Path Coupling Using Stopping Times

Magnus Bordewich[1], Martin Dyer[1], and Marek Karpinski[2]

[1] School of Computing, University of Leeds, Leeds LS2 9JT, UK
{dyer, magnusb}@comp.leeds.ac.uk
[2] Dept. of Computer Science, University of Bonn, 53117 Bonn, Germany
marek@cs.uni-bonn.de

Abstract. We analyse the mixing time of Markov chains using path coupling with stopping times. We apply this approach to two hypergraph problems. We show that the Glauber dynamics for independent sets in a hypergraph mixes rapidly as long as the maximum degree Δ of a vertex and the minimum size m of an edge satisfy $m \geq 2\Delta+1$. We also state results that the Glauber dynamics for proper q-colourings of a hypergraph mixes rapidly if $m \geq 4$ and $q > \Delta$, and if $m = 3$ and $q \geq 1.65\Delta$. We give related results on the hardness of exact and approximate counting for both problems.

1 Introduction

In this paper, we develop a new approach to using path coupling with stopping times to bound the convergence of time of Markov chains. Our interest is in applying these results to randomised approximate counting. For an introduction, see [16]. We illustrate our methods by considering the approximation of the numbers of independent sets and q-colourings in hypergraphs with upper-bounded degree Δ, and lower-bounded edge size m. These problems in hypergraphs are of interest in their own right but, while approximate optimisation in this setting has received considerable attention [5,6,13,17], there has been surprisingly little work on approximate counting other than in the graph case $m = 2$. The tools we develop here may also allow study of approximate counting for several related hypergraph problems. Note, for example, that independent sets in hypergraphs correspond to *edge covers* under hypergraph duality.

Our results are achieved by considering, in the context of path coupling [4], the stopping time at which the distance first changes between two coupled chains. The first application of these stopping times to path coupling was given by Dyer, Goldberg, Greenhill, Jerrum and Mitzenmacher [9]. Their analysis was later improved by Hayes and Vigoda [12], using results closely related to those which we develop in this paper. Mitzenmacher and Niklova [19] had earlier stated a similar theorem, but could not provide a conclusive proof. Their approach had many similarities to our Theorem 2.1, but conditioning problems necessitate a proof along somewhat different lines.

Our main technical result, Theorem 2.1, shows that the chain mixes rapidly if the expected distance between the two chains has decreased at this stopping time.

M. Liśkiewicz and R. Reischuk (Eds.): FCT 2005, LNCS 3623, pp. 19–31, 2005.

We note that a similar conclusion follows from [12, Corollary 4]. However, we give a simpler and more transparent proof of this result, without initially assuming bounded stopping times as is done in the approach of [12]. As a consequence, our Theorem 2.1 will usually give a moderate improvement in the mixing time bound in comparison with [12, Corollary 4]. See Remark 2.3 below.

The problem of approximately counting independent sets in graphs has been widely studied, see for example [8,10,18,20,22], but the only previous work on the approximate counting of independent sets in *hypergraphs* seems to that of Dyer and Greenhill [10]. They showed rapid mixing to the uniform distribution of a simple Markov chain on independent sets in a hypergraph with maximum degree 3 and maximum edge size 3. However, this was the only interesting case resolved. Their results imply rapid mixing only for $m \leq \Delta/(\Delta-2)$, which gives $m \leq 3$ when $\Delta = 3$ and $m \leq 2$ when $\Delta \geq 4$. In Theorem 3.1 we prove rapid mixing of a simple Markov chain, the *Glauber dynamics*, for any hypergraph such that $m \geq 2\Delta{+}1$, where m is the smallest edge size and Δ is the maximum degree. This is a marked improvement for large m. More generally, we consider the *hardcore distribution* on independent sets with *fugacity* λ. (See, for example, [10,18,22].) In [10], it is proved that rapid mixing occurs if $\lambda \leq m/((m-1)\Delta - m)$. Here we improve this considerably for larger values of m, to $\lambda \leq (m-1)/2\Delta$. We also give two hardness results: that computing the number of independent sets in hypergraphs is #P-complete except in trivial cases, and that there can be no approximation for the number of independent sets in a hypergraphs if the minimum edge size is at most logarithmic in Δ. It may be noted that our upper and lower bounds are exponentially different. We have no strong belief that either is close to the threshold at which approximate counting becomes possible, if such a threshold exists.

Counting q-colourings of hypergraphs was considered by Bubley [3], who showed that the Glauber dynamics was rapidly mixing if $q \geq 2\Delta$, generalising a result for graphs of Jerrum [15] and Salas and Sokal [21]. Much work has been done on improving this result for graph colourings, see [7] and its references, but little attention appears to have been given to the hypergraph case. Here we prove rapid mixing of Glauber dynamics for proper colourings of hypergraphs if $m \geq 4$, $q > \Delta$, and if $m = 3$, $q \geq 1.65\Delta$. For a precise statement of our result see Theorem 5.2. We give hardness results showing that computing the number of colourings in hypergraphs is #P-complete except in trivial cases, and that there can be no approximation for the number of colourings of hypergraphs if $q \leq (1-1/m)\Delta^{1/(m-1)}$. Again, there is a considerable discrepancy between the upper and lower bounds for large m.

The paper is organised as follows. Section 1.1 gives an intuitive motivation for the stopping time approach of the paper. Section 2 contains the full description and proof of Theorem 2.1 for path coupling with stopping times. We apply this to hypergraph independent sets in Section 3. Section 4 contains the related hardness results. Section 5 contains our results on the Glauber dynamics for hypergraph colouring. Finally, Section 6 gives the hardness results for counting colourings in hypergraphs. For reasons of space, most of the proofs are omitted in Sections 4–6. These may be found in [1].

1.1 Intuition

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 $\mathcal{H}$. Let λ be the *fugacity*, which weights independent sets. (See [10].) The most important case is $\lambda = 1$, which weights all independent sets equally and gives rise to the uniform distribution on all independent sets. 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.

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

This chain is easily shown to be ergodic with stationary probability proportional to $\lambda^{|I|}$ for each independent set $I \subseteq \mathcal{V}$. In particular, $\lambda = 1$ gives the uniform 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 two chains X_t and 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, if $\lambda = 1$, the change in the expected Hamming distance between X_t and Y_t after one step could be as high as $\frac{\Delta}{2n} - \frac{1}{n}$. Thus we obtain rapid mixing only in the case $\Delta = 2$. This case has some intrinsic interest, since the complement of an independent set corresponds, under hypergraph duality, to an *edge cover* [11] in a graph. Thus we may uniformly generate edge covers, but the scope for unmodified path coupling is obviously severely limited.

The insight on which this paper is based is as follows. Although in one step it could be more likely that a *bad vertex* (increasing Hamming distance) is chosen than a *good vertex* (decreasing Hamming distance), it is even more likely that one of the other vertices in an edge containing w is chosen and removed from the independent set. Once the edge has two unoccupied vertices other than w, then any vertex in that edge can be added in both chains. This observation enables us to show that, if T is defined to be the stopping time at which the distance between X_t and Y_t first changes, the expected distance between X_T and Y_T will be less than 1. Theorem 2.1 below shows that under these circumstances path coupling can easily be adapted to prove rapid mixing.

Having established this general result, we use it to prove that $\mathcal{M}(\mathcal{H})$ is rapidly mixing for hypergraphs with $m \geq 2\lambda\Delta + 1$. Note that, though all the results in this paper will be proved for uniform hypergraphs of edge size m, they carry through trivially for hypergraphs of minimum edge size m.

2 Path Coupling Using a Stopping Time

First we prove the main result discussed above.

Theorem 2.1. *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 $\mathcal{M}$, where S is the set of pairs of states (X, Y) such that $\mathrm{d}(X, Y) = 1$. For any initial states $(X_0, Y_0) \in S$ let T be the stopping time given by the minimum t such that $\mathrm{d}(X_t, Y_t) \neq 1$. Suppose, for some $p > 0$, that*

(i) $\Pr(T = t \,|\, T \geq t) \geq p$, *independently for each* t,
(ii) $\mathbf{E}[\mathrm{d}(X_T, Y_T)] \leq \alpha < 1$.

Then $\mathcal{M}$ mixes rapidly. In particular the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies

$$\tau(\varepsilon) \;\leq\; \frac{1}{p}\,\frac{3}{1-\alpha}\ln(eD_2)\ln\Big(\frac{2D_1}{\varepsilon(1-\alpha)}\Big),$$

where $D_1 = \max\{\mathrm{d}(X, Y) : X, Y \in \Omega\}$ and $D_2 = \max\{\mathrm{d}(X_T, Y_T) : X_0, Y_0 \in \Omega,\, \mathrm{d}(X_0, Y_0) = 1\}$.

Proof. Consider the following game. In each round a gambler either wins £1, loses some amount £$(l-1)$ or continues to the next round. If he loses £$(l-1)$ in a game, he starts l separate (but possibly dependent) games simultaneously in an effort to win back his money. If he has several games going and loses one at a certain time, he starts l more games, while continuing with the others that did not conclude. We know that the probability he finishes a game in a given step is at least p, and the expected winnings in each game is at least $1 - \alpha$. The question is: does his return have positive expectation at any fixed time ? We will show that it does. But first a justification for our interest in this game.

Each game represents a single step on the path between two states of the coupled Markov chain. We start with X_0 and Y_0 differing at a single vertex. The first game is won if the first time the distance between the coupled chains changes is in convergence. The game is lost if the distance increases to l. At that point we consider the distance l path X_t to Y_t, and the l games played represent the l steps in the path. Although these games are clearly dependent, they each satisfy the conditions given. The gambler's return at time t is one minus the length of the path at time t, so a positive expected return corresponds to an expected path length less than one. We will show that the expected path length is sufficiently small to ensure coupling.

For the initial game we define the *level* to be zero, for any other possible game we define the level to be one greater than the level of the game whose loss precipitated it. We define the random variables M_k, l_{jk} and $I_{jk(t)}$ as follows. M_k is the number of games at level k that are played, l_{jk}, for $j = 1 \dots M_k$, is the number of games in level $k+1$ which are started as a result of the outcome of game j in level k, and $I_{jk}(t)$ is an indicator function which takes the value 1 ifgame j in level k is active at time t, and 0 otherwise. Let $N(t)$ be the number of games active at time t. Then, by linearity of expectations,

$$\mathbf{E}[N(t)] = \sum_{k=0}^{\infty} \mathbf{E}\left[\sum_{j=1}^{M_k} I_{jk}(t)\right]. \tag{1}$$

We will bound this sum in two parts, splitting it at a point $k = K$ to be determined. For $k \leq K$ we observe that $M_k \leq D_2^k$. Since $\Pr(I_{jk}(t) = 1)$ is at most the probability that exactly $k-1$ games of a sequence are complete at time t, regardless of outcome, we have

$$\begin{aligned}\mathbf{E}\left[\sum_{j=1}^{M_k} I_{jk}(t)\right] &\leq D_2^k \max_j \mathbf{E}[I_{jk}(t)] \\ &\leq D_2^k \Pr(\text{exactly } k-1 \text{ games complete by time } t).\end{aligned}$$

So that

$$\begin{aligned}\sum_{k=0}^{K} \mathbf{E}\left[\sum_{j=1}^{M_k} I_{jk}(t)\right] &\leq \sum_{k=0}^{K} D_2^k \Pr(\text{exactly } k-1 \text{ games complete by time } t) \\ &\leq {D_2}^K \Pr(\text{at most } K \text{ games complete by } t). \qquad (2)\end{aligned}$$

On the other hand, for $k > K$ we observe that

$$\begin{aligned}\mathbf{E}\left[\sum_{j=1}^{M_k} I_{jk}(t)\right] \leq \mathbf{E}[M_k] &= \mathbf{E}_{M_{k-1}}\left[\mathbf{E}[M_k | M_{k-1}]\right] \\ &= \mathbf{E}_{M_{k-1}}\left[\mathbf{E}[\sum_{j=1}^{M_{k-1}} l_{jk-1} | M_{k-1}]\right]\end{aligned}$$

Since $\mathbf{E}[l_{jk-1}] \leq \alpha$ for any starting conditions, we may apply this bound even when conditioning on M_{k-1}. So

$$\mathbf{E}\left[\sum_{j=1}^{M_k} I_{jk}(t)\right] \leq \mathbf{E}[\alpha M_{k-1}] \leq \alpha^k, \tag{3}$$

using linearity of expectation, induction and $\mathbf{E}[M_1] \leq \alpha$. Putting (2) and (3) together we get

$$\begin{aligned}\mathbf{E}[N(t)] \leq &\ {D_2}^K \Pr(\text{at most } K \text{ games complete by } t) + \sum_{k=K+1}^{\infty} \alpha^k \\ = &\ {D_2}^K \Pr(\text{at most } K \text{ games complete by } t) + \frac{\alpha^{K+1}}{1-\alpha}. \qquad (4)\end{aligned}$$

We now set $K = \lfloor (\ln\alpha)^{-1}\ln(\frac{\varepsilon(1-\alpha)}{2D_1})\rfloor$, hence the final term is at most $\varepsilon/2D_1$. The probability that a game completes in any given step is at least p. If we select a time $\tau \geq c/p$ for $c \geq K+1 \geq 1$, then the probability that at most K games are complete is clearly maximised by taking this probability to be exactly p in all games. Hence, by Chernoff's bound (see, for example, [14, Theorem 2.1]),

$$\begin{aligned}\mathbf{E}[N(\tau)] \leq\ & {D_2}^K \sum_{k=0}^{K} \binom{\tau}{k} p^k (1-p)^{\tau-k} + \frac{\varepsilon}{2D_1} \\ \leq\ & e^{K\ln D_2 - \frac{(c-K)^2}{2c}} + \frac{\varepsilon}{2D_1} \\ \leq\ & e^{K\ln D_2 + K - c/2} + \frac{\varepsilon}{2D_1}.\end{aligned}$$

Choosing $c = 2K\ln(eD_2) + 2\ln\frac{2D_1}{\varepsilon}$, we obtain $\mathbf{E}[N(\tau)] < \frac{\varepsilon}{D_1}$, where $\tau = \lceil \frac{3\ln(eD_2)}{p(1-\alpha)} \ln\left(\frac{2D_1}{\varepsilon(1-\alpha)}\right)\rceil$.

We conclude that the gambler's expected return at time τ is positive. More importantly, for any initial states $X_0, Y_0 \in \Omega$, the expected distance at time τ is at most ε by linearity of expectations, and so the probability that the chain has not coupled is at most ε. The mixing time claimed now follows by standard arguments. See, for example, [16]. □

Remark 2.2. The assumption that the stopping time occurs when the distance changes is not essential. The assumption that S contains only pairs of states at distance 1 can be removed at the expense of a more complicated proof. We clearly cannot dispense with assumption (ii), or we cannot bound mixing time. Assumption (i) may appear a restriction, but appears to be naturally satisfied in most applications. It seems more natural than the assumption of bounded stopping time, used in [12]. Assumption (i) can easily be replaced by something weaker, for example by allowing p to vary with time rather than remain constant. Provided $p \neq 0$ sufficiently often, a similar proof will be valid.

Remark 2.3. Let $\gamma = 1/(1-\alpha)$. It seems likely that D_2 will be small in comparison to γ in most applications, so we might suppose $D_2 < \gamma < D_1$. The mixing time bound from Theorem 2.1 can then be written $O(p^{-1}\gamma\log D_2 \log(D_1/\varepsilon))$. We may compare this with the bound which can be derived using [12, Corollary 4]. This can be written in similar form as $O(p^{-1}\gamma\log\gamma\log(D_1/\varepsilon))$. In such cases we obtain a reduction in the estimate of mixing time by a factor $\log\gamma/\log D_2$. In the applications below, for example, we have $D_2 = 2$ and $\gamma = \Omega(\Delta)$, so the improvement is $\Omega(\log\Delta)$.

Remark 2.4. The reason for our improvement on the result of [12] is that the use of an upper bound on the stopping time, as is done in [12], will usually underestimate the number of stopping times which occur in a long interval, and hence the mixing rate.

3 Hypergraph Independent Sets

We now use the approach of path coupling via stopping times to prove that the chain discussed in Section 1.1 is rapidly mixing. The metric used in path coupling analyses throughout the paper will be Hamming distance between the coupled chains. We prove the following theorem.

Theorem 3.1. *Let λ, Δ be fixed, and let $\mathcal{H}$ be a hypergraph such that $m \geq 2\lambda\Delta + 1$. Then the Markov chain $\mathcal{M}(\mathcal{H})$ has mixing time $O(n \log n)$.*

Before commencing the proof itself, we analyse the stopping time T for this problem.

3.1 Edge Process

Let X_t and Y_t be copies of $\mathcal{M}$ which we wish to couple, with $Y_0 = X_0 \cup \{w\}$. Let e be any edge containing w, with $m = |e|$. We consider only the times at which some vertex in e is chosen. The progress of the coupling on e can then be modelled by the following "game". We will call the number of unoccupied vertices in e (excluding w) *units*. At a typical step of the game we have k units, and we either win the game, win a unit, keep the same state or lose a unit. These events happen with the following probabilities: we win the game with probability $1/m$, win a unit with probability at least $(m-k-1)/(1+\lambda)m$, lose a unit with probability at most $\lambda k/(1+\lambda)m$ and stay in the same state otherwise. If ever $k = 0$, we are bankrupt and we lose the game. Winning the game models the "good event" that the vertex v is chosen and the two chains couple. Losing the game models the "bad event" that the coupling increases the distance to 2. We wish to know the probability that the game ends in bankruptcy. We are most interested in the case where $k = 1$ initially, which models e being critical. Note that the value of k in the process on hypergraph independent sets dominates the value in our model, since we can always delete (win in the game), but we may not be able to insert (lose in the game) because the chosen vertex is critical in some other edge.

Let p_k denote the probability that a game is lost, given that we start with k units. We have the following system of simultaneous equations.

$$\begin{aligned} (m-1+2\lambda)p_1 - (m-2)p_2 &= \lambda \\ -k\lambda p_{k-1} + (m-k+(k+1)\lambda)p_k - (m-k-1)p_{k+1} &= 0 \quad (k=2,3,\ldots,m-1) \end{aligned} \tag{5}$$

Solving these yields the following, which may be confirmed by substituting into Equations (5).

$$p_k = \frac{1}{\binom{m-1}{k}}\left(\lambda^k - \frac{\sum_{i=1}^{k}\binom{m}{i}\lambda^{m+k-i}}{(1+\lambda)^m - \lambda^m}\right) = \frac{\sum_{i=k+1}^{m}\binom{m}{i}\lambda^{m+k-i}}{\left((1+\lambda)^m - \lambda^m\right)\binom{m-1}{k}} \quad (k=1,2,\ldots,m-1). \tag{6}$$

In particular

$$p_1 = \frac{\lambda}{m-1}\left(1 - \frac{m\lambda^{m-1}}{(1+\lambda)^m - \lambda^m}\right). \tag{7}$$

3.2 The Expected Distance Between X_T and Y_T

The stopping time for the pair of chains X_t and Y_t will be when the distance between them changes, in other words either a good or bad event occurs. The probability that we observe the bad event on a particular edge e with $w \in e$ is at most p_k as calculated above. Let ξ_t denote the number of empty vertices in e at time t when the process is started with $\xi_0 = k$. Now ξ_t can never reach 0 without first reaching $k-1$ and, since the process is Markovian, it follows that

$$p_k = \Pr(\exists t\, \xi_t = 0 | \xi_0 = k) = \Pr(\exists t\, \xi_t = 0 | \xi_s = k-1) \Pr(\exists s\, \xi_s = k-1 | \xi_0 = k) < p_{k-1}.$$

Since w is in at most Δ edges, the probability that we observe the bad event on any edge is at most Δp_1. The probability that the stopping time ends with the good event is therefore at least $1 - \Delta p_1$. The path coupling calculation is then

$$\mathbf{E}[\mathrm{d}(X_T, Y_T)] \leq 2\Delta p_1.$$

This is required to be less than 1 in ordered to apply Theorem 2.1. If $m \geq 2\lambda\Delta + 1$, then by (7)

$$2\Delta p_1 = 1 - \frac{(2\lambda\Delta+1)\lambda^{2\lambda\Delta}}{(1+\lambda)^{2\lambda\Delta+1} - \lambda^{2\lambda\Delta+1}}.$$

Proof of Theorem 3.1. The above work puts us in a position to apply Theorem 2.1. Let $m \geq 2\lambda\Delta + 1$. Then for $\mathcal{M}(\mathcal{H})$ we have

(i) $\Pr(\mathrm{d}(X_t, Y_t) \neq 1 | \mathrm{d}(X_{t-1}, Y_{t-1}) = 1) \geq \frac{1}{n}$ for all t, and

(ii) $\mathbf{E}[\mathrm{d}(X_T, Y_T)] < 1 - \dfrac{(2\lambda\Delta+1)\lambda^{2\lambda\Delta}}{(1+\lambda)^{2\lambda\Delta+1} - \lambda^{2\lambda\Delta+1}}$.

Also for $\mathcal{M}(\mathcal{H})$ we have $D_1 = n$ and $D_2 = 2$. Hence by Theorem 2.1, $\mathcal{M}(\mathcal{H})$ mixes in time

$$\tau(\varepsilon) \leq 6n \frac{(1+\lambda)^{2\lambda\Delta+1} - \lambda^{2\lambda\Delta+1}}{(2\lambda\Delta+1)\lambda^{2\lambda\Delta}} \ln\left(n\varepsilon^{-1} \frac{(1+\lambda)^{2\lambda\Delta+1} - \lambda^{2\lambda\Delta+1}}{(2\lambda\Delta+1)\lambda^{2\lambda\Delta}}\right).$$

This is $O(n \log n)$ for fixed λ, Δ. □

Remark 3.2. In the most important case, $\lambda = 1$, we require $m \geq 2\Delta + 1$. This does not include the case $m = 3$, $\Delta = 3$ considered in [10]. We have attempted to improve the bound by employing the chain proposed by Dyer and Greenhill in [10, Section 4]. However, this gives only a marginal improvement. For large $\lambda\Delta$, we obtain convergence for $m \geq 2\lambda\Delta + \frac{1}{2} + o(1)$. For $\lambda = 1$, this gives a better bound on mixing time for $m = 2\Delta + 1$, with dependence on Δ similar to Remark 3.3 below, but does not even achieve mixing for $m = 2\Delta$. We omit the details in order to deal with the Glauber dynamics, and to simplify the analysis.

Remark 3.3. The terms in the running time which are exponential in λ, Δ would disappear if we instead took graphs for which $m \geq 2\lambda\Delta + 2$. In this case the running time would be

$$\tau(\varepsilon) \leq 6(2\lambda\Delta + 1)n \ln(n\varepsilon^{-1}(2\lambda\Delta + 1)) \leq 12(2\lambda\Delta + 1)n \ln(n\varepsilon^{-1}).$$

Furthermore, if we took graphs such that $m > (2 + \delta)\lambda\Delta$, for some $\delta > 0$, then the running time would no longer depend on λ, Δ at all, but would be $\tau(\varepsilon) \leq c_\delta n \ln(n\varepsilon^{-1})$ for some constant c_δ.

Remark 3.4. It seems that path coupling cannot show anything better than m linear in $\lambda\Delta$. Suppose the initial configuration has edges $\{w, v_1, \ldots, v_{m-2}, x_i\}$ for $i = 1, \ldots, \Delta$, with $w, v_1, \ldots, v_{m-2} \in X_0$, $x_1, \ldots, x_\Delta \notin X_0$ and w the change vertex. Consider the first step where any vertex changes state. Let $\mu = (1 + \lambda)(m - 1 + \Delta)$. The good event occurs with probability $(1 + \lambda)/\mu$, insertion of a critical vertex with probability $\lambda\Delta/\mu$, and deletion of a non-critical vertex with probability $(m - 1)/\mu$. We therefore need $(m - 1) + (1 + \lambda) \geq \lambda\Delta$, i.e. $m \geq \lambda(\Delta - 1)$, to show convergence by path coupling.

Remark 3.5. It seems we could improve our bound $m \geq 2\lambda\Delta + 1$ for rapid mixing of the Glauber dynamics somewhat if we could analyse the process on all edges simultaneously. Examination of the extreme cases, where all edges adjacent to w are otherwise independent, or where they are dependent except for one vertex (as in Remark 3.4), suggests that improvement to $(1 + o(1))\lambda\Delta$ may be possible, where the $o(1)$ is relative to $\lambda\Delta$. However, the analysis in the general case seems difficult, since edges can intersect arbitrarily.

4 Hardness Results for Independent Sets

We have established that the number of independent sets of a hypergraph can be approximated efficiently using the Markov Chain Monte Carlo technique for hypergraphs with edge size linear in Δ. We next state hardness results, in particular that exact counting is unlikely to be possible, and that unless NP=RP, there can be no *fpras* for the number of independent sets of all hypergraphs with edge size $\Omega(\log \Delta)$.

Theorem 4.1. *Let $\mathcal{G}(m, \Delta)$ be the class of uniform hypergraphs with minimum edge size $m \geq 3$ and maximum degree Δ. Computing the number of independent sets of hypergraphs in $\mathcal{G}(m, \Delta)$ is #P-complete if $\Delta \geq 2$. If $\Delta \leq 1$, it is in P.*

Let $G = (V, E)$, with $|V| = n$, be a graph with maximum degree Δ and N_i independent sets of size i $(i = 0, 2, \ldots, n)$. For $\lambda > 0$ let $Z_G(\lambda) = \sum_{i=0}^{n} N_i \lambda^i$ define the *hard core partition function*. The following is a combination of results in Luby and Vigoda [18] and Berman and Karpinski [2].

Theorem 4.2. *If $\lambda > 694/\Delta$, there is no* fpras *for $Z_G(\lambda)$ unless NP=RP.*

We note that Theorem 4.2 could probably be strengthened using the approach of [8]. However, this has yet to be done.

Theorem 4.3. *Unless NP=RP, there is no* fpras *for counting independent sets in hypergraphs with maximum degree Δ and minimum edge size $m < 2\lg(1 + \Delta/694) - 1 = \Omega(\log \Delta)$.*

5 Hypergraph Colouring

In this section we present without proof another result obtained using Theorem 2.1. We now consider Glauber dynamics on the set of proper colourings of a hypergraph. Again our hypergraph $\mathcal{H}$ will have maximum degree Δ, minimum edge size m, and we will have a set of q colours. 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

(i) selecting a vertex $v \in \mathcal{V}$ and a colour $k \in \{1, 2, \ldots, q\}$ uniformly at random,
(ii) let X'_t be the colouring obtained by recolouring v colour k
(iii) 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. Again we may use Theorem 2.1 to prove rapid mixing of this chain under certain conditions, however first we note that the following result may be obtained using standard path coupling techniques.

Theorem 5.1. *For $m \geq 4$, $q > \Delta$, the Markov chain $\mathcal{C}(\mathcal{H})$ mixes in time $O(n \log n)$.*

This leaves little room for improvement in the case $m \geq 4$, indeed it is not clear whether the Markov chain described is even ergodic for $q \leq \Delta$. The following simple construction does show that the chain is not in general ergodic if $q \leq \frac{\Delta}{m} + 1$. Let $q = \frac{\Delta}{m} + 1$, and take a hypergraph $\mathcal{H}$ on $q(m-1)$ vertices. We will group the vertices into q groups $\mathcal{V} = \mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_q$, each of size $m-1$. Then the edge set of $\mathcal{H}$ is $E = \{\{v\} \cup \mathcal{V}_j : v \in \mathcal{V}, v \notin \mathcal{V}_j\}$. The degree of each vertex is $(q-1) + (q-1)(m-1) = \Delta$. If we now colour each group $\mathcal{V}_j$ a different colour, we obtain $q!$ distinct colourings, but for each of these the Markov chain is frozen (no transition is valid).

The case $m = 2$ is graph colouring and has been extensively studied. See, for example, [7]. This leaves the case $m = 3$, hypergraphs with 3 vertices in each edge. The standard path coupling argument only shows rapid mixing for $q \geq 2\Delta$, since there may be two vertices in each edge that can be selected and lead to a divergence of the two chains. This occurs if, of the two vertices in an edge which are not w, one is coloured red and the other blue. However, we can do better using Theorem 2.1. The proof is omitted, but again hinges on the fact that an edge is very unlikely to persist in such a critical state.

Theorem 5.2. *There exists Δ_0 such that, if $\mathcal{H}$ is a 3-uniform hypergraph with maximum degree $\Delta > \Delta_0$ and $q \geq 1.65\Delta$, the Markov chain $\mathcal{C}(\mathcal{H})$ mixes rapidly.*

6 Hardness Results for Colouring

Again we state a theorem that exact counting is #P-complete except in the few cases where it is clearly in P. Let $\mathcal{G}(m, \Delta)$ be as in Theorem 4.1.

Theorem 6.1. *Computing the number of q-colourings of hypergraphs in $\mathcal{G}(m, \Delta)$ is #P-complete if $\Delta, q > 1$. If $\Delta \leq 1$ or $q \leq 1$ it is in P.*

Again let $\mathcal{G}(m, \Delta)$ be as defined in Theorem 4.1. The hardness of approximation result, Corollary 6.3, follows directly from the following NP-completeness result.

Theorem 6.2. *Determining whether a hypergraph in $\mathcal{G}(m, \Delta)$ has any q-colouring is NP-complete for any $m > 1$ and $2 < q \leq (1 - 1/m)\Delta^{1/(m-1)}$.*

Corollary 6.3. *Unless NP=RP, there is no* fpras *for counting q-colourings of a hypergraphs with maximum degree Δ and minimum edge size m if $2 < q \leq (1 - 1/m)\Delta^{1/(m-1)}$.*

Remark 6.4. It is clearly a weakness that our lower bound for approximate counting is based entirely on an NP-completeness result. However, we note that the same situation pertains for graph colouring, which has been the subject of more intensive study.

7 Conclusions

We have presented an approach to the analysis of path coupling with stopping times which improves on the method of [12] in most applications. Our method may itself permit further development.

We apply the method to independent sets and q-colourings in hypergraphs with maximum degree Δ and minimum edge size m. In the case of independent sets, there seems scope for improving the bound $m \geq 2\Delta+1$, but anything better than $m \geq \Delta + o(\Delta)$ would seem to require new methods. For colourings, there is probably little improvement possible in our result $q > \Delta$ for $m \geq 4$, but many questions remain for $q \leq \Delta$. For example, even the ergodicity of the Glauber (or any other) dynamics is not clearly established. For the most interesting case, $m = 3$, the bound $q > 1.65\Delta$ (for large Δ) can almost certainly be reduced, but substantial improvement may prove difficult.

Our #P-completeness results seem best possible for both of the problems we consider. On the other hand, our lower bounds for hardness of approximate counting seem very weak in both cases, and are far from our upper bounds. These lower bounds can probably be improved, but we have no plausible conjecture as to what may be the truth.

Acknowledgments

We are grateful to Tom Hayes for commenting on an earlier draft of this paper, and to Mary Cryan for useful discussions at an early stage of this work.

References

1. M. Bordewich, M. Dyer and M. Karpinski, Path coupling using stopping times and counting independent sets and colourings in hypergraphs, (2005) http://arxiv.org/abs/math.PR/0501081.
2. P. Berman and M. Karpinski, Improved approximation lower bounds on small occurrence optimization, *Electronic Colloquium on Computational Complexity* **10** (2003), Technical Report TR03-008.
3. R. Bubley, *Randomized algorithms: approximation, generation and counting*, Springer-Verlag, London, 2001.
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. I. Dinur, V. Guruswami, S. Khot and O. Regev, A new multilayered PCP and the hardness of hypergraph vertex cover, in *Proc. 35th ACM Symposium on Theory of Computing (STOC 2003)*, ACM, 2003, pp. 595–601.
6. I. Dinur, O. Regev and C. Smyth, The hardness of 3-uniform hypergraph coloring, in *Proc. 43rd Symposium on Foundations of Computer Science (FOCS 2002)*, IEEE, 2002, pp. 33–42
7. 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 (FOCS 2004)*, IEEE, 2004, pp. 582–589.
8. M. Dyer, A. Frieze and M. Jerrum, On counting independent sets in sparse graphs, *SIAM Journal on Computing* **31** (2002), 1527–1541.
9. 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.
10. M. Dyer and C. Greenhill, On Markov chains for independent sets, *Journal of Algorithms* **35** (2000), 17–49.
11. M. Garey and D. Johnson, *Computer and intractability*, W. H. Freeman and Company, 1979.
12. T. Hayes and E. Vigoda, Variable length path coupling, in *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2004)*, SIAM, 2004, pp. 103–110.
13. T. Hofmeister and H. Lefmann, Approximating maximum independent sets in uniform hypergraphs, *Proc. 23rd International Symposium on Mathematical Foundations of Computer Science (MFCS 1998)*, Lecture Notes in Computer Science **1450**, Springer, 1998, pp. 562–570.
14. S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.
15. 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.
16. M. Jerrum, *Counting, sampling and integrating: algorithms and complexity*, ETH Zürich Lectures in Mathematics, Birkhäuser, Basel, 2003.
17. M. Krivelevich, R. Nathaniel and B. Sudakov, Approximating coloring and maximum independent sets in 3-uniform hypergraphs, in *Proc. 12th Annual ACM-SIAM Symposium on Discrete Algorithms, (SODA 2001)*, SIAM, 2001, pp. 327–328.
18. M. Luby and E. Vigoda, Fast convergence of the Glauber dynamics for sampling independent sets, *Random Structures and Algorithms* **15** (1999), 229–241.
19. M. Mitzenmacher and E. Niklova, Path coupling as a branching process, unpublished manuscript, 2002.

20. M. Molloy, Very rapidly mixing Markov chains for 2Δ-coloring and for independent sets in a graph with maximum degree 4, *Random Structures and Algorithms* **18** (2001), 101–115.
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, A note on the Glauber dynamics for sampling independent sets, *The Electronic Journal of Combinatorics* **8**, R8(1), 2001.