Metric Construction, Stopping Times and Path Coupling

Magnus Bordewich*, Martin Dyer* and Marek Karpinski†

Abstract

In this paper we examine the importance of the choice of metric in path coupling, and the relationship of this to *stopping time analysis*. We give strong evidence that stopping time analysis is no more powerful than standard path coupling. In particular, we prove a stronger theorem for path coupling with stopping times, using a metric which allows us to restrict analysis to standard one-step path coupling. This approach provides insight for the design of non-standard metrics giving improvements in the analysis of specific problems.

We give illustrative applications to hypergraph independent sets and SAT instances, hypergraph colourings and colourings of bipartite graphs. In particular we prove rapid mixing for Glauber dynamics on independent sets in hypergraphs whenever the minimum edge size m and degree Δ satisfy $m \geq \Delta + 2$, and for all edge sizes when $\Delta = 3$. Previously rapid mixing was only known for $m \geq 2\Delta + 1$. This result leads to approximation schemes for monotone SAT formulae in which the maximum number of occurrences of a variable (Δ) and the minimum number of variables per clause (m) satisfy the same condition. For Glauber dynamics on proper colourings of 3-uniform hypergraphs we prove rapid mixing whenever the number of colours q is at least $\left\lceil \frac{3}{2}\Delta + 1 \right\rceil$. Previously the best known result was for $q \geq 1.65\Delta$ and $\Delta \geq \Delta_0$ for some large Δ_0. Finally we prove rapid mixing of scan dynamics (where the order of vertex updates is deterministic) for proper colourings of bipartite graphs whenever $q > f(\Delta)$, where $f(\Delta) \to \beta\Delta$, as $\Delta \to \infty$, and β satisfies $\frac{1}{\beta}e^{1/\beta} = 1$, ($\beta \approx 1.76$). This gives rapid mixing with fewer colours than Vigoda's $11\Delta/6$ bound [22], whenever $\Delta \geq 31$, and equals this bound for $\Delta \geq 14$.

*School of Computing, University of Leeds, Leeds LS2 9JT. Email: {dyer,magnusb}@comp.leeds.ac.uk.
†Dept. of Computer Science, University of Bonn, 53117 Bonn. Email: marek@cs.uni-bonn.de.

1 Introduction

Path coupling [5] has proved to be a useful technique for analysing Markov chains. Analysis is carried out relative to a chosen *metric* on the state space, for example the Hamming distance on the independent sets in a graph or hypergraph. The limitations of the analysis are always caused by certain "bad" configurations. But these configurations may be unlikely in a typical realisation of the chain. Consequently, path coupling has been augmented by other techniques, such as *stopping time* analysis. See [2, 10, 15, 18] for some applications of this technique. A general theorem for applying stopping times was proved in [15], and improved somewhat in [2].

The stopping time approach is applicable when the bad configurations have a reasonable probability of becoming less bad as time passes. For example, the bad configurations for the Glauber dynamics on hypergraph independent sets involve almost full edges containing the change vertex. (See [2] for details.) However, it seems likely that the number of occupied vertices in these edges will have been reduced before we must either increase or decrease the distance between the coupled chains. This observation allows a greatly improved analysis [2].

The stopping time approach is a multistep analysis, and appears to give a powerful extension of path coupling. However, in this paper we provide strong evidence that the stopping time approach is no more powerful than single-step path coupling. We observe that, in cases where stopping times can be employed to advantage, equally good or better results can be achieved by using a suitably tailored *metric* in the one-step analysis. The intuition behind the choice of metric will be illustrated with several examples.

In fact, our first example is a proof of a theorem for path coupling using stopping times, relying on a particular choice of metric which enables us to work with the standard one-step path coupling. The resulting theorem is stronger than those in [2, 15]. The proof implies that all results obtained using stopping times can just as well be obtained using standard path coupling and the right choice of metric. This does not immediately imply that we can abandon the analysis of stopping times. Determining the metric used in our proof involves bounding the expected distance at a stopping time. However the proof does suggest that it may be better to carry out one-step analysis using a metric indicated directly by the stopping time intuition.

With this insight, we revisit the Glauber dynamics for hypergraph independent sets (or equivalently, satisfying assignments of monotone SAT formulas), and hypergraph colourings, analysed in [2] using stopping times. We find that we are able to obtain considerably stronger results than those obtained in [2], using metrics inspired by the stopping times considerations but then optimised to give the best results. The technical advantage arises from the possibility of using linearity of expectation where stopping time analysis must use concentration inequalities and union bounds.

We note that this paper does not contain the first uses of "clever" metrics with path coupling. See [6, 17] for examples. But we do give the first widely applicable rationale for choosing a good metric. While there have been instances in the literature of optimising the *chain* [12, 22], the only previous analysis of which we are aware which uses optimisation of the *metric* appeared in [17].

The organisation of the paper is as follows. In section 2 we prove a better stopping time theorem than previously known, using only standard path coupling. In section 3 we give our improved results for sampling independent sets in hypergraphs, and in section 4, applications to counting the number of satisfying assignments in monotone SAT formulas. In section 5 we give improved results for sampling colourings of 3-uniform hypergraphs. Finally, in section 6 we give a completely new application, to the "scan" chain for sampling colourings of bipartite graphs. For even relatively small values of Δ, our results improve Vigoda's [22] celebrated $11\Delta/6$ bound on the number of colours required for rapid mixing.

2 Path coupling and stopping times

We first deal with the most useful and applicable case, in which the stopping time for a pair of coupled chains is the first time that the distance between the two chains changes. This simplifies the proofs and makes the thrust of the argument clearer. In Section 2.2 we do deal with more general stopping times, however it should be noted that so far all applications of stopping times results in path coupling have only used this simple form of stopping time.

2.1 Distance-change stopping time

Let $\mathcal{M}$ be a Markov chain on state space Ω. Let d be an integer valued metric on $\Omega \times \Omega$, and let (X_t, Y_t) be a path coupling for $\mathcal{M}$. We define T_t, a stopping time for the pair $(X_t, Y_t) \in S$, to be the smallest $t' > t$ such that $\mathrm{d}(X_{t'}, Y_{t'}) \neq \mathrm{d}(X_t, Y_t)$. We will define a new metric d′ such that if we have contraction in the metric d at the stopping times, then we have contraction in the metric d′ at every step which has a positive probability of being a stopping time.

Let $\alpha > 0$ be a constant such that $\mathbb{E}[\mathrm{d}(X_{T_t}, Y_{T_t})] \leq \alpha \mathrm{d}(X_t, Y_t)$ for all (X_t, Y_t) in S. If $\alpha < 1$, then for any $(X_t, Y_t) \in S$, we simply define d′ as follows.

$$\mathrm{d}'(X_t, Y_t) = (1 - \alpha)\mathrm{d}(X_t, Y_t) + \mathbb{E}[\mathrm{d}(X_{T_t}, Y_{T_t})] \leq \mathrm{d}(X_t, Y_t). \tag{1}$$

The metric is extended in the usual way to pairs $(X_t, Y_t) \notin S$, using shortest paths. See, for example, [11]. We will apply path coupling with the metric d′ and the original coupling. First we show a contraction property for this metric.

Lemma 2.1. *If* $\mathbb{E}[\mathrm{d}(X_{T_t}, Y_{T_t})] \leq \alpha \mathrm{d}(X_t, Y_t) < \mathrm{d}(X_t, Y_t)$ *for all* (X_t, Y_t) *in* S*, then*

$$\mathbb{E}[\mathrm{d}'(X_k, Y_k) \,|\, X_0, Y_0] \leq \big(1 - (1 - \alpha) \Pr(T_0 \leq k)\big)\mathrm{d}'(X_0, Y_0).$$

Proof. We prove this by induction on k. It obviously holds for $k = 0$, since $T_0 > 0$. Using $\mathbb{1}_{\mathcal{A}}$ to denote the 0/1 indicator of any event $\mathcal{A}$, we may write (1) as

$$\mathrm{d}'(X_0, Y_0) \;=\; (1 - \alpha)\mathrm{d}(X_0, Y_0) + \mathbb{E}[\mathrm{d}(X_{T_k}, Y_{T_k})\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}(X_{T_0}, Y_{T_0})\mathbb{1}_{T_0 \leq k}], \tag{2}$$

since if $T_0 > k$ then $T_k = T_0$. Similarly, we have

$$\begin{aligned}
\mathbb{E}[\mathrm{d}'(X_k, Y_k)] \;&=\; \mathbb{E}[\mathrm{d}'(X_k, Y_k)\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}'(X_k, Y_k)\mathbb{1}_{T_0 \leq k}] \\
&=\; (1 - \alpha)\mathbb{E}[\mathrm{d}(X_k, Y_k)\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}(X_{T_k}, Y_{T_k})\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}'(X_k, Y_k)\mathbb{1}_{T_0 \leq k}]. \\
&=\; (1 - \alpha)\mathbb{E}[\mathrm{d}(X_0, Y_0)\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}(X_{T_k}, Y_{T_k})\mathbb{1}_{T_0 > k}] + \mathbb{E}[\mathrm{d}'(X_k, Y_k)\mathbb{1}_{T_0 \leq k}]. \qquad (3)
\end{aligned}$$

Subtracting (2) from (3), we have

$$\mathbb{E}[\mathrm{d}'(X_k,Y_k)] - \mathrm{d}'(X_0,Y_0) = -(1-\alpha)\mathbb{E}[\mathrm{d}(X_0,Y_0)\mathbb{1}_{T_0\le k}] + \mathbb{E}[(\mathrm{d}'(X_k,Y_k) - \mathrm{d}(X_{T_0},Y_{T_0}))\mathbb{1}_{T_0\le k}].$$

For $T_0 \le k$, since $k - T_0 \le k-1$ the inductive hypothesis implies $\mathbb{E}[\mathrm{d}'(X_k,Y_k)\,|\,X_{T_0},Y_{T_0}] \le \mathrm{d}'(X_{T_0},Y_{T_0}) \le \mathrm{d}(X_{T_0},Y_{T_0})$, (if $(X_k,Y_k)\notin S$ this follows by linearity). Hence we have

$$\mathbb{E}[\mathrm{d}'(X_k,Y_k)] - \mathrm{d}'(X_0,Y_0) \le -(1-\alpha)\mathbb{E}[\mathrm{d}(X_0,Y_0)\mathbb{1}_{T_0\le k}],$$

The conclusion follows, since $\mathbb{E}[\mathrm{d}(X_0,Y_0)\mathbb{1}_{T_0\le k}] = \Pr(T_0\le k)\mathrm{d}(X_0,Y_0) \ge \Pr(T_0\le k)\mathrm{d}'(X_0,Y_0)$. □

We may now prove the first version of our main result.

Theorem 2.2. *Let $\mathcal{M}$ be a Markov chain on state space Ω. Let d be an integer valued metric on Ω, and let (X_t,Y_t) be a path coupling for $\mathcal{M}$. Let T_t be the above stopping times. Suppose for all $(X_0,Y_0)\in S$ and for some integer k and $p>0$, that*

(i) $\Pr[T_0\le k]\ge p$,

(ii) $\mathbb{E}[\mathrm{d}(X_{T_0},Y_{T_0})/\mathrm{d}(X_0,Y_0)] \le \alpha < 1$.

Then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies

$$\tau(\varepsilon) \le \frac{k}{p(1-\alpha)}\ln\Big(\frac{eD}{\varepsilon(1-\alpha)}\Big).$$

where $D = \max\{\mathrm{d}(X,Y) : X,Y\in\Omega\}$.

Proof. ¿From Lemma 2.1, d' contracts by a factor $1-(1-\alpha)p \le e^{-(1-\alpha)p}$ for every k steps of $\mathcal{M}$. Note also that $\mathrm{d}'\le D$. It follows that, at time $\tau(\varepsilon)$, we have

$$\Pr(X_\tau \ne Y_\tau) \le \mathbb{E}[\mathrm{d}(X_\tau,Y_\tau)] \le \frac{\mathbb{E}[\mathrm{d}'(X_\tau,Y_\tau)]}{1-\alpha} \le \frac{De^{-(1-\alpha)p\tau/k}}{1-\alpha} \le \varepsilon,$$

from which the theorem follows. □

If $1-\alpha$ is small compared to ε, it is possible to do better than this. We will need the technical Lemma 2.3 below, which says that we will not have to wait too long for a stopping time to occur.

Lemma 2.3. *If $\mathcal{M}$ satisfies the conditions of Theorem 2.2 then* $\Pr[T_t > t+t'] \le (1-p)^{\lfloor t'/k\rfloor}$.

Proof. We prove this by induction on t'. It clearly holds for all t and $t'<k$ since $\lfloor t'/k\rfloor = 0$. Suppose inductively that $\Pr[T_t > s+t] \le (1-p)^{\lfloor s/k\rfloor}$ for all t and $s<t'$. Then, if $t'\ge k$,

$$\begin{aligned}\Pr[T_t > t+t'] &= \Pr[T_t > t+t'-k \text{ and } T_{t+t'-k} > t+t'] \\ &= \Pr[T_t > t+t'-k]\Pr[T_{t+t'-k} > t+t' \mid T_t > t+t'-k].\end{aligned}$$

Since the process is Markovian, and by condition (i),

$$\begin{aligned}\Pr[T_{t+t'-k} > t+t' \mid T_t > t+t'-k] &\leq \max\{\Pr[T_{t+t'-k} > t+t'] : (X_{t+t'-k}, Y_{t+t'-k}) \in S\} \\ &= \max\{\Pr[T_0 > k\ ] : (X_0, Y_0) \in S\} \\ &\leq 1-p.\end{aligned}$$

By the inductive hypothesis this gives

$$\Pr[T_t > t'+t] \leq (1-p)^{\lfloor (t'-k)/k \rfloor}(1-p) = (1-p)^{\lfloor t'/k \rfloor}.$$ □

Theorem 2.4. *Let $\mathcal{M}$ be a Markov chain on state space Ω. Let d be an integer valued metric on $\Omega \times \Omega$, and let (X_t, Y_t) be a path coupling for $\mathcal{M}$. Let T_t be the above stopping time. Suppose for all $(X_0, Y_0) \in S$ and for some integer k and $p > 0$, that*

(i) $\Pr[T_0 \leq k] \geq p$,

(ii) $\mathbb{E}[\mathrm{d}(X_{T_0}, Y_{T_0})/\mathrm{d}(X_0, Y_0)] \leq \alpha < 1$.

Then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies

$$\tau(\varepsilon) \leq \frac{k(2-\alpha)}{p(1-\alpha)} \ln\Big(\frac{2eD}{\varepsilon}\Big).$$

where $D = \max\{\mathrm{d}(X, Y) : X, Y \in \Omega\}$.

Proof. Let $X_t = Z_0^0, Z_0^1, \ldots, Z_0^r = Y_t$ be a shortest path from X_t to Y_t in the metric d', such that $(Z_0^i, Z_0^{i+1}) \in S$ $(i = 0, \ldots, r-1)$. If $\mathfrak{t}_i$ is the stopping time for (Z_0^i, Z_0^{i+1}) then, using Lemma 2.3,

$$\begin{aligned}\Pr(X_{t+t'} \neq Y_{t+t'} \mid X_t, Y_t) &\leq \Pr(\exists i : Z_{t'}^i \neq Z_{t'}^{i+1}) \\ &\leq \Pr(\exists i : Z_{\mathfrak{t}_i}^i \neq Z_{\mathfrak{t}_i}^{i+1} \text{ or } \mathfrak{t}_i > t') \\ &\leq \sum_{i=0}^{r-1} \big(\mathbb{E}[\mathrm{d}(Z_{\mathfrak{t}_i}^i, Z_{\mathfrak{t}_i}^{i+1})] + \Pr(\mathfrak{t}_i > t')\big) \\ &\leq \sum_{i=0}^{r-1} \Big(\mathrm{d}'(Z_0^i, Z_0^{i+1}) + (1-p)^{\lfloor t'/k \rfloor}\Big) \\ &\leq \mathrm{d}'(X_t, Y_t) + D(1-p)^{\lfloor t'/k \rfloor}.\end{aligned}$$

Hence

$$\begin{aligned}\Pr(X_{t+t'} \neq Y_{t+t'}) &\leq \mathbb{E}[\mathrm{d}'(X_t, Y_t)] + D(1-p)^{\lfloor t'/k \rfloor}, \\ &\leq De^{-(1-\alpha)p\lfloor t/k \rfloor} + D(1-p)^{\lfloor t'/k \rfloor}. \\ &\leq D\big(e^{-(1-\alpha)p\lfloor t/k \rfloor} + e^{-p\lfloor t'/k \rfloor}\big).\end{aligned}$$

Therefore

$$\Pr(X_{t+t'} \neq Y_{t+t'}) \leq \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon, \quad \text{if } t \geq k\left\lceil \frac{\ln(2D/\varepsilon)}{p(1-\alpha)} \right\rceil \text{ and } t' \geq k\left\lceil \frac{\ln(2D/\varepsilon)}{p} \right\rceil.$$

The statement of the theorem now follows easily. □

2.2 General stopping times

We now extend the results proved in this section to incorporate stopping times other than the first time at which the distance changes. In order to make sense in the context of path coupling, the stopping times must satisfy the following conditions.

STOPPING TIME CONDITIONS:

1. There must be a stopping time $T(X_0, Y_0)$ defined for each pair $(X_0, Y_0) \in S$, such that $\mathbb{E}[\mathrm{d}(X_{T_0(X,Y)}, Y_{T_0(X,Y)})] \leq \alpha \mathrm{d}(X_0, Y_0)$.
2. For all $(X_0, Y_0) \in S$ we have $\Pr[T(X_0, Y_0) \leq k] \geq p$.
3. The coupling should be Markovian.

We may assume that for $(X_0, Y_0) \in S$ if $X_t = Y_t$ then $T(X_0, Y_0) \leq t$. Since the future evolution of (X_t, Y_t) does not depend on the evolution up to time t, by 1 and 3 it follows that for all $t \geq 0$ there is a stopping time $T_t(X, Y)$ such that if $X_t = X, Y_t = Y$ then $\mathbb{E}[\mathrm{d}(X_{T_t(X,Y)}, Y_{T_t(X,Y)})] \leq \alpha \mathrm{d}(X_t, Y_t)$. Moreover, from 2 and 3, it follows that $\Pr[T_t(X, Y) \leq k + t] \geq p$.

When dealing with the first change in distance we had the benefit that for all $(X_t, Y_t) \in S$ and $t' > t$, if $T_t(X_t, Y_t) > t'$ then $(X_{t'}, Y_{t'}) \in S$ and also $T_{t'}(X_{t'}, Y_{t'}) = T_t(X_t, Y_t)$. This no longer necessarily holds. We must therefore be more careful about exactly which stopping time we are referring to at any time and regarding any pair of states.

Let (X_t, Y_t) be a coupled evolution of the chain, and let $P_t = (X_t = Z_t^0, Z_t^1, Z_t^2, \ldots, Z_t^{d_t} = Y_t)$ be the path-coupling path from X_t to Y_t, so that $(Z_t^i, Z_t^{i+1}) \in S$ for all i, t. We will inductively define a set of *starting pairs* in the paths P_t, $t \geq 0$ as follows.

1. For all $i, (Z_0^i, Z_0^{i+1})$ is a starting pair.
2. For each $(Z_{t_1}^i, Z_{t_1}^{i+1})$ if there is a time $t_0 \leq t_1$ and starting pair $(Z_{t_0}^j, Z_{t_0}^{j+1}) \in P_{t_0}$ such that $(Z_{t_1}^i, Z_{t_1}^{i+1})$ is in the subpath of P_{t_1} which evolved from $(Z_{t_0}^j, Z_{t_0}^{j+1})$ and $T_{t_0}(Z_{t_0}^j, Z_{t_0}^{j+1}) > t_1$, then $(Z_{t_1}^i, Z_{t_1}^{i+1})$ is *not* a starting pair but t_0 is the *starting time associated with* $(Z_{t_1}^i, Z_{t_1}^{i+1})$ and $(Z_{t_0}^j, Z_{t_0}^{j+1})$ is the *starting pair associated with* $(Z_{t_1}^i, Z_{t_1}^{i+1})$.
3. For each $(Z_{t_1}^i, Z_{t_1}^{i+1})$ such that there is no time and pair as above, then $(Z_{t_1}^i, Z_{t_1}^{i+1})$ is defined to be a starting pair. Note that in this case there must be a time $t_0 \leq t_1$ and starting pair $(Z_{t_0}^j, Z_{t_0}^{j+1}) \in P_{t_0}$ such that $(Z_{t_1}^i, Z_{t_1}^{i+1})$ is in the subpath of P_{t_1} which evolved from $(Z_{t_0}^j, Z_{t_0}^{j+1})$ and $T_{t_0}(Z_{t_0}^j, Z_{t_0}^{j+1}) = t_1$.

For a starting pair $(Z_{t_0}^j, Z_{t_0}^{j+1})$, we define the distance at time $t_1, t_0 \leq t_1 < T_{t_0}(Z_{t_0}^j, Z_{t_0}^{j+1})$ to be

$$d_{t_1}(Z_{t_0}^j, Z_{t_0}^{j+1}) = (1 - \alpha)\mathrm{d}(Z_{t_0}^j, Z_{t_0}^{j+1}) + \mathbb{E}\left[\mathrm{d}(Z^j_{T_{t_0}(Z_{t_0}^j Z_{t_0}^{j+1})}, Z^{j+1}_{T_{t_0}(Z_{t_0}^j Z_{t_0}^{j+1})}) \mid \mathcal{F}_{t_1}\right] \tag{4}$$

where $\mathcal{F}_t$ is the σ-algebra generated by $\{(X_{t'}, Y_{t'}) : t' \leq t\}$. Thus $\{\mathcal{F}_t : t' \geq 0\}$ is the filtration generated by the coupling. The distance at times not in the given range is zero. This is analogous to the definition of the new metric in equation (1). At a time t we are interested in the set $\mathcal{SP}_t$ of starting pairs $(Z^j_{t_0}, Z^{j+1}_{t_0})$ for which $T_{t_0}(Z^j_{t_0}, Z^{j+1}_{t_0}) > t$. We define the distance between X_t and Y_t to be

$$D(X_t, Y_t) = \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} d_t(Z^j_{t_0}, Z^{j+1}_{t_0}). \tag{5}$$

It is clear that if $\mathrm{d}(X_t, Y_t) \neq 0$ then $D(X_t, Y_t) \geq (1-\alpha)$. We now prove a contraction lemma analagous to Lemma 2.1.

Lemma 2.5. *Given the stopping times conditions, then for all (X_0, Y_0) and all $t \geq 0$*

$$\mathbb{E}[D(X_{t+k}, Y_{t+k}) \,|\, X_t, Y_t] \leq \left(1 - \frac{(1-\alpha)p}{\gamma+1}\right) D(X_k, Y_k),$$

where γ is the maximum value of $\mathbb{E}[\mathrm{d}(X_{T_0(X,Y)}, Y_{T_0(X,Y)}) \mid \mathcal{F}_t]/\mathrm{d}(X_0, Y_0)$ over all pairs in S and evolutions $\mathcal{F}_t$ such that $t < T_0(X, Y)$.

Proof. The set $\mathcal{SP}_{t+k}$ is the union of the starting pairs from $\mathcal{SP}_t$ which did not reach their stopping time by time $t+k$, and those starting pairs arising from a pair in $\mathcal{SP}_t$ which did stop by time $t+k$. Hence, writing T_{t_0} for $T_{t_0}(Z^j_{t_0} Z^{j+1}_{t_0})$,

$$D(X_{t+k}, Y_{t+k}) \;=\; \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} \mathbb{1}_{T_{t_0} > t+k} d_{t+k}(Z^j_{t_0}, Z^{j+1}_{t_0}) + \mathbb{1}_{T_{t_0} \leq t+k} \sum d_{t+k}(Z^l_{t_l}, Z^{l+1}_{t_l})$$

where the second sum is over starting pairs arising from the stopping of pair $(Z^j_{t_0}, Z^{j+1}_{t_0})$. As in Lemma 2.1, we may assume inductively that $\mathbb{E}[d_{t+k}(Z^l_{t_l}, Z^{l+1}_{t_l}) \mid t < t_l \leq t+k] \leq \mathrm{d}(Z^l_{t_l}, Z^{l+1}_{t_l})$. Then, given $\mathcal{F}_t$, the expected value of $D(X_{t+k}, Y_{t+k})$ is

$$\begin{aligned}
\mathbb{E}[D(X_{t+k}, Y_{t+k})] \;&\leq\; \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} \mathbb{E}[\mathbb{1}_{T_{t_0} > t+k} d_{t+k}(Z^j_{t_0}, Z^{j+1}_{t_0})] + \mathbb{E}[\mathbb{1}_{T_{t_0} \leq t+k} \mathrm{d}(Z^j_{T_{t_0}}, Z^{j+1}_{T_{t_0}})] \\
&\leq\; \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} \mathbb{E}\left[\mathbb{1}_{T_{t_0} > t+k} (1-\alpha) \mathrm{d}(Z^j_{t_0}, Z^{j+1}_{t_0})\right] + \mathbb{E}[\mathrm{d}(Z^j_{T_{t_0}}, Z^{j+1}_{T_{t_0}})]. \qquad (6)
\end{aligned}$$

So subtracting (5) from (6) we get

$$\begin{aligned}
\mathbb{E}[D(X_{t+k}, Y_{t+k})] - D(X_t, Y_t) &\leq \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} -(1-\alpha)\mathbb{E}[\mathrm{d}(Z^j_{t_0}, Z^{j+1}_{t_0}) \mathbb{1}_{T_0 \leq t+k}] \\
&\leq\; -(1-\alpha)p \sum_{(Z^j_{t_0}, Z^{j+1}_{t_0}) \in \mathcal{SP}_t} \mathrm{d}(Z^j_{t_0}, Z^{j+1}_{t_0}) \qquad (7) \\
&\leq\; -\frac{(1-\alpha)p}{1-\alpha+\gamma} D(X_t, Y_t).
\end{aligned}$$

The final inequality follows since, by (4), we have $d_t(Z^j_{t_0}, Z^{j+1}_{t_0}) \leq (1-\alpha+\gamma)\mathrm{d}(Z^j_{t_0}, Z^{j+1}_{t_0})$. □

The γ term arises because although we have contraction in inequality (7), we need to express this as a proportion of $D(X_t, Y_t)$. The expected value at the stopping time is only guaranteed to be at most $\alpha\mathrm{d}$ at the outset. If we have already evolved, possibly adversely, the expected value at the stopping time could be larger than this, and the proportional changes correspondingly smaller. However γ is bounded by the maximum distance (in the original metric) that can occur at the stopping time; in practice this is very likely to be a small constant.

By following the same arguments as in Section 2.1, with this contraction lemma we obtain the following theorem.

Theorem 2.6. *Let $\mathcal{M}$ be a Markov chain on state space Ω. Let d be an integer valued metric on Ω, and let (X_t, Y_t) be a path coupling for $\mathcal{M}$. Let $T(X_0, Y_0)$ be stopping times satisfying the stopping times conditions. Then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies*

$$\tau(\varepsilon) = O\left(\frac{k(1-\alpha+\gamma)}{p(1-\alpha)} \ln\left(\frac{D}{\varepsilon}\right)\right).$$

Remark 2.7. One of the most interesting features of Theorems 2.4 and 2.6 is that their proofs employ only standard path coupling (applied to the k-step chain), but with a metric which has some useful properties. Thus, for any problem to which stopping times might be applied, there exists a metric from which the same result could be obtained using one-step path coupling.

Remark 2.8. Stopping times condition 2 may appear a restriction, but appears to be naturally satisfied in most applications, even with $k = 1$. The alternative, though less natural, assumption of uniformly bounded stopping times [15] is also included. (See Remark 2.9.)

Remark 2.9. We may compare this stopping time theorem with those in [2, 15]. The main result of [15, Theorem 3] concerns bounded stopping times, where $T_0 \leq M$ for all $(X_0, Y_0) \in S$, and gives a mixing time of $O(M(1-\alpha)^{-1} \log D)$. By setting $k = M$ and $p = 1$ in Theorem 2.4, we obtain the same mixing time up to minor changes in constants, but with a proof that does not involve defining a multistep coupling. For unbounded mixing times, [15, Corollary 4] gives a bound $O(\mathbb{E}[T](1-\alpha)^{-2} W \log D)$ by truncating the stopping times, where W denotes the maximum of $\mathrm{d}(X_t, Y_t)$ over all $(X_0, Y_0) \in S$ and $t \leq T$. In most applications $\mathbb{E}[T] \leq k/p$, so in Theorem 2.4 we obtain an improvement of order $W(1-\alpha)^{-1}$. By comparison with [2], we obtain a more modest improvement, of order $\log W \log(D(1-\alpha)^{-1})/\log D$. For the more general stopping times, comparing Theorem 2.6 and [15, Corollary 4], we obtain an improvement of order $\frac{W}{\gamma(1-\alpha)}$. It should be noted that $\gamma \leq W$.

Remark 2.10. Further improvements to Theorem 2.4 seem unlikely, other than in constants. The term k/p must be present, since it bounds a single stopping time. A term $1/(1-\alpha)\log(D/\varepsilon) = \Theta(\log_\alpha(D/\varepsilon))$ also seems essential, since it bounds the number of stopping times required. Likewise improvements to Theorem 2.6 are likely restricted to changing the dependence on γ, although it seems plausible that some dependence is required.

3 Hypergraph independent sets

We now turn our attention to hypergraph independent sets. These were previously studied in [2]. Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be a hypergraph of maximum degree Δ and minimum edge size m. A subset $S \subseteq \mathcal{V}$

of the vertices is *independent* if no edge is a subset of S. Let $\Omega(\mathcal{H})$ be the set of all independent sets of $\mathcal{H}$. We define the Markov chain $\mathcal{M}(\mathcal{H})$ with state space $\Omega(\mathcal{H})$ by the following transition process (*Glauber dynamics*). If the state of $\mathcal{M}$ at time t is X_t, the state at $t+1$ is determined by the following procedure.

1. Select a vertex $v \in \mathcal{V}$ uniformly at random,
2. (i) if $v \in X_t$ let $X_{t+1} = X_t \backslash \{v\}$ with probability 1/2,
   (ii) if $v \notin X_t$ and $X_t \cup \{v\}$ is independent, let $X_{t+1} = X_t \cup \{v\}$ with probability 1/2,
   (iii) otherwise let $X_{t+1} = X_t$.

This chain is easily shown to be ergodic with uniform stationary distribution. The natural coupling for this chain is the "identity" coupling, the same transition is attempted in both copies of the chain. If we try to apply standard path coupling to this chain, we immediately run into difficulties. Consider a state of the coupled chain at a time t, (X_t, Y_t), such that $Y_t = X_t \cup \{w\}$, where $w \notin X_t$ (the *change vertex*) is of degree Δ. An edge $e \in \mathcal{E}$ is *critical* in Y_t if it has only one vertex $z \in \mathcal{V}$ which is not in Y_t, and we call z *critical for* e. If each of the edges through w is critical for Y_t, then there are Δ choices of v in the transition which can be added in X_t but not in Y_t. Thus the change in the expected *Hamming distance* between X_t and Y_t after one step could be as high as $\frac{\Delta}{2n} - \frac{1}{n}$, and we obtain rapid mixing only in the case $\Delta = 2$.

For $(\sigma, \sigma \cup \{w\}) \in S$, let $E_i(w, \sigma)$ be the set of edges containing w which have i occupied vertices in σ. Using a result like Theorem 2.2 above, it is shown in [3] that, for the stopping time T given by the first epoch at which the Hamming distance between the coupled chains changes,

$$\mathbb{E}[\mathrm{d}_{\mathrm{Ham}}(X_T, Y_T | X_0 = \sigma, Y_0 = \sigma \cup \{w\})] \leq 2 \sum_{i=0}^{m-2} p_i |E_i| \leq 2p_1\Delta,$$

where the p_i is the probability that $\mathrm{d}(X_T, Y_T) = 2$ if w is in a single edge with i occupied vertices. Since $p_1 < 1/(m-1)$, we obtain rapid mixing when $2\Delta/(m-1) \leq 1$, i.e. when $m \geq 2\Delta + 1$. See [3] for details.

The approach of Section 2 would lead us to define a metric for which the distance between σ and $\sigma \cup \{w\}$ is $(1 - 2p_1\Delta) + 2\sum_{i=0}^{m-2} p_i|E_i|$. By Lemma 2.1, we know that this metric contracts in expectation. However, prompted by the form of this metric, but retaining the freedom to optimise constants, we will instead define the new metric d to be

$$\mathrm{d}(\sigma, \sigma \cup \{w\}) = \sum_{i=0}^{m-2} c_i |E_i|,$$

where $0 < c_i \leq 1$ $(0 \leq i \leq m-2)$ are a nondecreasing sequence of constants to be determined. Using this metric, we obtain the following theorem.

Theorem 3.1. *Let Δ be fixed, and let $\mathcal{H}$ be a hypergraph such that $m \geq \Delta + 2 \geq 5$, or $\Delta = 3$ and $m \geq 2$. Then the Markov chain $\mathcal{M}(\mathcal{H})$ has mixing time $O(n \log n)$.*

Proof. Without loss of generality, we take $c_{m-2} = 1$ and we will define $c_{-1} = c_0, c_{m-1} \geq \Delta + 1$. Note that c_{-1} has no real role in the analysis, and is chosen only for convenience, but c_{m-1} is chosen so that $c_{m-1} - c_{m-2} \geq \Delta \geq \mathrm{d}(\sigma, \sigma')$ for any pair $(\sigma, \sigma') \in S$. We require $c_i > 0$ for all i so that we will always have $\mathrm{d}(\sigma, \sigma') > 0$ if $\sigma \neq \sigma'$.

Now consider the expected change in distance between σ and $\sigma \cup \{w\}$ after one step of the chain.

If w is chosen, then the distance decreases by $\sum_{i=0}^{m-2} c_i|E_i|$. The contribution to the expected change in distance is $-\frac{2}{2n}\sum_{i=0}^{m-2} c_i|E_i|$.

If we insert a vertex v in an edge containing w, then we increase the distance by $(c_{i+1} - c_i) \geq 0$ for each edge in E_i containing v. This holds for $i = 0, \ldots, m-2$, by the choice of $c_{m-1} = \Delta + 1$. Let U be the set of unoccupied neighbours of w, and $\nu_i(v)$ be the number of edges with i occupants containing w and v. Then the contribution is

$$\sum_{v \in U} \frac{1}{2n} \sum_{i=0}^{m-2} \nu_i(v)(c_{i+1} - c_i) = \frac{1}{2n} \sum_{i=0}^{m-2} (c_{i+1} - c_i)(m - i - 1)|E_i|,$$

since

$$\sum_{v \in U} \nu_i(v) = \sum_{v \in U} \sum_{e \in E_i} \mathbb{1}_{v \in e} = \sum_{e \in E_i} \sum_{v \in e \cap U} 1 = \sum_{e \in E_i} (m - i - 1) = (m - i - 1)|E_i|.$$

If we delete a vertex v in an edge containing w, then we decrease the distance by $(c_i - c_{i-1})$ for each edge in E_i containing v. This holds for $i = 0, \ldots, m-2$, by the choice of c_{-1}. Let O be the set of occupied neighbours of w, and $\nu_i(v)$ be the number of edges with i occupants containing w and v. Then the contribution is

$$-\sum_{v \in O} \frac{1}{2n} \sum_{i=0}^{m-2} \nu_i(v)(c_i - c_{i-1}) = -\frac{1}{2n} \sum_{i=0}^{m-2} (c_i - c_{i-1}) i |E_i|,$$

since, as for U above,

$$\sum_{v \in O} \nu_i(v) = \sum_{v \in O} \sum_{e \in E_i} \mathbb{1}_{v \in e} = \sum_{e \in E_i} \sum_{v \in e \cap O} 1 = \sum_{e \in E_i} i = i|E_i|.$$

Let $\mathrm{d}_0 = \mathrm{d}(\sigma, \sigma \cup \{w\})$, and let d_1 be the distance between the evolved states after one step of the chain. The change in expected distance $\mathbb{E}[\mathrm{d}_1 - \mathrm{d}_0]$ satisfies

$$\begin{aligned}
2n\mathbb{E}[\mathrm{d}_1 - \mathrm{d}_0] \leq \; & -2\sum_{i=0}^{m-2} c_i|E_i| + \sum_{i=0}^{m-2} (c_{i+1} - c_i)(m - i - 1)|E_i| - \sum_{i=0}^{m-2} (c_i - c_{i-1}) i |E_i| \\
= \; & \sum_{i=0}^{m-2} \left(-2c_i + (m - i - 1)(c_{i+1} - c_i) - i(c_i - c_{i-1})\right)|E_i| \\
= \; & \sum_{i=0}^{m-2} \left(ic_{i-1} - (m+1)c_i + (m - i - 1)c_{i+1}\right)|E_i|.
\end{aligned}$$

We require $\mathbb{E}[\mathrm{d}_1 - \mathrm{d}_0] \leq -\gamma$, for some $\gamma \geq 0$, which holds for all possible choices of E_i if and only if $(m-i-1)c_{i+1} - (m+1)c_i + ic_{i-1} \leq -\gamma$ for all $i = 0, 1, \ldots, m-2$. Thus we need a solution to

$$\begin{aligned} &ic_{i-1} - (m+1)c_i + (m-i-1)c_{i+1} \leq -\gamma \qquad (i = 0, \ldots, m-2), \\ &0 = c_{-1} < c_0 \leq c_1 \leq \cdots \leq c_{m-3} \leq c_{m-2} = 1, \\ &c_{m-1} \geq \Delta + 1, \ \gamma \geq 0, \end{aligned} \tag{8}$$

with $\gamma > 0$ if possible. Adding (8) from i to $m-2$ gives

$$\begin{aligned} ic_{i-1} - (m-i)c_i - (m-1)c_{m-2} + c_{m-1} \ &\leq \ -(m-i-1)\gamma \qquad (i = 0, \ldots, m-2), \\ \text{i.e.} \ ic_{i-1} \ &\leq \ (m-i)c_i + (m - \Delta - 2) - (m-i-1)\gamma \qquad (i = 0, \ldots, m-1). \end{aligned} \tag{9}$$

Substitute $u_i = \binom{m-1}{i} c_i$ in (9), so $u_{m-1} \geq \Delta + 1$, $u_{m-2} = m-1$ and $u_{-1} = 0$. Then we have

$$u_{i-1} \leq u_i + \frac{m - \Delta - 2 + \gamma}{m}\binom{m}{i} - \gamma\binom{m-1}{i} \qquad (i = 0, \ldots, m-2).$$

Using the boundary condition $u_{-1} = 0$, these give

$$u_i \leq \gamma \sum_{j=0}^{i} \binom{m-1}{i} - \frac{m - \Delta - 2 + \gamma}{m} \sum_{j=0}^{i} \binom{m}{j} \quad (i = 0, \ldots, m-2).$$

The boundary condition $u_{m-2} = m-1$ now implies

$$\gamma \ \leq \ \frac{2^m - 1 - m}{(m-2)2^{m-1} + 1}\left(m - \Delta - 2 + \frac{m(m-1)}{2^m - 1 - m}\right).$$

Let

$$f(m) = m - 2 + \frac{m(m-1)}{2^m - 1 - m},$$

then we can have $\gamma \geq 0$ if and only if $f(m) \geq \Delta$, and $\gamma > 0$ if and only if $f(m) > \Delta$. Then

$$c_i \ = \ \frac{\gamma \sum_{j=0}^{i} \binom{m-1}{j} - \frac{m - \Delta - 2 + \gamma}{m} \sum_{j=0}^{i} \binom{m}{j}}{\binom{m-1}{i}} \qquad (i = 0, \ldots, m-2).$$

In order to satisfy the conditions of (8), we need to establish that $0 < c_i \leq c_{i+1}$ $(i = 0, \ldots, m-3)$.

$$\begin{aligned} c_i \ &= \ \frac{\gamma \sum_{j=0}^{i} \binom{m-1}{j} - \frac{m - \Delta - 2 + \gamma}{m} \sum_{j=0}^{i} \binom{m}{j}}{\binom{m-1}{i}} \qquad (i = 0, \ldots, m-2) \\ &= \ \gamma \frac{\sum_{j=0}^{i} \binom{m-1}{j} - \kappa \sum_{j=0}^{i} \binom{m}{j}}{\binom{m-1}{i}}, \text{ where } \kappa \ = \ \frac{m - \Delta - 2 + \gamma}{m\gamma}. \end{aligned}$$

$$
\begin{aligned}
&= \gamma \frac{\sum_{j=0}^{i}\binom{m-1}{j} - \kappa \sum_{j=0}^{i}\left(\binom{m-1}{j}+\binom{m-1}{j-1}\right)}{\binom{m-1}{i}}, \\
&= \gamma \frac{\sum_{j=0}^{i}\binom{m-1}{j} - \kappa \left(\sum_{j=0}^{i}\binom{m-1}{j}+\sum_{j=0}^{i-1}\binom{m-1}{j}\right)}{\binom{m-1}{i}}, \\
&= \gamma \frac{\sum_{j=0}^{i}\binom{m-1}{j} - \kappa \left(2\sum_{j=0}^{i}\binom{m-1}{j}-\binom{m-1}{i}\right)}{\binom{m-1}{i}}, \\
&= \gamma \frac{(1-2\kappa)\sum_{j=0}^{i}\binom{m-1}{j} + \kappa\binom{m-1}{i}}{\binom{m-1}{i}}, \\
&= \gamma(1-2\kappa)\frac{\sum_{j=0}^{i}\binom{m-1}{j}}{\binom{m-1}{i}} + \gamma\kappa, \\
&= \gamma(1-2\kappa)g_i + \gamma\kappa, \quad \text{say.}
\end{aligned}
$$

Now $2\kappa < 1$ is equivalent to $2(m-\Delta-2)/(m-2) < \gamma$, i.e.

$$
\frac{2(m-\Delta-2)}{m-2} < \frac{(2^m-1-m)(m-\Delta-2)+m(m-1)}{(m-2)2^{m-1}+1},
$$

which holds for all $\Delta > 0$. Finally, g_i is strictly increasing, since

$$
\begin{aligned}
\frac{g_{i-1}}{g_i} &= \frac{\frac{m-i}{i}\sum_{j=0}^{i-1}\binom{m-1}{j}}{\sum_{j=0}^{i}\binom{m-1}{j}} \\
&= \frac{\sum_{j=1}^{i}\frac{m-i}{i}\binom{m-1}{j-1}}{\sum_{j=0}^{i}\binom{m-1}{j}} \\
&\le \frac{\sum_{j=1}^{i}\binom{m-1}{j}}{\sum_{j=0}^{i}\binom{m-1}{j}}, \quad \text{since } j \le i, \\
&< 1.
\end{aligned}
$$

Hence c_i is strictly increasing. It only remains to verily that $c_0 > 0$. This is clearly equivalent to $\gamma > (m-\Delta-2)/(m-1)$. If $m = \Delta+2$, it follows from $\gamma > 0$. If $m > \Delta+2$, it follows from $\gamma > 2(m-\Delta-2)/(m-2)$, which we have already established.

If $m \ge 5$ then $m(m-1)/(2^m-1-m) < 1$, so we will have $f(m) > \Delta$ exactly when $m \ge \Delta+2$. For smaller values of m,

m	2	3	4
$f(m)$	2	$2\frac{1}{2}$	$3\frac{1}{11}$

The new case here is $\Delta = 3, m \ge 4$. In any case for which $f(m) > \Delta$, standard path coupling arguments yield the mixing times claimed since we have contraction in the metric and the minimum distance is at least c_0. Since we can show mixing for $\Delta = 3, m \le 3$ by other means (see [12]), we have mixing for $\Delta = 3$ and every m. □

Remark 3.2. The independent set problem here has a natural *dual*, that of sampling an *edge cover* from a hypergraph with edge size Δ and degree m. An edge cover is a subset of $\mathcal{E}$ whose union contains V. For the graph case of this sampling problem, with arbitrary m, see [4]. By duality this gives the case $\Delta = 2$ of the independent set problem here.

4 Satisfying assignments of SAT instances

The set of *independent* sets in a hypergraph with edge size m and degree Δ corresponds in a natural way to the set of *satisfying* assignments in a SAT instance with clause size equal to m and number of each variable occurrences bounded by Δ, cf. [12]. The optimisation problems connected to small (variable) occurrence number instances of SAT were studied recently in [1] (see [1] also for additional references).

Given a hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ with n vertices, k hyperedges, and edge size m and degree Δ. We construct an mSAT formula f, over n variables $X = \{x_1, \dots, x_n\}$ corresponding to vertices of $\mathcal{H}$ as follows. If $e = \{v_1, \dots, v_m\}$ is an hyperedge of $\mathcal{H}$, we associate with e a clause $C_e = \bigvee_{i=1}^{m} \bar{x}_i$, and furthermore we set $f = \bigwedge_{e \in \mathcal{E}} C_e$. Notice that the number of satisfying assignments of f is precisely the same as a number of all independent sets of $\mathcal{H}$, and a number of occurrences of variables in f is less than or equal to the degree of $\mathcal{H}$. We can moreover replace the literals $\bar{x}_i$ by x_i, to obtain a monotone mSAT formula f' with the same number of variable occurrences as f and with the same number of satisfying assignments. The above construction is reversable, showing the equivalence of corresponding counting problems of hypergraph independent sets and monotone SAT formulas.

Let us denote by $\#(m, \Delta)\mu$SAT the problem of counting number of satisfying assignments in monotone mSAT instances with at most Δ variable occurrences. Theorem 3.1. yields the first FPRASs (*Fully Polynomial Randomized Approximation Schemes*) for a large class of monotone mSAT formulas.

Theorem 4.1. *Let Δ be fixed, and $m \geq \Delta + 2 \geq 5$, or if $\Delta = 3$ then $m \geq 2$. Then the associated Markov chain $\mathcal{M}(\mathcal{H})$ yields an FPRAS for the $\#(m, \Delta)\mu SAT$ problem.*

The above result improves vastly the hitherto known results for approximate counting the number of satisfying assignments of general monotone SAT formulas.

5 Colouring 3-uniform hypergraphs

In our second application, also from [2], we consider proper colourings of 3-uniform hypergraphs. We again use Glauber dynamics. Our hypergraph $\mathcal{H}$ will have maximum degree Δ, uniform edge size 3, and we will have a set of q colours. For a discussion of the easier problem of colouring hypergraphs with larger edge size see [3]. A colouring of the vertices of $\mathcal{H}$ is proper if no edge is monochromatic. Let $\Omega'(\mathcal{H})$ be the set of all proper q-colourings of $\mathcal{H}$. We define the Markov chain $\mathcal{C}(\mathcal{H})$ with state space $\Omega'(\mathcal{H})$ by the following transition process. If the state of $\mathcal{C}$ at time t is X_t, the state at $t+1$ is determined by

1. selecting a vertex $v \in \mathcal{V}$ and a colour $k \in \{1, 2, \dots, q\}$ uniformly at random,

2. let X'_t be the colouring obtained by recolouring v colour k
3. if X'_t is a proper colouring let $X_{t+1} = X'_t$
   otherwise let $X_{t+1} = X_t$.

This chain is easily shown to be ergodic with the uniform stationary distribution. For some large enough constant Δ_0, it was shown in [3] to be rapidly mixing for $q > 1.65\Delta$ and $\Delta > \Delta_0$, using a stopping times analysis. Here we improve this result, and simplify the proof, by using a carefully chosen metric which is prompted by the new insight into stopping times analyses. If w is the change vertex, the intuition in [3] was that edges which contain both colours of w are initially "dangerous" but tend to become less so after a time. Thus our metric will be a function of the numbers of edges containing w with various relevant colourings.

Theorem 5.1. *Let Δ be fixed, and let $\mathcal{H}$ be a 3-uniform hypergraph of maximum degree Δ. Then if $q \geq \lceil \frac{3}{2}\Delta + 1 \rceil$, the Markov chain $\mathcal{C}(\mathcal{H})$ has mixing time $O(n \log n)$.*

Proof. Consider two proper colourings X and Y differing in a single vertex w. Without loss of generality let the change vertex w be coloured 1 in X and 2 in Y. We will partition the edges $e \in \mathcal{E}$ containing w into four classes E_1, E_2, E_3, E_4, determined by the colouring of $e \setminus \{w\}$, as follows:

$$E_1 : \{1,2\} \qquad E_2 : \{1,i\} \text{ or } \{2,i\} \ \ (2 < i) \qquad E_3 : \{i,i\} \ \ (2 < i) \qquad E_4 : \{i,j\} \ \ (2 < i < j).$$

Instead of using Hamming distance, we will take a new metric defined by

$$\mathrm{d}(X,Y) = \sum_{i=1}^{4} c_i |E_i|,$$

where $1 = c_1 \geq c_2 \geq c_3 \geq c_4 > 0$, and for convenience $c_0 = \Delta + 1$. Note that $\mathrm{d}(X,Y) \leq \Delta$ if X, Y have Hamming distance 1. The diameter is therefore at most Δn in the metric d.

Arguing as in Section 3, we have

$$\begin{aligned} nq\mathbb{E}[\mathrm{d}_1 - \mathrm{d}_0] \ \leq \ & -(q - |E_3|)\big(c_1|E_1| + c_2|E_2| + c_3|E_3| + c_4|E_4|\big) \\ & +|E_1|\big(-2(q-\Delta-1)(c_1-c_2) + 2(c_0-c_1)\big) \\ & +|E_2|\big(-(q-\Delta-2)(c_2-c_4) - (c_2-c_3) + (c_0-c_2) + (c_1-c_2)\big) \\ & +|E_3|\big(-2(q-\Delta-2)(c_3-c_4) + 4(c_2-c_3)\big) \\ & +|E_4|\big(2(c_3-c_4) + 4(c_2-c_4)\big). \end{aligned} \tag{10}$$

If, in (10), we set

$$\begin{aligned} 2(q-\Delta-1)(c_1-c_2) - 2(c_0-c_1) + c_1(q-|E_3|) &= \gamma \\ (q-\Delta-2)(c_2-c_4) + (c_2-c_3) - (c_0-c_2) - (c_1-c_2) + c_2(q-|E_3|) &= \gamma \\ 2(q-\Delta-2)(c_3-c_4) - 4(c_2-c_3) + c_3(q-|E_3|) &= \gamma \\ -2(c_3-c_4) - 4(c_2-c_4) + c_4(q-|E_3|) &= \gamma, \end{aligned} \tag{11}$$

where $\gamma \geq 0$, we have

$$\mathbb{E}[\mathrm{d}_1] \leq \ \mathrm{d}_0 - \frac{\gamma\Delta}{nq} \leq \Big(1 - \frac{\gamma}{nq}\Big)\mathrm{d}_0. \tag{12}$$

Note, that if we put $q' = q - |E_3|$, $\Delta' = \Delta - |E_3|$ in (11), we have

$$\begin{aligned} 2(q' - \Delta' - 1)(c_1 - c_2) - 2(c_0 - c_1) + c_1 q' \quad &= \ \gamma \\ (q' - \Delta' - 2)(c_2 - c_4) + (c_2 - c_3) - (c_0 - c_2) - (c_1 - c_2) + c_2 q' \quad &= \ \gamma \\ 2(q' - \Delta' - 2)(c_3 - c_4) - 4(c_2 - c_3) + c_3 q' \quad &= \ \gamma \\ -2(c_3 - c_4) - 4(c_2 - c_4) + c_4 q' \quad &= \ \gamma. \end{aligned} \tag{13}$$

This corresponds to a system like (11) with degree Δ', q' colours and $|E_3| = 0$. But, since $q'/\Delta' = (q - |E_3|)/(\Delta - |E_3|) \geq q/\Delta$, the smallest ratio for q/Δ is given by setting $|E_3| = 0$ in (11). Also, putting $c_3 = c_4$ makes the third and fourth equations in (13) identical, so $c_3 = c_4$ must be a solution. With these simplifications, and putting $c_0 = \Delta + 1$, $c_1 = 1$, we have

$$\begin{aligned} 2(q - \Delta - 1)(1 - c_2) - 2\Delta + q \quad &= \ \gamma \\ (q - \Delta - 1)(c_2 - c_4) - 2(1 - c_2) - \Delta + c_2 q \quad &= \ \gamma \\ -4(c_2 - c_4) + c_4 q \quad &= \ \gamma. \end{aligned}$$

Now the linear equations (11) may be solved for c_2, c_4 and γ, giving

$$c_1 = 1, \quad c_2 = \frac{2q - 2\Delta + 1}{2q - \Delta + 1}, \quad c_3 = c_4 = \frac{2q - 3\Delta + 1}{2q - \Delta + 1}, \quad \gamma = \frac{2q^2 - q(3\Delta - 1) - 4\Delta}{2q - \Delta + 1}.$$

The condition $\gamma \geq 0$ is equivalent to

$$q \ \geq \ \tfrac{3\Delta - 1}{4}\Big(1 + \sqrt{1 + \tfrac{32\Delta}{(3\Delta-1)^2}}\Big), \qquad \text{i.e. } \ q \ \geq \lceil \tfrac{3}{2}\Delta \rceil + 1.$$

Note that we have $c_i > 0$ ($i = 1, \ldots, 4$) under this condition. Note also that $\gamma > 0$ and hence, using (12), the mixing time satisfies

$$\tau(\varepsilon) \ \leq \ \frac{2q^2 - q\Delta + q}{2q^2 - q(3\Delta - 1) - 4\Delta} \ n \ln\Big(\frac{\Delta n}{\varepsilon}\Big). \qquad \square$$

6 Colouring bipartite graphs

Our final application is to colouring bipartite graphs. Several recent papers have used a stopping times or "burn in" analysis to prove rapid mixing for Glauber dynamics of graph colouring, starting with [8]. These are largely based upon the idea that although a vertex can have only $q - \Delta$ colours with which to be properly recoloured, it is very unlikely for any vertex to have so few colours available after a period of "burn in". Subject to more stringent girth and degree restrictions than used here, rapid mixing has been proved for fewer colours [9, 14, 19]. Here we capture this intuition by using a metric which directly incorporates the number of colours available to a vertex. In order to simplify the analysis, we do not consider Glauber dynamics here. Instead we prove that a Markov chain SCAN which uses the same method for recolouring a vertex as Glauber dynamics, but recolours the vertices in a deterministic order, mixes rapidly. In order to show this we first prove results for a closely related Markov chain, MULTICOLOUR, which is of interest in its own right.

Let $G = (V, E)$ be a bipartite graph with bipartition V_1, V_2, and maximum degree Δ. For $v \in V$, let $\mathcal{N}(v) = \{w : \{v, w\} \in E\}$ denote the neighbourhood of v. Let $Q = [q]$ be a colour set, and $X : V \to Q$ be a colouring of G, not necessarily proper. Let $C(v) = \{X(w) : w \in \mathcal{N}(v)\}$ be the set of colours occurring in the neighbourhood of v, and $c(v) = |C(v)|$. We consider the Markov chain MULTICOLOUR on colourings of G, which in each step picks one side of the bipartition at random, and then recolours every vertex on that side, followed by recolouring every vertex in the other half of the bipartition. If the state of MULTICOLOUR at time t is X_t, the state at time $t + 1$ is given by

MULTICOLOUR

1. choosing $r \in \{1, 2\}$ uniformly at random,
2. for each vertex $v \in V_r$,
   (i) choosing a colour $q(v) \in Q\backslash C(v)$ uniformly at random,
   (ii) setting $X_{t+1}(v) = q(v)$. (Heat bath recolouring)
3. for each vertex $v \in V\backslash V_r$,
   (i) choosing a colour $q(v) \in Q\backslash C(v)$ uniformly at random,
   (ii) setting $X_{t+1}(v) = q(v)$.

Note that the order in which the vertices are processed in steps 2 and 3 is immaterial. This chain is a single-site dynamics intermediate between Glauber and scan. It is easy to see that it is ergodic if $q > \Delta + 1$, and has equilibrium distribution uniform on all proper colourings of G. Observe also that it requires considerably fewer random bits than Glauber, and only slightly more than scan. We prove the following theorem.

Theorem 6.1. *For $q > f(\Delta)$ the mixing times of* SCAN *and* MULTICOLOUR *are $O(\log(n))$, where f is a function such that*

1. *$f(\Delta) \to \beta\Delta$, as $\Delta \to \infty$, where β satisfies $\frac{1}{\beta}e^{1/\beta} = 1$,*
2. *$f(\Delta) \leq \lceil 11\Delta/6 \rceil$ for $\Delta \geq 14$,*
3. *$f(\Delta) < \lceil 11\Delta/6 \rceil$ for $\Delta \geq 31$,*
4. *in particular $f(22) = 40 < \lceil 11\Delta/6 \rceil$.*

We will require the following lemmas.

Lemma 6.2. *For $1 \leq i \leq \Delta$ let S_i be a subset of $(Q - q_0)$ such that $m_i = |S_i| \geq q - \Delta$. Let s_i be selected uniformly at random from S_i, independently for each i. Finally let $C = \{s_i : 1 \leq i \leq \Delta\}$ and $c = |C|$. Then*

$$\mathbb{E}[q - c \mid s_1 = q_1] \geq 1 + (q - 2)\left(1 - \frac{1}{q - \Delta}\right)^{\frac{(\Delta-1)(q-\Delta)}{q-2}} = \alpha.$$

Proof. This follows from [8, Lemma 2.1] with minor adjustments as follows. Let $a_{ij} = 1$ if $j \in S_i$ and 0 otherwise. Thus $m_i = \sum_{j \in (Q-q_0)} a_{ij}$ and

$$\mathbb{E}[q-c] = 1 + \sum_{j \in (Q-q_0)} \prod_{i=1}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{a_{ij}}.$$

However if we are given that $s_1 = q_1$, then

$$\begin{aligned}
\mathbb{E}[q-c \mid s_1 = q_1] = & \ 1 + \sum_{j \in (Q-q_0-q_1)} \prod_{i=2}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{a_{ij}} \\
\geq & \ 1 + (q-2) \left(\prod_{j \in (Q-q_0-q_1)} \prod_{i=2}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{a_{ij}} \right)^{\frac{1}{q-2}} \\
\geq & \ 1 + (q-2) \left(\prod_{i=2}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{m_i} \right)^{\frac{1}{q-2}} \\
\geq & \ 1 + (q-2) \left(1 - \frac{1}{q-\Delta}\right)^{\frac{(\Delta-1)(q-\Delta)}{q-2}}.
\end{aligned}$$

Where the final inequality follows because $(1-1/m_i)^{m_i}$ in increasing with m_i and $m_i \geq q - \Delta$ for all i. □

Lemma 6.3. *For $1 \leq i \leq \Delta$ let S_i be a subset of $(Q - q_0)$ such that $m_i = |S_i| \geq q - \Delta$. Let s_i be selected uniformly at random from S_i, independently for each i. Finally let $C = \{s_i : 1 \leq i \leq \Delta\}$ and $c = |C|$. Then*

$$\mathbb{E}\left[\frac{1}{q-c} \mid s_1 = q_1\right] \leq \frac{1}{\alpha}\left(1 + \frac{(q-\alpha-1)(\alpha-1)}{(q-\Delta)(q-2)\alpha}\right) = \alpha'.$$

Proof. We will write $\bar{c}$ for $\mathbb{E}[c \mid s_1 = q_1]$. Let $Z = \frac{c-\bar{c}}{q-c}$, so that

$$\frac{1}{q-c} = \frac{1}{q-\bar{c}}\left(\frac{1}{1-Z}\right). \tag{14}$$

Note that $(1-Z)^{-1} = \frac{q-\bar{c}}{q-c} \leq \frac{q-\bar{c}}{q-\Delta}$. Now

$$\frac{1}{1-Z} = 1 + Z + \frac{Z^2}{1-Z} \leq 1 + Z + \frac{(q-\bar{c})Z^2}{q-\Delta}.$$

Hence

$$\mathbb{E}[(1-Z)^{-1} \mid s_1 = q_1] \leq 1 + \frac{q-\bar{c}}{q-\Delta}\frac{\mathrm{Var}(c \mid s_1 = q_1)}{(q-\bar{c})^2} = 1 + \frac{\mathrm{Var}(c \mid s_1 = q_1)}{(q-\Delta)(q-\bar{c})}. \tag{15}$$

We now turn our attention to bounding $\mathrm{Var}(c \mid s_1 = q_1)$. Let $c = \sum_{j\in(Q-q_0)} I_j$, where I_j indicates that colour j is in C. Now, conditional on $s_1 = q_1$, we have

$$\mathrm{Var}(\sum_{j\in(Q-q_0)} I_j) = \sum_{j\in(Q-q_0)} \mathrm{Var}(I_j) + 2\sum_{j<k} \mathrm{Cov}(I_j, I_k) \le \sum_{j\in(Q-q_0)} \mathrm{Var}(I_j),$$

since I_j and I_k are negatively correlated for all j and k. Let $p_j = \Pr(I_j = 1)$, then I_j has variance $p_j(1-p_j)$ and $\sum_{j\in(Q-q_0)} p_j = \bar{c}$. Also note that $p_{q_1} = 1$, hence $\mathrm{Var}(I_{q_1}) = 0$. By convexity, the maximum of $\sum_{j\in(Q-q_0-q_1)} p_j(1-p_j)$ such that $\sum_{j\in(Q-q_0-q_1)} p_j = \bar{c}-1$ is given by setting $p_j = (\bar{c}-1)/(q-2)$. Hence, using $\bar{c} = q - \alpha$,

$$\mathrm{Var}(c \mid s_1 = q_1) \le (\bar{c}-1)\left(1 - \frac{\bar{c}-1}{(q-2)}\right) = \frac{(q-\alpha-1)(\alpha-1)}{q-2}. \tag{16}$$

Putting together equations (14), (15) and (16) we have

$$\mathbb{E}\left[\frac{1}{q-c} \mid s_1 = q_1\right] \le \frac{1}{\alpha}\left(1 + \frac{(q-\alpha-1)(\alpha-1)}{(q-2)}\frac{1}{(q-\Delta)\alpha}\right).$$

□

Proof of Theorem 6.1. We first prove the theorem for MULTICOLOUR. In the path coupling setting, we will take S to be the set of pairs colourings which differ at exactly one vertex. Let v be the change vertex for some pair $(X, Y) \in S$, and assume without loss that $v \in V_1$. The distance between X and Y is defined to be $\mathrm{d}(X, Y) = \sum_{w\in\mathcal{N}(v)} \frac{1}{q-c_{X,Y}(w)}$, where $c_{X,Y}(w)$ is taken to be $\min\{c_X(w), c_Y(w)\}$ in the case that they differ. We couple as follows (the usual path coupling for Glauber dynamics). If we are recolouring a vertex which is not a neighbour of v, then the sets of available colours in X and Y are the same, and we use the same colour in both copies of the chain. If we are recolouring a vertex $w \in \mathcal{N}(v)$ then there are three cases to consider:

1. $|\{X(v), Y(v)\} \cap \{X(z) : z \in \mathcal{N}(w)\backslash\{v\}\}| = 2$.
   The colours $X(v)$ and $Y(v)$ are not available for recolouring w in either copy of the chain, hence the sets of available colours are the same, and we use the same colour in both copies of the chain.

2. $|\{X(v), Y(v)\} \cap \{X(z) : z \in \mathcal{N}(w)\backslash\{v\}\}| = 1$.
   Without loss assume colour $X(v)$ is not available to w in either copy of the chain. Colour $Y(v)$ is only available in X. We couple recolouring w in X with any colour other than $Y(v)$, with recolouring using the same colour in Y. We couple recolouring w in X with colour $Y(v)$, uniformly between recolouring w with each available colour in Y.

3. $|\{X(v), Y(v)\} \cap \{X(z) : z \in \mathcal{N}(w)\backslash\{v\}\}| = 0$.
   Here colour $Y(v)$ is only available in chain X, and $X(v)$ in only available in Y. We couple together recolouring with these colours respectively, and for each other colour (that is available to both copies), we recolour w with the same colour in both X and Y.

Note that in case 1, there is no probability of w being coloured differently in the two chains. In the other cases, the probability of disagreement at w is $\frac{1}{q-c_{X,Y}(w)}$.

Let X', Y' be the colourings after recolouring V_r (half a step of MULTICOLOUR) and X'', Y'' be the colourings after the full step of MULTICOLOUR. If we randomly select V_1 to be recoloured first, then the two copies of the chain have coupled in X' and Y' since the vertices in V_1 have the same set of available colours in each chain.

So suppose that we select V_2 to be recoloured first. The only vertices in V_2 that have different sets of available colours are those which are neighbours of v. Let $\mathcal{N}(v) = \{w_1, \dots, w_k\}$ and consider the path $W_0, W_1, \dots, W_{k+1}$ from X' to Y', where for $1 \le i \le k$, W_i agrees with X' on all vertices except $w_1, \dots, w_i$ which are coloured as in Y', and $W_0 = X'$ and $W_{k+1} = Y'$. Then for $i \le k$ we have

$$\mathrm{d}(W_{i-1}, W_i) = \mathbb{1}_{w_i} \sum_{z \in \mathcal{N}(w_i)} \frac{1}{q - c_{W_{i-1},W_i}(z)} \le \mathbb{1}_{w_i} \sum_{z \in \mathcal{N}(w_i)} \frac{1}{q - c_{W_i}(z)}, \tag{17}$$

where $\mathbb{1}_{w_i}$ indicates whether X' and Y' differ on w_i. Note that $\Pr[\mathbb{1}_{w_i} = 1] \le \frac{1}{q-c_{X,Y}(w_i)}$. Furthermore, by the construction of the coupling either conditioning on $\mathbb{1}_{w_i} = 1$ is the same as conditioning that $W_{i-1}(w_i) = q_1$, or that $W_i(w_i) = q_1$, for some q_1. We assume without loss that this is W_i. Then for each $z \in \mathcal{N}(w_i) - v$ the selection of colours in $C_{W_i}(z)$ satisfies the conditions of Lemma 6.3, since we may take $q_0 = X(z)$ and q_1 as above. For v, there is no colour q_0 which is necessarily unavailable for all its neighbours, since some are coloured as in X' and some as in Y'. Hence we use a slightly weaker bound on α and α', given by

$$\alpha_v = (q-1)\left(1 - \frac{1}{q-\Delta}\right)^{\frac{(\Delta-1)(q-\Delta)}{q-1}} \quad \text{and} \quad \alpha'_v = \frac{1}{\alpha_v}\left(1 + \frac{(q-\alpha_v)(\alpha_v)}{(q-\Delta)(q-1)\alpha_v}\right).$$

Hence for $i \le k$, $\mathbb{E}[\mathrm{d}(W_{i-1}, W_i)] \le \frac{1}{q-c_{X,Y}(w_i)}((\Delta-1)\alpha' + \alpha'_v)$. The value of $\mathrm{d}(W_k, W_{k+1})$ is still $\mathrm{d}(X, Y)$ since the vertices in V_1 have not yet been recoloured.

Now we consider the vertices in V_1. We apply the same analysis as above to each path segment W_{i-1}, W_i, but augment the analysis using the fact that at the time a vertex $z \in V_1$ is recoloured, its neighbours (in V_2) will already have been randomly recoloured. Let the neighbours of w_i be $z_1, z_2, \dots z_l$, and consider the path $Z_0, Z_1, \dots Z_{l+1}$, where for $1 \le j \le l$, Z_j agrees with W_{i-1} on all vertices except $z_1, \dots, z_j$ which are coloured as in W_i, and $Z_0 = W_{i-1}$ and $Z_{l+1} = W_i$. Arguing as above, for $j \le l$ we have

$$\mathrm{d}(Z_{j-1}, Z_j) = \mathbb{1}_{z_j} \sum_{w \in \mathcal{N}(z_j)} \frac{1}{q - c_{Z_{i-1},Z_i}(w)}.$$

But now $\Pr[\mathbb{1}_{z_j} = 1 |\, W_{i-1}, W_i] \le \frac{1}{q-c_{W_{i-1},W_i}(z_j)} \mathbb{1}_{w_i}$. This is similar to equation (17), and the same argument gives $\mathbb{E}[\mathbb{1}_{z_j} = 1] \le \frac{1}{q-c_{X,Y}(w_i)}\alpha'$, for $z_j \ne v$ and $\mathbb{E}[\mathbb{1}_{z_j} = 1] \le \frac{1}{q-c_{X,Y}(w_i)}\alpha'_v$ if $z_j = v$. Also, since it depends only on the colouring of V_2, we have $\mathrm{d}(Z_l, Z_{l+1}) = \mathrm{d}(W_{i-1}, W_i)$. So

$$\mathbb{E}[\sum_{j=1}^{l+1} \mathrm{d}(Z_{j-1}, Z_j)] \le \frac{1}{q - c_{X,Y}(w_i)}((\Delta-1)\alpha' + \alpha'_v)(((\Delta-1)\alpha' + \alpha'_v) + 1).$$

Δ	q	$\lceil 11\Delta/6 \rceil$	q/Δ
9	17	17	1.89
10	19	19	1.90
11	21	21	1.91
12	23	22	1.92
13	25	24	1.92
14	26	26	1.86
15	28	28	1.87
16	30	30	1.88
17	32	32	1.88
18	33	33	1.83
19	35	35	1.84
20	37	37	1.85
21	39	39	1.86
22	40	41	1.82
23	42	43	1.83
24	44	44	1.83
25	46	46	1.84
26	48	48	1.85
27	49	50	1.81
28	51	52	1.82
29	53	54	1.83
30	55	55	1.83
31	56	57	1.81
32	58	59	1.81
33	60	61	1.82
34	61	63	1.79
35	63	65	1.80
36	65	66	1.81
37	67	68	1.81
38	68	70	1.79
39	70	72	1.79
40	72	74	1.80
41	74	76	1.80
42	75	77	1.79
43	77	79	1.79
44	79	81	1.80
45	81	83	1.80
46	83	85	1.80
47	84	87	1.79
48	86	88	1.79
49	88	90	1.80
50	90	92	1.80
10000	17634	18334	1.76

Table 1: Minimum values of q for contraction.

Finally note that W_k and W_{k+1} differ only in V_1, so after recolouring V_1 they have coupled. Hence

$$\mathbb{E}[\mathrm{d}(X'', Y'')] = \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{l+1}\mathbb{E}[\mathrm{d}(Z_{j-1}, Z_j)] \tag{18}$$

$$\leq \frac{1}{2}\sum_{i=1}^{k}\frac{1}{q - c_{X,Y}(w_i)}((\Delta-1)\alpha' + \alpha'_v)(((\Delta-1)\alpha' + \alpha'_v) + 1) \tag{19}$$

$$= \mathrm{d}(X,Y)((\Delta-1)\alpha' + \alpha'_v)\frac{(((\Delta-1)\alpha' + \alpha'_v) + 1)}{2}. \tag{20}$$

This gives contraction as long as $((\Delta - 1)\alpha' + \alpha'_v)$ is less than 1. For large Δ, we see that α' and α'_v both approach $\frac{1}{q}e^{\Delta/q}$. Hence we have contraction when $\frac{\Delta}{q}e^{\Delta/q} < 1$. For small values of Δ it is possible to compute the smallest integral value of q for which there is contraction. These values are shown in Table 1. When there is contraction, standard path coupling arguments give the mixing time bounds claimed.

We now argue that SCAN mixes as rapidly as MULTICOLOUR. The Markov chain SCAN recolours the two sides of the bipartition in order, $(V_1, V_2), (V_1, V_2)\ldots.$ The Markov chain MULTICOLOUR recolours a random side first in each step. However, recolouring the same side twice in a row has exactly the same effect as recolouring it once, since vertices in the same side of the bipartition are independent. The recolouring given by a run of MULTICOLOUR with order $(V_1, V_2), (V_2, V_1), (V_1, V_2)$ has exactly the same result as if the reversed pair was omitted. Hence any randomly chosen sequence can be replaced with a purely alternating sequence. Should the purely alternating sequence corresponding to the random choices of MULTICOLOUR start with V_2 or finish with V_1, we can augment the sequence with a recolouring of V_1 at the beginning or V_2 at the end respectively. The result follows, since the former is equivalent to taking a different starting position in MULTICOLOUR, and the latter cannot increase the total variation distance from stationarity. □

Remark 6.4. Our analysis shows that one-step analysis of a single-site chain on graph colourings need not break down at $q = 2\Delta$ [16, 21]. This apparent "boundary" seems merely to be an artefact of using Hamming distance.

Remark 6.5. Our scan chain can be used to prove polynomial mixing time for the Glauber dynamics (with the same values of q and Δ) by comparison techniques [7, 20]. However, the proof is not completely straightforward and will appear elsewhere.

Remark 6.6. We note that many of the infinite graphs studied in statistical physics are bipartite, for example cubic grids and trees. Therefore our results imply, for example, absence of phase transition in the antiferromagnetic Potts model in the cubic grid with q colours and dimension $d = \Delta/2$. A proof follows the lines of that given by Vigoda [22, §5] with obvious modifications. Since results with similar q, d have been proved by different arguments in [13], we omit the details.

References

[1] P. Berman, M. Karpinski and A. D. Scott, Approximation hardness of short symmetric instances of MAX-3SAT, Elec. Coll. on Comp. Compl. ECCC TR03-049 (2003); submitted to *Discrete Applied Mathematics*.

[2] M. Bordewich, M. Dyer and M. Karpinski, Path coupling using stopping times, *Proceedings of the 15th International Symposium on Fundamentals of Computation Theory*, Springer Lecture Notes in Computer Science **3623**, pp. 19–31, 2005. (Full version appears as [3].)

[3] M. Bordewich, M. Dyer and M. Karpinski, Path coupling using stopping times and counting independent sets and colourings in hypergraphs, `http://arxiv.org/abs/math.PR/0501081`, 2005.

[4] R. Bubley and M. Dyer, Graph orientations with no sink and an approximation for a hard case of #SAT, in *Proc. 8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 1997)*, SIAM, 1997, pp. 248–257.

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

[6] R. Bubley and M. Dyer, Faster random generation of linear extensions, in *Proc. 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM-SIAM, pp. 350–354, 1998.

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

[8] M. Dyer and A. Frieze, Randomly colouring graphs with lower bounds on girth and maximum degree, *Random Structures and Algorithms* **23** (2003), 167–179.

[9] M. Dyer, A. Frieze, T. Hayes and E. Vigoda, Randomly coloring constant degree graphs, in *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004, pp. 582–589.

[10] M. Dyer, L. Goldberg, C. Greenhill, M. Jerrum and M. Mitzenmacher, An extension of path coupling and its application to the Glauber dynamics for graph colorings, *SIAM Journal on Computing* **30** (2001), 1962–1975.

[11] M. Dyer and C. Greenhill, Random walks on combinatorial objects, in *Surveys in Combinatorics* (J. D. Lamb and D. A. Preece, Eds.), London Mathematical Society Lecture Note Series **267**, Cambridge University Press, Cambridge, 1999, pp. 101–136.

[12] M. Dyer and C. Greenhill, On Markov chains for independent sets, *Journal of Algorithms* **35** (2000), 17–49.

[13] L. Goldberg, R. Martin and M. Paterson, Strong spatial mixing for lattice graphs with fewer colours, in *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004, pp. 562–571.

[14] T. Hayes and E. Vigoda, A Non-Markovian Coupling for Randomly Sampling Colorings, in *Proc. 44th Annual IEEE Symposium on Foundations of Computer Science*, 2003.

[15] T. Hayes and E. Vigoda, Variable length path coupling, in *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM-SIAM, 2004, pp. 103–110.

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

[17] M. Luby and E. Vigoda, Fast convergence of the Glauber dynamics for sampling independent sets, *Random Structures and Algorithms* **15** (1999), 229–241.

[18] M. Mitzenmacher and E. Niklova, Path coupling as a branching process, unpublished manuscript, 2002.

[19] M. Molloy, The Glauber dynamics on the colourings of a graph with large girth and maximum degree, in *Proc. 34th ACM Symposium on Theory of Computing*, 2002.

[20] D. Randall and P. Tetali, Analyzing Glauber dynamics by comparison of Markov chains, *Journal of Mathematical Physics* **41** (2000), 1598-1615.

[21] J. Salas and A. Sokal, Absence of phase transition for anti-ferromagnetic Potts models via the Dobrushin uniqueness theorem, *Journal of Statistical Physics* **86** (1997), 551–579.

[22] E. Vigoda, Improved bounds for sampling colorings, *Journal of Mathematical Physics* **41** (1999), 1555–1569.