1$ of our dynamics chooses for $X$ the vertex $a$ and the colour $G$, and so for $Y$ chooses $a$ and $R$. There are a few situations for which such a choice will not cause the Hamming distance to increase, namely if in $X_{i_1-1}$ either:

- $a$ has a neighbour coloured $G$ and a neighbour other than $w$ coloured $R$,
- $a$ is coloured $R$ and has a neighbour coloured $G$, or
- $a$ is coloured $G$ and has a neighbour other than $w$ coloured $R$.

We define $B_1$ to be the event that one of these three possibilities holds for $X_{i_1-1}$. Note that the vertex in $w \cup N(w)$ and colour in $\{1, ..., 2\Delta\}$ chosen for step $i_1$ is independent of $B_1$, i.e. for any possible pair $(v, c)$ the probability, conditional on $B_1$ holding (or, for that matter, on any property of $X_{i_1-1}, Y_{i_1-1}$), of $(v, c)$ being chosen is exactly $\frac{1}{2\Delta|w \cup N(w)|} \geq \frac{1}{2\Delta(\Delta+1)}$.

If $B_1$ holds then there are at most $\Delta - 1$ choices by the Glauber dynamics which can increase $H(X, Y)$. From this, it is straightforward to use the argument in Claim 4 to prove that $\mathbf{Exp}(Z_1 - Z_0|B_1) \leq -\frac{1}{2\Delta(\Delta+1)}$. This implies that if $\mathbf{Pr}(B_1) \geq \frac{1}{2}$, then $E_1 \leq -\frac{1}{4\Delta(\Delta+1)}$ and so $A_1$ holds.

Suppose then that $\mathbf{Pr}(B_1) < \frac{1}{2}$.

With probability at least $\frac{1}{2\Delta(\Delta+1)}$, step $i_1$ of our dynamics chooses $(a, G)$ for $X$ and thus chooses $(a, R)$ for $Y$. Suppose that $X_{i_1-1}(a) = Y_{i_1-1}(a) = T$. If $B_1$ does not hold then there are only 3 possibilities for the new colourings:

(i) $X(w) = R, Y(w) = G; X(a) = G, Y(a) = T$.
This occurs if $a$ has a neighbour other than $w$ coloured $R$.

(ii) $X(w) = R, Y(w) = G; X(a) = T, Y(a) = R$.
This occurs if $a$ has a neighbour other than $w$ coloured $G$.

(iii) $X(w) = R, Y(w) = G; X(a) = G, Y(a) = R$.
This occurs if no neighbours of $a$ other than $w$ are coloured R or G,

Note that these situations are not disjoint, since possibly $T = R$ or $G$. In either of these situations, a straightforward count will show that $E_2 \leq -\frac{1}{2\Delta^2}$, as we shall now see:

We start by considering situation (i) where $T \neq R, G$ (the situation where $T = R$ is identical to (iii), and the situation where $T = G$ cannot hold unless $B_1$ holds). If we are in this situation, then until step $i_2$ our coupling $\mathcal{C}$ is defined as follows. If we choose $(v, c)$ for $X$ then we choose $(v, c)$ for $Y$, unless:

- $v \in N_w - N_a - a$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$;
- $v \in N_a - N_w - w$ and $c \in \{G, T\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{G, T\}$ is not equal to $c$;
- $v \in N_w \cap N_a$ and $c \in \{R, T\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, T\}$ is not equal to $c$.

If we choose $(w, R)$ for $X$ then $X(w)$ and $Y(w)$ will agree. Also, if we choose $(w, c)$ for $x$ where $c$ is one of the at least $\Delta - 2$ colours not in $\{R, G, T\}$ or appearing on $N(w) - a$ then $X(w), Y(w)$ will agree. A similar argument yields that there are at least $\Delta$ colours $c$ for which choosing $(a, c)$ for $X$ will cause $X(a), Y(a)$ to agree - here we take advantage of the fact that $G$ appears on $N(a) - w$. Thus there are at least $2\Delta - 1$ choices which will cause $H(X, Y)$ to decrease by 1.

If a vertex $\alpha$ is adjacent to exactly one of $w, a$, there is at most one colour $c$ for which choosing $(\alpha, c)$ for $X$ will cause $X(\alpha), Y(\alpha)$ to disagree. If $\alpha$ is adjacent to both $x$ and $a$ then only $c = T$ will cause $X(\alpha), Y(\alpha)$ to disagree. Thus, since there is at least one vertex in $N_w \cap N_v$ (namely $b$), there are at most $2\Delta - 3$ choices which will cause $H(X, Y)$ to increase by 1.

Furthermore, there are at most $2\Delta$ vertices eligible to be chosen in step $i_2$ and so fewer than $4\Delta^2$ possible choices for a vertex and a colour. Therefore, $E_2 \leq -\frac{2}{4\Delta^2}$, as claimed.

If we are in situation (ii) where $T \neq R, G$, then we define our coupling $\mathcal{C}$ until step $i_2$ in nearly the same manner as in situation (i), the only difference being that we replace "$\{G, T\}$" by "$\{R, T\}$" in the rule for when $v \in N_a - N_w - w$. Our analysis is virtually identical, again yielding $E_2 \leq -\frac{2}{4\Delta^2}$.

As we remarked earlier, situation (i) with $T = R$ and situation (ii) with $T = G$ are identical to situation (iii), and there are no other possibilities if $B_1$ does not hold. So situation (iii) is the only remaining situation to be analysed. In this situation, we define our coupling $\mathcal{C}$ until step $i_2$ as follows. If we choose $(v, c)$ for $X$ then we choose $(v, c)$ for $Y$, unless:

- $v \in N_w - N_a - a$ or $v \in N_a - N_w - w$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$.

The analysis is very similar to that of (i) and (ii). Here there are at least $2\Delta - 2$ choices which cause $H(X, Y)$ to decrease by 1, and at most $2\Delta - 4$ choices which cause $H(X, Y)$ to increase by 1. The key observation is that there is no choice which will cause $X(b), Y(b)$ to disagree. Therefore, once again we have $\mathbf{Exp}(Z_2 - Z_1) \leq -\frac{2}{4\Delta^2}$.

Thus, if the choice in step $i_1$ is $(a, G)$ for $X$ and if $B_1$ doesn't hold then $E_2 \leq -\frac{1}{4\Delta^2}$. Therefore, assuming that $\mathbf{Pr}(B_1) < \frac{1}{2}$, we have $\mathbf{Pr}(A_2) > \frac{1}{8\Delta^2}$.

*Case 2: $w$ does not have 2 adjacent neighbours.* In this case, consider any neighbour $a$ of $w$. Again we define $B_1$ as in Case 1, and we can assume $\mathbf{Pr}(B_1) < \frac{1}{2}$ as otherwise $A_1$ holds.

Again, with probability at least $\frac{1}{2\Delta(\Delta+1)}$, step $i_1$ of the dynamics chooses $(a, G)$ for $X$ and $(a, R)$ for $Y$. Again, we have the same three possibilities as before, and we define $\mathcal{C}$ until step $i_2$ as before. The analysis of the first 2 (with $T \neq R, G$) is virtually the same, the only difference being that this time the number of choices which cause $H(X, Y)$ to increase is at most $2\Delta - 2$ rather than $2\Delta - 3$. Both yield $E_2 \leq -\frac{1}{4\Delta^2}$. In situation (iii), however, it is not true that $E_2$ is neccessarily strictly less than 0. So we consider the next change in $H$.

Consider any neighbour $b \neq a$ of $w$. Now we will focus on the possibility that $X$ chooses the vertex $b$ and colour $G$ in step $i_2$, and so $Y$ chooses $b$ and $R$. Our first step is to consider the possibilities by which this choice will produce no change in $H(X, Y)$, i.e. for which it will yield $H(X_{i_2}, Y_{i_2}) = H(X_{i_1}, Y_{i_1})$. Those possibilities are that in $X_{i_2 - 1}$:

- $b$ has a neighbour coloured $G$ and a neighbour other than $w$ coloured $R$,
- $b$ is coloured $R$ and has a neighbour coloured $G$, or
- $b$ is coloured $G$ and has a neighbour other than $w$ coloured $R$.

We define $B_2$ to be the event that $X_{i_2 - 1}$ satisfies one of these three possibilities. As was the case when we considered $B_1$, it is straightforward to compute that if the conditional probability of $B_2$, conditioned on the values of $X_{i_1}, Y_{i_1}$ is at least $\frac{1}{2}$, then $E_2 \leq -\frac{1}{8\Delta^2}$ and so $A_2$ holds.

Suppose then that this conditional probability of $B_2$ is less than $\frac{1}{2}$.

With probability at least $\frac{1}{2\Delta(2\Delta)}$, step $i_2$ chooses $(b, G)$ for $X$ and $(b, R)$ for $Y$. Supposing that $b$ was coloured $L$ before this step and that $B_2$ does not hold, we have 3 possible situations for the 2 new colourings:

(i) $X(w) = R, Y(w) = G; X(a) = G, Y(a) = R; X(b) = G, Y(b) = L.$

(ii) $X(w) = R, Y(w) = G; X(a) = G, Y(a) = R; X(b) = L, Y(b) = R.$

(iii) $X(w) = R, Y(w) = G; X(a) = G, Y(a) = R; X(b) = G, Y(b) = R.$

By now, no doubt, the reader has noticed the pattern of our coupling. Until step $i_3$, we define $\mathcal{C}$ as follows. (Note that, since we are in Case 2, $N_w \cap N_a = N_w \cap N_b = \emptyset$.)

*In situation (i):* If we choose $(v, c)$ for $X$ then we choose $(v, c)$ for $Y$, unless:

- $v \in N_w - \{a, b\}$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$;
- $v \in N_a - N_b - w$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$;
- $v \in N_b - w$ and $c \in \{G, L\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{G, L\}$ is not equal to $c$.

*In situation (ii):* If we choose $(v, c)$ for $X$ then we choose $(v, c)$ for $Y$, unless:

- $v \in N_w - \{a, b\}$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$;
- $v \in N_a - N_b - w$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$;
- $v \in N_b - w$ and $c \in \{R, L\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, L\}$ is not equal to $c$.

*In situation (iii):* If we choose $(v, c)$ for $X$ then we choose $(v, c)$ for $Y$, unless:

- $v \in N_a \cup N_b \cup N_w - \{a, b, w\}$ and $c \in \{R, G\}$ - here we choose $(v, c')$ for $Y$ where $c' \in \{R, G\}$ is not equal to $c$.

In situations (i) and (ii) when $L \neq R, G$, if, in $X$, either $b$ or $w$ gets assigned a colour other than $R, G, L$ which does not appear in its neighbourhood, then $H(X, Y)$ decreases by 1. The same is true if $a$ gets assigned a colour other than $R, G$ which does not appear in its neighbourhood. Using the fact that either $G$ or $R$ must appear in $N(b) - w$, we see that there are at least $3\Delta - 3$ such possible choices. For any vertex $u \in N(a) \cap N(b)$, there are at most 2 colours $c$ for which choosing $(u, c)$ for $X$ will cause $H$ to increase by 1. If $u$ is adjacent to exactly one of $a, b, w$ then there is at most one such colour $c$. Since $N(w) \cap N(a) = N(w) \cap N(b) = \emptyset$, ther are no other choices which can affect $H$. Thus, there are a total of at most $3\Delta - 4$ choices which could cause $H$ to increase by 1.

As before, if $B_2$ does not hold then situations (i) and (ii) with $L \in \{R, G\}$ are equivalent to situation (iii). Therefore, the only remaining situation to be analysed is (iii). In situation (iii), if either $a, b$ or $w$ gets assigned a colour other than $R, G$ which does not appear in its neighbourhood, then $H(X, Y)$ decreases by 1. There are at least $3\Delta - 2$ such possible choices, but at most $3\Delta - 4$ possible choices which increase $H(X, Y)$ by 1.

In either situation, $E_3 \leq -\frac{1}{6\Delta^2}$ and so $A_3$ holds.

Therefore in Case 2, if $\mathbf{Pr}(B_1) < \frac{1}{2}$ then with probability at least $\frac{1}{4\Delta(\Delta+1)} \times \frac{1}{4\Delta(2\Delta)} \geq \frac{1}{64\Delta^4}$, either $A_2$ or $A_3$ occurs, and so at least one of the following must hold: $\mathbf{Pr}(A_2) \geq \frac{1}{128\Delta^4}$, or $\mathbf{Pr}(A_3) \geq \frac{1}{128\Delta^4}$. $\square$

**Proof of Lemma 5** First we prove the lemma for the case where $X, Y$ initially differ on exactly one vertex.

Thus far, we have only defined our coupling $\mathcal{C}$ for the case that at first, $H(X, Y) = 1$ and for the initial steps of the procedure - up until step $i_2$ and, in some cases step $i_3$. Using path coupling, it is straightforward to define $\mathcal{C}$ for all other steps of the procedure such that the expected change in $H(X_i, Y_i)$ is always at most 0. Indeed, at any other step of our process, if $H(X_i, Y_i) = t$ then we define a sequence $X_i = X^0, X^1, ..., X^t = Y_i$ such that $H(X^j, X^{j+1}) = 1$. We couple each pair $X^j, X^{j+1}$ using the same rule that we used for Claim 4 and that we also used to define $\mathcal{C}$ until step $i_1$. This yields an expected change of at most 0 in $H(X^j, X^{j+1})$ and thus, by the Linearity of Expectation, $\mathbf{Exp}(H(X_{i+1}, Y_{i+1})) \leq H(X_i, Y_i)$.

That done, we next show that $\mathbf{Exp}(Z_3 - Z_0) \leq -\frac{1}{1600\Delta^6}$.

Summing over all possible values $(x,y)$ for $(X_{i_2}, Y_{i_2})$, we have

$$\begin{aligned}\mathbf{Exp}(Z_3 - Z_0) &= \mathbf{Exp}(Z_2 - Z_0) + \mathbf{Exp}(Z_3 - Z_2) \\ &= \mathbf{Exp}(Z_2 - Z_0) + \sum_{x,y} \mathbf{Pr}((X_{i_2}, Y_{i_2}) = (x,y)) \times E_3(x,y)\end{aligned}$$

We set $\mathcal{A}_j = \{(x,y) : \text{if } (X_{i_{j-1}}, Y_{i_{j-1}}) = (x,y) \text{ then } A_j \text{ holds}\}$. Since $\mathbf{Exp}(Z_2 - Z_0) \leq 0$ and $E_3$ is always at most 0, if $\mathbf{Pr}(A_3) \geq \frac{1}{128\Delta^4}$ then

$$\begin{aligned}\mathbf{Exp}(Z_3 - Z_0) &\leq \sum_{(x,y)\in\mathcal{A}_3} \mathbf{Pr}((X_{i_2}, Y_{i_2}) = (x,y)) \times E_3(x,y) \\ &\leq \tfrac{1}{128\Delta^4} \times \left(-\tfrac{1}{12\Delta^2}\right) < -\tfrac{1}{1600\Delta^6}.\end{aligned}$$

If $\mathbf{Pr}(A_2) \geq \frac{1}{128\Delta^4}$ then the analogous argument yields that $\mathbf{Exp}(Z_2 - Z_0) \leq -\frac{1}{1600\Delta^6}$, and since $\mathbf{Exp}(Z_3 - Z_2) \leq 0$ we have $\mathbf{Exp}(Z_3 - Z_0) \leq -\frac{1}{1600\Delta^6}$.

Finally, if $A_1$ holds, then we have $\mathbf{Exp}(Z_1 - Z_0) \leq -\frac{1}{12\Delta^2}$ and since $\mathbf{Exp}(Z_3 - Z_1) \leq 0$ we have $\mathbf{Exp}(Z_3 - Z_0) \leq -\frac{1}{12\Delta^2}$.

We have shown that the expected total change up to step $i_3$ is small. Now we must show that with sufficiently high probability, we will reach step $i_3$ before time $Kn$. So we let $C$ be the event that either $i_3 \leq Kn$ or $X_{i_j} = Y_{i_j}$ and $i_j < Kn$ for some $j < 3$ (in which case $i_3 = \infty$).

For any pair $(X_i, Y_i)$ which differs on at least one vertex, $u$, the probability that the next step chooses $u$ or a neighbour of $u$ is at least $\frac{1}{n}$. Therefore, $\mathbf{Pr}(\overline{C})$ is at most the probability that the binary variable $BIN(Kn, \frac{1}{n})$ is less than 3. As $n$ tends to infinity, this approaches the probability that the Poisson variable with mean $K$ is less than 3, which is $\mathrm{e}^{-K}(1 + K + \frac{1}{2}K^2)$. In particular, for $n$ sufficiently large,

$$\mathbf{Pr}(\overline{C}) < 2\mathrm{e}^{-K}(1 + K + \frac{1}{2}K^2) < \tfrac{1}{10000\Delta^6},$$

for $K \geq 12\log(10\Delta)$.

Since

$$\begin{aligned}-\tfrac{1}{1600\Delta^6} \geq \mathbf{Exp}(Z_3 - Z_0) &= \mathbf{Exp}(Z_3 - Z_0|C) \times \mathbf{Pr}(C) + \mathbf{Exp}(Z_3 - Z_0|\overline{C}) \times \mathbf{Pr}(\overline{C}) \\ &\geq \mathbf{Exp}(Z_3 - Z_0|C) \times (1 - \tfrac{1}{10000\Delta^6}) + (-1) \times \tfrac{1}{10000\Delta^6},\end{aligned}$$

we have $\mathbf{Exp}(Z_3 - Z_0|C) \leq -\frac{1}{2000\Delta^6}$. Since conditional on any value of $i_3 \leq Kn$ and any value of $(X_{i_3}, Y_{i_3})$ we have $\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - Z_3) \leq 0$, this yields

$$\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - H(X_0, Y_0)|C) \leq -\tfrac{1}{2000\Delta^6},$$

which, along with our bound on $\mathbf{Pr}(\overline{C})$ and the absolute upper bound $Z_3 - Z_0 \leq 3$ yields

$$\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - H(X_0, Y_0)) \leq -\tfrac{1}{2000\Delta^6} + 3 \times \tfrac{1}{10000\Delta^6} = -\tfrac{1}{5{,}000\Delta^6}.$$

This proves our lemma for the case where $X, Y$ initially differ on exactly one vertex.

We still must prove our lemma for the general case. The first step is to define the coupling $\mathcal{C}$ for the general case where initially $H(X, Y) = t$. This is done, as has been done many times before in this paper, using path coupling, and this yields

$$\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - H(X, Y) \leq -\tfrac{1}{5000\Delta^6} \times t.$$

We omit what are, by this point, straightforward (and highly repetitive) path coupling details. □

# 2 Independent Sets

In this section we prove Theorem 2. Again we use $H(X, Y)$ to denote the Hamming distance between two independent sets $X, Y$, i.e. the number of vertices that are in one set but not the other.

Suppose that $X, Y$ are two independent sets of a graph $G$ with maximum degree 4, and that $X, Y$ differ only in that a single vertex $w$ is in $X$ but not in $Y$. We couple a single iteration of IDD on $X$ and on $Y$ by choosing the same vertex $v$ for $X$ and $Y$. If $v \notin w \cup N(w)$ then we choose to modify $Y$ iff we choose to modify $X$. If $v \in N(w)$ and $v$ has a neighbour in $Y$ then it has at least 2 neighbours in $X$ and so there is no decision to be made for $X$, and we just decide whether to drag in $Y$ if the number of neighbours is exactly 1. Otherwise, we couple the decisions as to whether to modify $X$ and $Y$ as follows:

(1) if $v = w$ then we choose to add it to $Y$ iff we choose not to remove it from $X$;

(2) if $v \in N(w)$ and $v$ has no neighbours in $Y$ then we pick a uniform real $p \in (0, 1)$ and we

(i) drag $v$ in $X$ and add $v$ to $Y$ if $p < \frac{1}{8}$,

(ii) do not modify $X$ and add $v$ to $Y$ if $\frac{1}{8} \leq p < \frac{1}{2}$,

(iii) do not modify $X$ or $Y$ otherwise;

We remark that it is straightforward to check that under this coupling, the expected change in $H(X,Y)$ after one iteration is at most 0. Indeed, if we choose $v = w$ then the expected change is -1, and if we choose $v$ to be any neighbour $u \in N_w$ such that $u$ has at most one neighbour in $Y$, then the expected change is $+\frac{1}{4}$. For any other choice of $v$, there can be no change. Since $w$ has only 4 neighbours, the overall expected change is at most 0. Using the same ideas as in the earlier section, it is straightforward to extend this coupling via path coupling so that for any initial pair $X, Y$, the expected change in $H(X,Y)$ is at most 0. However, we need to show that after $Kn$ steps, the expected change is bounded away from 0:

**Lemma 8** *There exists a constant $K$ and a coupling $\mathcal{C}$ such that for any initial unequal pair $(X,Y)$, coupled by $\mathcal{C}$, $\mathbf{Exp}(H(X_{Kn}, Y_{Kn}) - H(X,Y)) \leq -\frac{1}{10^{11}} \times H(X,Y)$.*

Theorem 2 follows from Lemma 8 by Lemma 6 or by using the same argument as in the proof of Theorem 1. We omit the straightforward and repetitive details.

We define $D$ to be the symmetric difference of $X, Y$, i.e. the set of vertices on which $X$ and $Y$ differ. We will define $\mathcal{C}$ so that $H(X,Y)$ cannot change unless we choose a vertex in $D \cup N(D)$. So we define $i_j$ to be the $j$th step which chooses a vertex of $D \cup N(D)$ (if $X_{i_j} = Y_{i_j}$ then we define $i_{j'} = \infty$ for every $j' > j$. We define $Z_j, E_j$ as in the previous section. We define $A_j$ to be the event that $E_j \leq -\frac{1}{100}$.

The main step is analagous to Lemma 7:

**Lemma 9** *If $(X,Y)$ initially differ on exactly one vertex $w$, then either $A_1$ holds or $\mathbf{Pr}(A_i) > \frac{1}{4\times 10^8}$ for some $i \in \{2,3,4\}$.*

**Proof** The proof follows along the same lines as that of Lemma 7.

*Case 1: $w$ has two adjacent neighbours $a, b$.* We define $B_1$ to be the event that in $Y_{i_1-1}$, $\deg_Y(a) \geq 2$. Conditional on $B_1$ holding, it is easy to check that the conditional expected value $\mathbf{Exp}(H(X_{i_1}, Y_{i_1}) - H(X_0, Y_0)) \leq$

$-\frac{1}{4} \times \frac{1}{|D \cup N(D)|} \leq -\frac{1}{20}$. It follows that if $\mathbf{Pr}(B_1) \geq \frac{1}{2}$ then $A_1$ holds. Suppose then that $\mathbf{Pr}(B_1) < \frac{1}{2}$. With probability at least $\frac{1}{5}$, in step $i_1$ we choose vertex $a$. If $B_1$ does not hold, then with probability at least $\frac{1}{8}$ we will either add $a$ to $Y$ or drag $a$ in $Y$. I.e., with probability at least $\frac{1}{2} \times \frac{1}{5} \times \frac{1}{8} = \frac{1}{80}$, after step $i_1$ we will be one of the following two situations:

(i) $D = \{w, a\}$ where $w \in X, a \in Y$; or

(ii) for some $u \in N_a$, $D = \{w, a, u\}$ where $w, u \in X, a \in Y$.

We now describe our coupling rule. We choose our random vertex $v$ which we use for both $X$ and $Y$. If $v$ has at least 2 neighbours in $X$ or at least 2 neighbours in $Y$ then there is no choice to be made for at least one of the two sets, and so the decision of what to do with $v$ is not coupled. Otherwise:

- if $v \in X$ and $v \in Y$ then we remove $v$ from $X$ iff we remove $v$ from $Y$.
- if $v \notin X$ and $v \notin Y$ then
  - if $\deg_X(v) = \deg_Y(v) \leq 1$, then we modify $X$ iff we modify $Y$;
  - if $\deg_X(v) = 0, \deg_Y(v) = 1$, then the choice is coupled as described in (2) above; the case $\deg_X(v) = 1, \deg_Y(v) = 0$ is coupled in the symmetric manner;
- if $v \in X$ and $v \notin Y$ then
  - if $\deg_Y(v) = 0$ then the choice is coupled as in (1) above;
  - if $\deg_Y(v) = 1$ then with probability $\frac{1}{2}$, $v$ is deleted from $X$ and $Y$ is not modified; with probability $\frac{1}{8}$ we drag $v$ in $Y$ and $X$ is not modified.

The case $v \in Y$ and $v \notin X$ is symmetric.

It is straightforward to compute that in situation (i) we have $E_2 \leq -\frac{1}{2} \times \frac{1}{|D \cup N(D)|} \leq -\frac{1}{14}$ and in situation (ii) $E_2 \leq -\frac{1}{2} \times \frac{1}{|D \cup N(D)|} \leq -\frac{1}{20}$. Therefore, if $\mathbf{Pr}(B_1) < \frac{1}{2}$ then with probability at least $\frac{1}{80}$, $A_2$ holds.

*Case 2: $w$ does not have any 2 adjacent neighbours.* Let $a, b, c$ be any 3 neighbours of $w$. We define $B_1$ as in Case 1; $B_2$ is the event that in $Y_{i_2-1}$, $\deg_Y(b) \geq 2$; $B_3$ is the event that in $Y_{i_3-1}$, $\deg_Y(c) \geq 2$.

If $\mathbf{Pr}(B_1) \geq \frac{1}{2}$ then, as in Case 1, $A_1$ holds; otherwise, with probability at least $\frac{1}{80}$, step $i_1$ will either add or drag $a$ in $Y$. Suppose that step $i_1$ does this. If $\mathbf{Pr}(B_2|X_{i_1}, Y_{i_1}) \geq \frac{1}{2}$, then again, it is easy to compute that $A_2$ holds; otherwise, with probability at least $\frac{1}{2\times 8\times 11}$, step $i_2$ will either add or drag $b$ in $Y$. Suppose that step $i_2$ does this. If $\mathbf{Pr}(B_3|X_{i_2}, Y_{i_2}) \geq \frac{1}{2}$, then again, it is easy to compute that $A_3$ holds; otherwise, with probability at least $\frac{1}{2\times 8\times 17}$, step $i_3$ will either add or drag $c$ in $Y$.

Thus, with probability at least $\frac{1}{80} \times \frac{1}{2\times 8\times 11} \times \frac{1}{2\times 8\times 17} > 10^{-8}$, either one of $A_1, A_2, A_3$ holds, or the first 3 changes to $H(X,Y)$ either add or drag $a, b, c$ in $Y$. If the latter happens then we use the same coupling as in Case 1, and it is straightforward to verify that $E_4 \leq -\frac{1}{4} \times \frac{1}{|D \cup N(D)|} \leq -\frac{1}{92}$, and so $A_4$ holds. $\square$

The remainder of the proof of Lemma 8 now follows in the same manner as that of Lemma 5. The main steps are: (i) $\mathbf{Exp}(Z_4 - Z_0) \leq -\frac{1}{4*10^{10}}$, and (ii) for $K$ sufficiently large, with probability at least $\frac{1}{10^{12}}$ we have $i_4 \leq Kn$ or $X_{i_j} = Y_{i_j}$ and $i_j < Kn$ for some $j < 4$. We omit the straightforward and repetitive details.

**Acknowledgements:** I would like to thank Catherine Greenhill and an anonymous referee for many very helpful suggestions, and for pointing out a minor error in an earlier draft. Also, some discussions with Catherine Greenhill and Martin Dyer were very helpful.

# References

[1] D. 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** (1986), 243 - 297.

[2] R. Bubley and M. Dyer, *Path coupling: a technique for proving rapid mixing in Markov chains.* 28th Annual Symposium on Foundations of Computer Science (1997), 223 - 231.

[3] A. Czumaj, P. Kanarek, M. Kutylowski and K. Lorys, *Delayed path coupling and generating ranom permutations via distributed stochastic processes.* Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (1999), 271 - 280.

[4] M. Dyer, L. Goldberg, C. Greenhill, M. Jerrum and M. Mitzenmacher, *An extension of path coupling and its application to the Glauber dynamics for path colourings,* Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms (2000), 355 - 363.

[5] M. Dyer and C. Greenhill, *On Markov chains for independent sets*, J. Algorithms **35** (2000), 17 - 49.

[6] M. Dyer and C. Greenhill, *A more rapidly mixing Markov chain for graph colourings.* Random Structures and Algorithm **13** (1998), 285 - 317.

[7] M. Dyer and C. Greenhill, *Random walks on combinatorial objects.* Surveys in Combinatorics 1999, J. Lamb and D. Preece eds., vol 267 of London Math. Soc. Lecture Note Series, Cambridge U. Press, 1999, 101 - 136.

[8] M. Dyer, C. Greenhill and M. Molloy, in preparation.

[9] M. Dyer, A. Frieze and M. Jerrum, *On counting independent sets in sparse graphs.* 30th Annual Symposium on Foundations of Computer Science (1999), 210 - 217.

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

[11] M. Jerrum, *Mathematical foundations of the Markov chain Monte Carlo method.* Probabilistic Methods for Algorithmic Discrete Mathematics, eds. M. Habib, C. McDiarmid, J. Ramirez-Alfosin and B. Reed. Springer, 1998, 116 - 165.

[12] M. Luby and E. Vigoda, *Approximately counting up to four.* 29th Annual Symposium on Foundations of Computer Science (1998), 682 - 687.

[13] J. Salas and A. Sokal, *Absence of phase transition for antiferromagnetic Potts model via the Dobrushin uniqueness theorem*, J. Statistical Physics **86** (1997), 551 - 579.

[14] E. Vigoda, *Improved bounds for sampling colorings.* J. Mathematical Physics **41** (2000), 1555 - 1569.