# MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold

Charilaos Efthymiou*
Goethe University, Mathematics Institute, Frankfurt 60054, Germarny
efthymiou@math.uni-frankfurt.de

**Abstract**

Sampling from the *Gibbs distribution* is a well studied problem in computer science as well as in statistical physics. In this work we focus on the *k-colouring model* and the *hard-core model* with fugacity $\lambda$ when the underlying graph is an instance of Erdős-Rényi random graph $G(n, p)$, where $p = d/n$ and $d$ is fixed.

We use the *Markov Chain Monte Carlo* method for sampling from the aforementioned distributions. In particular, we consider *Glauber (block) dynamics*. We show a dramatic improvement on the bounds for *rapid mixing* in terms of the number of colours and the fugacity for the corresponding models. For both models the bounds we get are only within small constant factors from the conjectured ones by the statistical physicists.

We use *Path Coupling* to show rapid mixing. For $k$ and $\lambda$ in the range of our interest the technical challenge is to cope with the high degree vertices, i.e. vertices of degree much larger than the expected degree $d$. The usual approach to this problem is to consider block updates rather than single vertex updates for the Markov chain. Taking appropriately defined blocks the effect of high degree vertices diminishes. However devising such a block construction is a *non trivial* task.

We develop for a first time a *weighting schema* for the paths of the underlying graph. Only, vertices which belong to "light" paths can be placed at the boundaries of the blocks. The tree-like local structure of $G(n, d/n)$ allows the construction of simple structured blocks.

## 1 Introduction

Sampling from the *Gibbs distribution* is well studied problem in computer science as well as in statistical physics. Examples include sampling from the uniform (or a weighted) distribution over combinatorial structures like $k$-colourings, independent sets, matchings of a graph $G$ e.t.c. In this work we focus on colourings and independent sets when the underlying graph is an instance of Erdős-Rényi random graph $G(n, p)$, where $p = d/n$ and $d$ is 'large' but remains bounded as $n \to \infty$. We say that an event occurs *with high probability (w.h.p.)* if the probability of the event to occur tends to 1 as $n \to \infty$.

For this kind of problems, the most powerful algorithms and somehow the most natural ones are based on the Markov Chain Monte Carlo (MCMC) method. The setup is an ergodic, time-reversible Markov chain over the $k$-colourings (or independent sets) of the underlying graph. The updates guarantee that the equilibrium distribution of the chain is the desired one. Here we use standard *Glauber block updates* (in the course of this paper we refer to the chains as *Glauber dynamics*). The main technical challenge is to establish that the underlying Markov chain has *rapid mixing*, i.e. it converges sufficiently fast to the equilibrium distribution (see [10, 17, 16]).

Given the input graph $G(n, d/n)$, the focus is on two distributions. The first one is the *colouring model*, i.e. the uniform distribution over the $k$-colourings of the input graph. The second one is the *hard-core model* with *fugacity* $\lambda$, i.e. each independent set $\sigma$ is assigned probability measure proportional to $\lambda^{|\sigma|}$. The parameters of interest are the number of colours $k$ and the fugacity $\lambda$, respectively. The aim is to show rapid mixing for $k$ as small as possible and $\lambda$ as large as possible for the corresponding models.

For MCMC algorithms to converge, typically, the bounds for both $k$ and $\lambda$ are expressed in terms of the *maximum degree* of the underlying graph. Examples of such bounds are [5, 11, 12, 14, 21, 22, 27] for colouring and [6, 7, 21, 28] for independent sets. In that terms, what makes the case of $G(n, d/n)$ special is the (relatively) big fluctuation in the degree of the vertices. To be more specific, w.h.p. the vast majority of vertices in $G(n, d/n)$ are of degree close to $d$, while the maximum degree is as huge as $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$. In such a situation, it is natural to expect that the rapid mixing bounds for both $k$, $\lambda$ depend on the *expected degree* $d$,

*Partially supported by EPSRC grant EP/G039070/2.

rather than the maximum degree.

Sophisticated but mathematically non rigorous arguments from statistical physics (e.g. in [19, 30]) support this picture. Roughly speaking, they suggest that w.h.p. over the instances of $G(n, d/n)$ the Glauber (block) dynamics on $k$-colouring has rapid mixing for any $k > d$. Furthermore, for $k < d$ the chain is expected to be *non-ergodic* [1]. To our knowledge, there are no predictions for the fugacity as far as the hard-core model is concerned. However, using the result in [18] and standard arguments we could conjecture that we have rapid mixing as long as $\lambda < \frac{(d-1)^{d-1}}{(d-2)^d} \approx \frac{e}{d}$. These conjectured threshold values for $k$ and $\lambda$ are what we call *Gibbs Uniqueness thresholds*.

So far, the best bounds for Monte Carlo sampling appeared in [24] (which improved on [4]). The authors in [24] provide for a first time rapid mixing bounds, for both $k$ and $\lambda$, which depend on the expected degree $d$. That is, w.h.p. over $G(n, d/n)$ there are functions $f(d)$ and $h(d)$ such that Glauber dynamics has rapid mixing for $k$-colourings and hard-core as long as $k \geq f(d)$ and $\lambda \leq h(d)$, respectively[2]. However, the values for $k$ and $\lambda$ that are allowed there are many orders of magnitude off the conjectured bounds. Here we improve on these bounds substantially. We show that w.h.p. over the underlying graph $G(n, d/n)$ we have rapid mixing for $k \geq \frac{11}{2}d$ and for $\lambda \leq \frac{1-\epsilon}{2}\frac{1}{d}$. That is, we approach the conjectured bounds for rapid mixing only within small constants.

We use the well-known *Path Coupling* technique, from [3], to show rapid mixing. Path Coupling is also used in both of the previous papers on the problem, i.e. [4, 24]. For $k$ and $\lambda$ in the range of our interest the technical challenge is to cope with the high degree vertices, i.e. vertices of degree much larger than $d$. The natural approach is to consider block updates rather than single vertex updates for the Markov chain. In particular we use the observation that the effect of high degree vertices somehow diminishes when they are away from the boundary of their block. Devising such a block construction is a highly complex task. We introduce for a first time a weighting schema for the paths of the underlying graph which allows a desired block construction.

To be more specific, we use our weighting schema to assign weight to paths in $G(n, d/n)$. These weights allow distinguishing which vertices can be used for the boundaries of the blocks. We call such vertices *break-points*. A break point should have all the paths emanating from it of sufficiently small weight. It turns out that w.h.p. there is a plethora of break-points in $G(n, d/n)$. This allows creating small, simple structured blocks.

Compared to [24], one could remark that our weighting schema allows a more specific characterization of the blocks. I.e. we have more information about the position of high degree vertices inside the blocks. Somehow, this allows better results from the path coupling analysis. Also, we should mention that our setting for path coupling analysis is closer to [4] rather than [24].

**Non Monte Carlo Approaches.** In the literature, there are approximate sampling algorithms for both colouring and hard-core model on $G(n, d/n)$ which do not use the MCMC approach. Usually their efficiency requirements are weaker than those for the MCMC ones. In practice this means that we get efficiency guarantees for a wider range of the parameters $k$ and $\lambda$. However, the approximation guarantees we get are *weaker*.

In [8] the author of this work proposed an algorithm for approximate sampling $k$-colourings of $G(n, d/n)$ which has a notable performance in terms of minimum $k$, it requires $k \geq (2 + \epsilon)d$. The error of the output is a vanishing function of the size of the graph $n$. Recently Sinclair, Srivastava and Yin in [25] presented a non-MCMC algorithm for sampling from the hard core-model on $G(n, d/n)$ for $\lambda < e/d$. Essentially, [25] improves (non trivially) on [29] for the case where the underlying graph has bounded connectivity[3]. For typical instances of $G(n, d/n)$, the algorithm requires $poly(n) \times poly(m)$ steps to return a sample within distance $1/m$ from the target distribution.

Given rapid mixing, the output of a MCMC algorithm is within distance $1/m$ from the target distribution in a $\ln(m) \times poly(n)$ number of steps.

**Notation.** We use small letters of the greek alphabet to indicate colourings or independent sets, e.g. $\sigma, \tau$. Also, by $\sigma(v)$ we indicate the assignment of the vertex $v$ under the configuration $\sigma$. For a vertex set $B$ we call *(outer) boundary* of $B$ the vertices outside $B$ which are adjacent to some vertex inside $B$.

## 2 The Result

We consider the *colouring model* and the *hard core model*. For each of these two models we consider a graph $G = (V, E)$ and a set of *spins* $C$. We define a *configuration space* $\Omega \subseteq C^V$. Given $\Omega$, the model

[1] When $k < d$ there is a connected block which contains all but an exponentially small fraction the $k$-colourings of $G(n, d/n)$. It is conjectured that restricting the states of the chain to the $k$-colourings in this block we should have rapid mixing.

[2] Even though these functions are not given explicitly it is conceivable from the analysis that it holds that $f(d) \geq d^c$ and $h(d) \leq d^{-c'}$ for fixed $c, c' > 2$.

[3] It is not hard to see that typical instances of $G(n, d/n)$ are such graphs.

specifies a distribution $\mu : \Omega \to [0, 1]$. This distribution is also called *Gibbs distribution*.

`Colouring Model:` Given a graph $G = (V, E)$ and a sufficiently large integer $k$, the *colouring model* specifies the following: The configuration space $\Omega$ consist of all the proper $k$-colourings of $G$. The Gibbs distribution is the uniform distribution over $\Omega$. That is, each $\sigma \in \Omega$ is assigned probability measure

$$\mu(\sigma) = \frac{1}{|\Omega|}.$$

`Hard Core Model:` Given $G = (V, E)$ and $\lambda > 0$, the hard core model with *fugacity* $\lambda$ specifies the following: The configuration space $\Omega$ is all the independent sets of $G$. The Gibbs distribution specifies that each $\sigma \in \Omega$, is assigned probability measure $\mu(\sigma)$ which is proportional to $\lambda^{|\sigma|}$, where $|\sigma|$ is the cardinality of $\sigma$. That is

$$\mu(\sigma) = \frac{1}{Z}\lambda^{|\sigma|},$$

$Z$ is a normalizing quantity, i.e. $Z = \sum_{\sigma\in\Omega} \lambda^{|\sigma|}$, which is usually called *partition function*.

In this work we propose a Markov Chain Monte Carlo algorithm for approximate sampling from the two models above. The error in the distribution of the output sample is expressed in terms of *total variation distance*.

For two distribution $\nu, \xi$ on some discrete space $S$ we define the total variation distance as follows:

$$||\nu - \xi||_{TV} = \max_{A\subseteq S} |\nu(A) - \xi(A)|.$$

A high level description of the algorithm is as follows:

`Sampling Algorithm:` The input is a graph $G = (V, E)$, a number $err > 0$ which is the error in the distribution of the output sample and the target Gibbs distribution[4] $\mu$.

First, the algorithm partitions the set of vertices $V$ into an appropriate set of blocks $\mathcal{B}$. Given $\mathcal{B}$, it simulates the following Markov chain and returns the configuration of the chain after $T = T(err)$ transitions.

- Start from an *arbitrary* configuration.
- At each transition, the chain chooses uniformly at random a block $B \in \mathcal{B}$. If $X_t$ is the current state of the chain, then the next one, $X_{t+1}$, is acquired as follows:
  - For every vertex $u \notin B$ set $X_{t+1}(u) = X_t(u)$.
  - Set $X_{t+1}(B)$ according to Gibbs distribution conditional that the spins outside $B$ are set $X_{t+1}(V\backslash B)$.

Given a set of technical conditions known as *ergodicity*, the above chain converges to Gibbs distribution $\mu$. For the range of $k$ and $\lambda$ we consider here ergodicity is well known to hold for our chains (see more details in full version [9]).

The performance of our algorithm both in terms of time efficiency and accuracy depends on the following conditions:

1. The construction of the set of blocks $\mathcal{B}$ should be done in polynomial time.
2. We need an efficient algorithm that provides the initial configuration of the chain.
3. We need an efficient algorithm which implements the transitions of the chain.
4. The chain should converge to stationarity sufficiently fast (we have rapid mixing).

For further details on how we deal with each of the above issues see in Section 3 and Section 4. The convergence rate, which we consider to be the most important of the four, is treated separately in Section 3. The rest are treated in Section 4. We are going to show that our algorithm indeed satisfies the above conditions with high probability over the instances of the input graph. In particular, the main result of this work is stated in the following theorem.

THEOREM 2.1. *We let $\epsilon > 0$, sufficiently small $err > 0$, and sufficiently large $c, d > 0$. W.h.p. over the input instances $G(n, d/n)$ the* `Sampling Algorithm` *returns a configuration within total variation distance $err$ from the target distribution $\mu$ in $\ln\left(\frac{1}{err}\right) \times O(n^c)$ number of steps, as long as the following conditions are met:*

**colouring model:** *It holds that $k \geq \frac{11}{2}d$*

**hard-core model:** *It holds that $\lambda \leq \frac{1-\epsilon}{2d}$.*

Theorem 2.1 follows as a corollary of the results in Sections 3 and 4.

## 3 Rapid Mixing Results & Proof Technique

We use *mixing time*, $\tau_{mix}$, as a measure of the speed of convergence of Markov chains. The mixing time is defined as the number of transitions needed in order to guarantee that the chain starting from an arbitrary configuration, is within total variation distance $1/e$ from the stationary distribution (see [20]).

[4] colouring or hard-core model with the appropriate parameters

REMARK 1. *In our context, we say that a Markov chain is* rapidly mixing *if $\tau_{mix}$ is polynomial in $n$.*

REMARK 2. *It is not hard to see that the number of transitions that are required to get within error err from the stationary distribution is $T(err) = \ln\left(\frac{1}{err}\right) \times \tau_{mix}$.*

As far as the mixing rate of the two chains is regarded we show the following theorems.

THEOREM 3.1. *Let $\mathcal{M}_c$ be the Markov chain on the k-colourings of $G(n, d/n)$ we define in Section 2. With probability $1 - o(1)$ over the graph instances $G(n, d/n)$ and for $k \geq \frac{11}{2}d$, the mixing time of $\mathcal{M}_c$ is $O(n \ln n)$.*

A large part of the proof of Theorem 3.1 appears in Section 6. For the full proof see in the full version of this work in [9].

THEOREM 3.2. *Let $\mathcal{M}_{hc}$ be the Markov chain of the hard core model on $G(n, d/n)$ with fugacity $\lambda$, we define in Section 2. Let $\epsilon > 0$ be fixed. With probability $1-o(1)$ over the graph instances $G(n, d/n)$ and for $\lambda \leq \frac{1-\epsilon}{2d}$ the mixing time of $\mathcal{M}_{hc}$ is $O(n \ln n)$.*

For the proof of Theorem 3.2, see in the full version of this work in [9].

In the two theorems rapid mixing holds w.h.p. over the instances of $G(n, d/n)$ because (among other restrictions) the underlying graph should admit the appropriate partition $\mathcal{B}$. In Section 3.1 we present a *high level* description of the proof technique we use for the two theorems above.

**3.1 Proof technique** We show rapid mixing by using the well-known *Path Coupling* technique in [3]. The technique goes as follows: W.l.o.g. we consider the colouring model. Assume that the underlying graph $G = (V, E)$ is of maximum degree $\Delta$ and, for the moment, let $k > \Delta$. Finally, assume that we have the Markov chain $M$ on the $k$-colourings of $G$ with single vertex updates.

Consider any two copies of $M$ at state $X_0, Y_0$, respectively. We take $X_0, Y_0$ so that they have exactly one disagreement, i.e. their Hamming distance $H(X_0, Y_0)$ is equal to 1. The coupling carries out one transition of each copy of $M$. Let $X_1, Y_1$ be the colouring after each transition, respectively. A sufficient condition for rapid mixing is the following one

$$E[H(X_1, Y_1)|X_0, Y_0] \leq 1 - \Theta\left(n^{-1}\right). \quad (3.1)$$

To study the technique further, assume now that for $w \in V$ we have $X_0(w) \neq Y_0(w)$. It is natural to use a coupling that updates the same vertex in both copies. The cases that matter are only those where the coupling chooses to update either the disagreeing vertex $w$ or one of its neighbours. If the update involves the vertex $w$, then we get that $X_1 = Y_1$. This happens with probability $1/n$, where $|V| = n$. On the other hand, if the update involves a neighbour of $w$, then $X_1, Y_1$ may have an extra disagreement. In particular, the update of a neighbour of $w$ can generate an extra disagreement with probability at most $\frac{1}{k-\Delta}$. Since the disagreeing vertex $w$ has at most $\Delta$ neighbours, the probability of having an extra disagreement is at most $\frac{\Delta}{n}\frac{1}{k-\Delta}$. For $k \geq 2\Delta + 1$, it is direct that (3.1) is satisfied.

W.h.p. $G(n, d/n)$ is of maximum degree $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$. That is, a vanilla path coupling would require an unbounded number of colours. Otherwise, i.e. if $k$ is smaller than the maximum degree, there is no control on the expected number of disagreements generated in the coupling. However, it is possible to gain some control on the disagreements by using (appropriate) block updates rather than single vertex updates. In particular, the blocks should be constructed in such a manner that the high degree vertices are somehow "hidden" well inside them[5].

In our setting, we consider two copies of $\mathcal{M}_c$ at states $X_0, Y_0$. The states differ only on the assignment of the vertex $w$. We have block updates. The coupling chooses uniformly at random a block $B$ from the set of blocks $\mathcal{B}$ and updates the colouring of $B$ in both chains. It turns out that the crucial case for proving (3.1) is when the outer boundary of $B$ is not the same for both chains, i.e. the disagreeing vertex $w$ is not inside $B$ but it is adjacent to some vertices in $B$. There, we need to upper bound the expected number of disagreements generated in the update of $B$. The construction of $\mathcal{B}$ should minimize the expected number of disagreements.

To this end, we build on an idea from [4] for bounding the expected number of disagreements inside the block. The authors there use the well-known "*disagreement percolation*" coupling construction, [2], to express the expected number of disagreements in terms of percolation probabilities. The idea assumes that the block $B$ is a tree with at most one extra vertex (which is the case here as well) and goes as follows: Clearly, the disagreement at the boundary prohibits identical coupling of $X_1(B)$ and $Y_1(B)$. The disagreement percolation assembles the coupling in a stepwise fashion, i.e. couples the colouring of each vertex of $B$ at a time. It gives priority to vertices which are next to a disageeing vertex, if any. Disagreements propagates into $B$ along paths from $w$. A disagreement at vertex $u' \in B$ at (edge) distance $r$ from $w$ propagates

[5] This is the approach that is used in the analysis in both [4, 24].

to a neighbour $u \in B$ at distance $r+1$ if $X_1(u) \neq Y_1(u)$. The probability of the event $X_1(u) \neq Y_1(u)$ in the coupling is upper bounded by $\varrho_u$, where

$$\varrho_u = \begin{cases} \frac{2}{k-(1+\alpha)d} & \text{if } \Delta(u) \leq (1+\alpha)d \\ 1 & \text{otherwise,} \end{cases} \tag{3.2}$$

with $\alpha > 0$ and $k \geq (1+\alpha)d + 2$.

The disagreement percolation is dominated by an independent process, i.e. the disagreement propagates over the path $L$ that starts from $w$ with probability at most $\prod_{u \in L \setminus \{w\}} \varrho_u$. Clearly, the expected number of disagreements is at most the expected number of paths of disagreements that start from $w$ and propagate inside $B$.

Intuitively, high degree vertices are expected to have an increased contribution to the number of disagreements. I.e. if a high degree vertex is disagreeing, it has an increased number of neighbours to propagate the disagreement. However, for typical instances of $G(n, d/n)$ and $k \geq \frac{11}{2}d$, it turns out that the larger the distance between a high degree vertex from $w$ the less probable is for the disagreement to reach it. This, somehow, can balance the increased contribution that high degree vertices have. We exploit this observation in the block construction so as to control the overall number of disagreeing vertices in the updates.

To be more specific, we use the following weighting schema: Each vertex $u$, of degree $\Delta(u)$ in $G(n, d/n)$, is assigned weight $W(u)$ such that

$$W(u) = \begin{cases} (1+\gamma)^{-1} & \text{if } \Delta(u) \leq (1+\alpha)d \\ d^c \cdot \Delta(u) & \text{otherwise,} \end{cases} \tag{3.3}$$

for appropriate real numbers $\alpha, \gamma, c > 0$. Given the weights of the vertices, each block $B \in \mathcal{B}$ should satisfy (among others) the following two properties:

(a) $B$ is either a tree or a unicyclic graph

(b) For every path $L$ between a vertex at the outer boundary of $B$ and a high degree vertex[6] inside $B$ it should hold that $\prod_{u \in L} W(u) \leq 1$.

In the weighting schema, observe that the low degree vertices reduce the weight of the path $L$, while the high degree vertices increase it. Restricting the weight of a path between a high degree vertex in the block $B$ and a boundary vertex, somehow, guarantees that the high degree vertices are sufficiently far from the boundary. I.e. so as to keep the weight of the path low we require a sufficiently large number of low degree vertices between the boundary vertex and the high degree vertex. As we describe in the following paragraphs, choosing appropriately the parameters $\alpha, \gamma, c$ in the weighting schema we can control the expected number of disagreement and get (3.1), for $k$ down to $\frac{11}{2}d$.

**Where $\mathbf{k \geq \frac{11}{2}d}$ comes from.** So as to derive the bound $\frac{11}{2}d$ for $k$ we need to show how do we bound the number of disagreements in the update of the configuration of a block. Due to its construction, the block has only one vertex adjacent to the disagreeing vertex $w$[7]. Let us call this vertex $v$. Due to (b), above, any path that connects $w$ and a high degree vertex in $B$ should be of "low weight".

We consider $T$, the *tree of self-avoiding walks* of $B$ rooted at vertex $v$. That is, at level $i$ of $T$ we have the vertices of $B$ which are reachable from $v$ by a path (within $B$) of length $i$. In the independent process, let $L_i^T$ denote the expected number of paths of disagreements in $T$, that connect the root and the vertices at level $i$ of $T$. The disagreement probabilities are specified in (3.2).

For the rapid mixing condition (3.1), it will suffice to show that

$$L_i^T \leq \beta(1-\theta)^i \qquad \text{for } i \geq 0, \tag{3.4}$$

for appropriate $\theta < 1, \beta > 0$. Since $L_i^T$ depends only on the first $i$ levels of $T$, we can neglect all but the first $i$ levels of $T$.

Assume that the degree of the root of $T$ is equal to $s$, for some integer $s > 0$. The condition in (3.4) reduces to the following: For each subtree $T'$ rooted at a child of the root of $T$ it should hold that

$$L_{i-1}^{T'} \leq \frac{\beta(1-\theta)^i}{s \cdot \varrho_v}, \tag{3.5}$$

where $\varrho_v$ is the probability of disagreement for the root of $T$. That is, instead of (3.4) it suffices to show (3.5) for every $T'$.

Applying the same inductive argument many times we get an increased number of conditions each of which considers a smaller subtree of $T$. After sufficiently many applications the conditions consider only single vertex subtrees. In particular each of these subtrees $T''$ contains exactly one vertex which belongs to the level $i$ of $T$. Then, instead of (3.4), it suffices to have the following, conditions: For every subtree $T''$, which contains only the vertex $u$, it should hold that

$$L_0^{T''} \leq \frac{\beta(1-\theta)^i}{\prod_{x \in L \setminus \{u\}} (\Delta(x) \cdot \varrho_x)}, \tag{3.6}$$

[6] degree greater than $(1+\alpha)d$

[7] This will become apparent to the reader once we describe in detail how do we construct the blocks.

where $L$ is the path in $T$ from the root to vertex $u$. Observe that on the l.h.s. of (3.6) we have the expected number of paths of disagreement in $T''$. It holds that $L_0^{T''} \leq p_u$.

From all the above, we conclude that it suffices that the following conditions hold: For every vertex $u$ at level $i$ of $T$ it should hold that

$$\frac{\beta(1-\theta)^i}{\prod_{x\in L\setminus\{u\}}(\Delta(x)\cdot\varrho_x)} \geq \varrho_u. \tag{3.7}$$

The reader should observe that to a certain extent the denominator in (3.7) can be controlled by choosing appropriately the parameters of the weighting schema and $k$. Indeed, setting appropriately these parameters and having $k \geq \frac{11}{2}d$, (3.7) holds for every $u$ at level $i$ of $T$.

REMARK 3. *We follow exactly the same approach to show rapid mixing of Markov chains for the hard core model. The partitioning of the vertices of the underlying graph is exactly the same. The only difference is the probabilities that the disagreements propagate inside the block $B$.*

## 4 Detailed description of the algorithm

**4.1 Block Creation** Consider the graph $G(n, d/n)$ and the weighting schema from (3.3). That is, each vertex $u$, of degree $\Delta(u)$, is assigned weight

$$W(u) = \begin{cases} (1+\gamma)^{-1} & \text{if } \Delta(u) \leq (1+\alpha)d \\ d^c \cdot \Delta(u) & \text{otherwise,} \end{cases}$$

for appropriate $\alpha, \gamma, c > 0$. Given the weight for each vertex, we introduce the concept of "influence".

DEFINITION 1. (INFLUENCE) *For a vertex $v$, let $\mathcal{P}(v)$ denote the set of all paths of length at most $\frac{\ln n}{d^{2/5}}$ that start from $v$. We call "influence" on the vertex $v$, denoted as $E(v)$, the following quantity:*

$$E(v) = \max_{L\in\mathcal{P}(v)} \left\{ \prod_{v\in L} W(v) \right\}.$$

When for some vertex $v$ it holds that $E(v) \leq 1$, then this means that "it is on light paths only". Such vertices are special for the block construction as the boundaries of the blocks are specified by using exclusively this kind of vertices.

DEFINITION 2. (BREAK-POINTS & INFLUENCE PATHS) *A vertex $v$ such that $E(v) \leq 1$ is called* "break-point". *Also, a path $L$ that does not contain break-points is called* "influence path".

Observe that in order to check whether some vertex $v$ is a break-point we only need to check paths of length at most $\frac{\ln n}{d^{2/5}}$ around it. Clearly this leaves open the possibility that there is a longer heavy path that emanates from $v$. W.h.p. over $G(n, d/n)$, it turns out that this is not the case for every break point.

Given the set of break points in $G(n, d/n)$ we can proceed with the creation of blocks.

**Block Creation:** We have two different kinds of blocks. For this, let $\mathcal{C}$ denote the set of all cycles of length at most $4\frac{\ln n}{\ln^5 d}$ in $G(n, d/n)$.

1. For each cycle $C \in \mathcal{C}$ we have a block which contains every vertex $v \in C$ as well as all the vertices that are reachable from $v$ through an influence path that does not use vertices of $C\setminus\{v\}$.
2. The remaining blocks are created as follows: Pick a vertex $v$ whose block is not specified yet.
   - if $v$ is a break-point, then $v$ is a block by itself
   - otherwise, the block of $v$, contains $v$ and all the vertices that are reachable from $v$ through an influence path.

REMARK 4. *Observe that the outer boundary of a multi-vertex block $B$ contains only break points.*

A typical instance of $G(n, d/n)$ contains a plethora of break points. Taking a path $L$ in $G(n, d/n)$ we expect that the product of the weights of its vertices is rather low. Mainly, this is due to the fact that only a very small fraction of vertices has large weight. E.g. Chernoff bounds imply that for each $u \in L$ it holds that $Pr[\Delta(u) > (1+\alpha)d] \leq \exp\left(-\alpha^2 d/3\right)$.

REMARK 5. *For a path $L$ in $G(n, d/n)$ with $|L|$ vertices, we show that the probability that its weight is greater than 1 is at most* $\exp(-d^{0.8}|L|)$.

The above remark implies that w.h.p. the following scenario holds: When we construct the blocks that contain the cycles in $\mathcal{C}$, the paths we add around each cycle are rather short. Since the distance between any two cycles in $\mathcal{C}$ is large no two cycles in $\mathcal{C}$ are going to end up in the same block. That is, the set of block $\mathcal{B}$ consist of blocks that are trees with at most one extra edge. In Section 5 we give a high level description of our arguments which show that the influence paths w.h.p. cannot be too long.

The fact that w.h.p. $\mathcal{B}$ contains blocks that are trees with at most an extra edge, turns out to be important when we consider the algorithms that implement the updates of the chains.

The main result of this section is the following theorem.

THEOREM 4.1. *There is a small fixed $s > 0$ such that with probability $1 - o(1)$ over the instances of $G(n, d/n)$ the following holds: The construction of the set of blocks $\mathcal{B}$ can be made in time $O(n^s)$. Each block in $\mathcal{B}$ is either a tree or a unicyclic graph.*

The proof of Theorem 4.1 see in full version of this work in [9].

**4.2 Acquiring the initial state** For the chain over the independent sets (hard core model) acquiring an initial state is trivial. I.e. it suffices to consider the *empty* independent set as the initial state of the chain.

For the chain over the colourings we can get an initial state by using the algorithm suggested in [13]. The authors there provide a greedy, polynomial time algorithm which $k$-colours typical instances of $G(n, d/n)$ for any $k \geq (1 + c)d/\ln d$ and any fixed $c > 0$. Clearly here we consider much larger number of colours.

**4.3 Update Algorithm** For the update algorithm we need to bear in mind that Theorem 4.1 implies that, w.h.p. the set $\mathcal{B}$ contains only blocks that are trees with at most one extra edge. For such simple structured blocks one may find many different ways of implementing the transitions of the chain efficiently. In what follows we describe such an approach.

For both $\mathcal{M}_c$ and $\mathcal{M}_{hc}$, the update of the configuration of a block $B \in \mathcal{B}$ is done by assigning appropriately spins to the vertices in some predefined order. That is, having fixed the spin of the vertices up to some vertex $u$, we assign spin to the next vertex $v$ by working as follows: We compute the distribution of the spin on the vertex $v$ given the configuration on the previous vertices in the ordering and the boundary conditions of $B$. Once we have this distribution, we can assign spin to the vertex $v$ appropriately. The critical issue is how do we compute this distribution. Depending on the model we follow a different approach.

For the chain over the $k$-colouring of $G(n, d/n)$ it suffices to count the number of $k$-colouring of $B$ which assign $v$ colour $c$, for every $c \in [k]$. So as to achieve that we use the Dynamic Programming algorithm for counting colourings suggested in [4] (See the algorithm in Section 3.3 of [4]). For fixed $k$ this algorithm is polynomial in $n$.

For the chain over the, two spin, independent sets it suffices to compute $p^v_{occ}$ the probability of the vertex $v$ to be "occupied" (to be in the independent set). We can use the algorithm by D. Weitz in [29] to compute this probability. It is straightforward to show that this algorithm computes exactly $p^v_{occ}$ for the kind of blocks we consider here in polynomial time [8].

## 5 Short Influence Paths

In this section we present a high level description of the arguments we use for proving that the influence paths we consider in the construction of $\mathcal{B}$ is not too long w.h.p. This would imply that the blocks in $\mathcal{B}$ are trees with at most one extra edge.

We call *elementary* every path $L$ in $G(n, d/n)$ such that there is no cycle shorter than $10 \ln n/d^{2/5}$ which contains two vertices of $L$.

Consider $G(n, d/n)$ and let the parameters of the weighting schema be fixed numbers $\gamma, c > 0$, $\alpha \in (0, 3/2)$. Let the set $\mathbb{U}$ contain all the *elementary* paths in $G(n, d/n)$ of length $\frac{\ln n}{\ln^5 d}$ that do not have any break-point. Then, we show that it holds that

$$Pr[\mathbb{U} \neq \emptyset] \leq n^{-\frac{1}{3}\frac{\gamma}{1+\gamma}\ln d}. \tag{5.8}$$

As w.h.p. the distance between any two cycles in $\mathcal{C}$ is much larger than $\frac{\ln n}{\ln^2 d}$, the above condition is sufficient to provide the desired kind of blocks in $\mathcal{B}$.

We show (5.8) by using the first moment method. Let $L = v_1, \ldots, v_{|L|}$ be an *elementary* path in $G(n, d/n)$ of length $T = \frac{\ln n}{\ln^5 d}$. Also, let $\varrho_L$ be the probability that $L$ does not have a break point. It holds that

$$Pr[\mathbb{U} \neq \emptyset] \quad \leq \quad E[|\mathbb{U}|] \leq \binom{n}{T+1} \cdot \left(\frac{d}{n}\right)^T \cdot \varrho_L \leq n^{1.1} \cdot \varrho_L.$$

It suffices to bound appropriately $\varrho_L$.

So as to figure out whether the path $L$ has a break point or not we should examine not only the weights of the subpaths of $L$ but the weights of paths that intersect $L$, as well. That is, for a vertex $v \in L$ we should examine all the paths of length at most $\ln n/d^{2/5}$ that emanate from it. For technical reasons we make the following distinction of the vertices on $L$.

DEFINITION 3. *A vertex $v_i \in L$ is called* left-break *or* right-break *for $L$ if it has the corresponding property below:*

**left-break:** *There is no path $L' \in \mathcal{P}(v_i)$ such that $\prod_{v_s \in L'} W(v_s) > 1$ and $L' \cap L$ contains $v_j$ for $j \leq i$,*

**right-break:** *There is no path $L' \in \mathcal{P}(v_i)$ such that $\prod_{v_s \in L'} W(v_s) > 1$ and $L' \cap L$ contains $v_j$ for $j \geq i$.*

*($W(\cdot)$ is the weight defined in Section 3.1. For the definition of the set $\mathcal{P}(v_i)$ see in Definition 1.)*

[8] Actually Weitz's algorithm is efficient for computing marginals even when someone considers even larger neighbourhoods of $G(n, d/n)$ than what we consider here, e.g. see [23].

REMARK 6. *All the paths coming from the "right side" of a right-break should be "light". We don't know what happens with its paths coming from the left-side (since $L$ is elementary). An analogous statement holds for the left-breaks. It is clear that $v_i \in L$ is break point* if and only if *it is both a left-break and a right-break for $L$.*

Let $Y_l$ and $Y_r$ be the number of left breaks and the number of right breaks for $L$, respectively. Let the event

$$\mathcal{S} = \text{"either } Y_l \text{ or } Y_r \text{ is smaller than } 0.9|L|\text{"}.$$

It holds that $\varrho_L \leq Pr[\mathcal{S}]$. To see this, observe that if $\mathcal{S}$ does not hold then the number of vertices which are at the same time right-break points and left-break points is at least $0.8|L|$. That is, the number of break point is at least $0.8|L|$.

It is not hard to see that the random variables $Y_l$ and $Y_r$ are symmetric, i.e. identically distributed. Using this symmetry and a simple union bound we get that $Pr[\mathcal{S}] \leq 2Pr[Y_l < 0.9|L|]$. The focus now is only on $Y_l$. It suffices to show that

$$Pr[Y_l < 0.9|L|] \leq 4n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}. \tag{5.9}$$

To derive the probability bound above we have to build on the weighting schema of Section 3.1. Somehow we need to accommodate the weight of paths that intersect $L$. For this reason we introduce a new weighting schema which is slightly different than the one we have in Section 3.1.

For each $v_i \in L$ let $N_i$ denote the set of the vertices outside $L$ which are adjacent to $v_i$, i.e. $|N_i| = \Delta(v_i) - 2$. For every $w \in N_i$ let $E_{out}(w)$ denote the maximum influence on vertex $w$, only from paths of length at most $\ln n/d^{2/5}$ that do not use vertices in $L$. For every $v_i$ let

$$Q(v_i) = \max_{w \in N_i} \{E_{out}(w)\}.$$

Now, we associate each $v_i \in L$ with the following quantity

$$U(v_i) = \begin{cases} \frac{\max\{1, Q(v_i)\}}{1+\gamma} & \text{if } \Delta(v_i) \leq (1+\alpha)d \\ \max\{1, Q(v_i)\}\,\Delta(v_i)d^c & \text{otherwise,} \end{cases}$$

where $\alpha, \gamma, c$ are specified in (3.3).

By definition, every $v_i \in L$ such that $U(v_i) > 1$ cannot be a left break for $L$. Let $H$ be the set of vertices in $L$ with large weight, i.e. $H = \{v_j \in L | U(v_j) > 1\}$. Towards distinguishing the left breaks we consider the following: for each $v_j \in H$ let $\mathcal{R}_j = v_j, v_{j+1}, \ldots, v_s$ be the maximal subpath of $L$ such that for any $j' \in [j, s]$ it holds $\prod_{r=j}^{j'} U(v_r) > 1$. Somehow, $\mathcal{R}_j$s contain vertices which are on some heavy path[9]. Furthermore, we show that any vertex $v_{j'} \notin H$ which also does not belong to a set $\mathcal{R}_j$, for any $j$, is a left break. That is the vertices that do not belong to $\bigcup_j \mathcal{R}_j$ are left breaks.

Letting $\mathcal{R} = \cup_j \mathcal{R}_j$, the above implies that $Y_l \geq |L| - |\mathcal{R}|$. It suffices to get an appropriate tail bound for $|\mathcal{R}|$, i.e. we need to bound $Pr[|\mathcal{R}| \geq 0.1|L|]$. For computing this tail bound we make heavy use of moment generating function of $|\mathcal{R}|$. There we need to prove the following, very interesting tail bound on the weight of an elementary path $P$, i.e. $\prod_{u \in P} U(u)$

Let $P$ be an *elementary* path in $G(n, d/n)$ with a number of vertices $|P| \leq \frac{\ln n}{\ln^2 d}$. Let $AC(P) = \prod_{u \in P} U(u)$. For any $\delta > 0$ it holds that

$$Pr[AC(P) \geq \delta] \leq 2\exp\left[-d^{7/10}\left(|P| + \ln \delta\right)\right].$$

## 6 Proof of Theorem 3.1

**Ergodicity.** For Theorem 3.1 first we need to address the problem of *ergodicity* for $\mathcal{M}_c$. From [4] we have that the Glauber dynamics (and hence the block dynamics we have here) is ergodic with probability $1 - o(1)$ over the instances $G(n, d/n)$ when $k \geq d + 2$.

REMARK 7. The proof for ergodicity in [4] goes as follows: They show that if a graph $G$ has no $t$-core[10], then for any $k \geq t + 2$ the Glauber dynamics over $k$-colourings is ergodic. Then the authors use the result in [26], i.e. w.h.p. $G(n, d/n)$ has no $t$-core for $t \geq d$.

So as to prove rapid mixing, we require that the underlying graph $G(n, d/n)$ has the following three properties: (A) It can be coloured with $d$ colours or less. (B) Take $0 < \gamma \leq \alpha \leq 10^{-2}$ and $c = 10$. Each block $B \in \mathcal{B}(\alpha, \gamma, c)$ is either a tree or a unicyclic graph. (C) For the same $\alpha, \gamma, c$ and for each $B \in \mathcal{B}$ the following holds: between every vertex $u$ at the outer boundary of $B$ and every high degree[11] vertex $u' \in B$ there is no path $L$ such that $\prod_{u'' \in L} W(u'') > 1$.

LEMMA 6.1. *With probability $1 - o(1)$ the graph $G(n, d/n)$ satisfies all the conditions (A), (B) and (C).*

For a proof of Lemma 6.1 see in the full version of this work in [9].

Let $\mathcal{G}_d^A$ denote the family of graphs such that they have no $t$-core for $t \geq d$ and they satisfy conditions (A), (B) and (C). We show that the chain $\mathcal{M}_c$ has a rapid mixing for any $k \geq \frac{11}{2}d$, as long as the underlying graph $G(n, d/n) \in \mathcal{G}_d^A$. We show rapid mixing by using

[9] "Heavy" is w.r.t. the weighting schema of Section 3.1.

[10] A graph without a $t$-core can have its vertices ordered $v_1, \ldots v_n$ such that $v_i$ has fewer than $t$ neighbours in $\{v_1, \ldots, v_{i-1}\}$.

[11] degree greater than $(1+\alpha)d$

path coupling. In particular, Theorem 3.1 follows as a corollary of Lemma 6.1 and the following result.

THEOREM 6.1. *Let $\mathcal{M}_c$ be such that the underlying graph $G(n, d/n) \in \mathcal{G}_d^A$ and $k \geq \frac{11}{2}d$. Consider $X$ and $Y$, two states of $\mathcal{M}_c$ such that $H(X,Y) = 1$. Let $P_X, P_Y$ be the one step transition probabilities of $\mathcal{M}_c$ given that it starts from $X$, $Y$, respectively. There is a coupling $\nu_{X,Y}$ of $P_X$ and $P_Y$ such that for $(X', Y')$ distributed as in $\nu_{X,Y}$ it holds that*

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - 0.1225/n, \tag{6.10}$$

*where $E_{\nu_{X,Y}}[\cdot]$ denotes the expectation w.r.t. the distribution $\nu_{X,Y}$.*

The proof of Theorem 6.1 appears in Section 6.1.

**6.1 Proof of Theorem 6.1** Consider the chain $\mathcal{M}_c$ with underlying graph $G(n, d/n) \in \mathcal{G}_d^A$ and $k \geq \frac{11}{2}d$. Let $X, Y$ be two states of $\mathcal{M}_c$ at hamming distance 1 from each other, i.e. $H(X,Y) = 1$. Furthermore, assume for the vertex $w$ that we have $X(w) \neq Y(w)$. We are going to show that there is a coupling $\nu_{X,Y}$ of $P_X$ and $P_Y$ that has the property stated in (6.10). For this, we construct a pair $(X', Y')$ distributed as in $\nu_{X,Y}$ by considering the two copies of $\mathcal{M}_c$ at states $X, Y$, respectively, and by coupling the next step of these two chains. The coupling is such that both chains choose to update the configuration of the same block.

There are cases where we update the colouring of a single vertex block, i.e. a break-point. There we use the notion of *maximal coupling* transition, as defined in [22]. That is, assume that the single vertex block is the vertex $u$[12]. The update of the colouring of $u$ is such that the probability of having $X'(u) \neq Y'(u)$ is minimized. To do this we work as follows: Let $L_X, L_Y$ be the set of colours not appearing in the neighbourhood of $u$ under the colouring $X$ and $Y$, respectively. Since we assume $k > (1+\alpha)d$, both sets are not empty. Take two mappings $f_X : [0,1] \to L_X$ and $f_Y : [0,1] \to L_Y$ such that

- for each $c \in L_X$, $|f^{-1}(c)| = \frac{1}{|L_X|}$ and similarly for $Y$,
- $\{x : f_X(x) \neq f_Y(x)\}$ is as small as possible in measure.

Then, take a uniformly random real $\mathcal{U} \in [0,1]$ and choose colour $f_X(\mathcal{U})$ for $X'(u)$ and $f_Y(\mathcal{U})$ for $Y'(u)$.

Consider now the coupling of the transitions of two chains at state $X, Y$ so as to construction of $(X', Y')$. Let $B$ be the block whose colouring is chosen to be updated in the coupling. Let $N = |\mathcal{B}|$, i.e. $B$ is chosen with probability $1/N$. We need to consider three cases for the relative position of the block $B$ and the disagreeing vertex $w$.

`Case 1:` The disagreeing vertex $w$ is internal in a block $B'$, i.e. it is not adjacent to any vertex outside $B'$. With probability $1/N$ in the coupling we have $B = B'$. Then, we can coupe $X'$ and $Y'$ such that $H(X', Y') = 0$, as the boundary conditions of $B$ in both chains are identical. Also, with the remaining probability we have $H(X', Y') = 1$. It follows that

$$E_{\nu_{X,Y}}[H(X', Y')] = 1 - 1/N.$$

`Case 2:` The disagreeing vertex $w$ belongs to some block $B'$ but it is adjacent to some vertices outside $B'$. Also, assume that $B'$ contains more than one vertices. It is easy to check that $\Delta(w) \leq (1+\alpha)d$ (otherwise condition (C) would be violated). Also, it holds that all the blocks that are adjacent to $w$ are single vertex blocks, i.e. they are break-points.

For this case the following holds: If the coupling does not choose $B'$ or some single vertex block adjacent to $B'$, then we have $H(X', Y') = 1$. On the other hand, with probability $1/N$ the coupling chooses to update the block $B'$. Then, as in Case 1, we get that $H(X', Y') = 0$. Also, with probability at most $(1+\alpha)d/N = 1.01d/N$ the block $B$ that the coupling chooses to update is a break-point adjacent to $w$. Then, the probability that the break-point gets different colour assignment in $X'$ and $Y'$ is at most $\frac{1}{k-(1+\alpha)d} = \frac{1}{4.49d}$, as $k \geq \frac{11}{2}d$. It follows that

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - 0.75/N.$$

The update of the colouring of $u$, a neighbour of $w$, yields a disagreement with probability at most $\frac{1}{k-(1+\alpha)d}$ for the following reason: We use maximal coupling. For both $L_X, L_Y$, the two lists of available colours for $u$. It holds that $|L_X|, |L_Y| \geq k-(1+\alpha)d$. Either $|L_X| = |L_Y|$ and each list contains exactly one colour that does appear in the other, or only one of $L_X, L_Y$ contains one extra colour that the other does not have. In any case, maximal coupling implies that the probability of picking a different random element from the lists $L_X$ and $L_Y$ cannot be bigger than $\frac{1}{4.49d}$.

`Case 3:` The vertex $w$ is itself a block, i.e. $w$ is a break-point. If we don't have $B = \{w\}$ or $B$ adjacent to $w$, then we have $H(X', Y') = 1$. Otherwise, with probability $1/N$ the coupling chooses to update $w$, i.e. $B = \{w\}$. Then we get $H(X', Y') = 0$. Also, with probability at most $(1+\alpha)d/N = 1.01d/N$, $B$ is one

[12] This vertex should have small degree, i.e. at most $(1+\alpha)d$.

of the blocks that are adjacent to $w$. W.l.o.g. assume that $B$ consist of more than one vertex[13]. We need to bound the expected number of disagreements $R_B$ that are generated inside $B$ when its colouring is updated. To this end we use the following proposition.

PROPOSITION 6.1. *For Case 3, there is a coupling such that* $E[R_B] \leq \frac{0.8688}{d}$.

Using Proposition 6.1 and the fact that $\alpha = 10^{-2}$ and $N \leq n$, we get that

$$\begin{aligned} E_{\nu_{X,Y}}[H(X',Y')] &\leq 1 - \frac{1}{N} + \frac{(1+\alpha)d}{N} E[R_B] \\ &\leq 1 - \frac{0.1225}{n}. \end{aligned}$$

The theorem follows.

**6.2 Proof of Proposition 6.1** For the proof of the proposition we use an approach similar to one used in [4] (the proof of Theorem 1 (a)). That is, we use the well-known "disagreement percolation" coupling construction in [2]. We wish to couple $X'(B)$ and $Y'(B)$ as close as possible. However, identical coupling is precluded due to the disagreement at the boundary of $B$. The disagreement percolation assembles the coupling in a stepwise fashion, i.e. couples the colouring of each vertex of $B$ at a time. It gives priority to vertices which are next to a disageeing vertex, if any. Disagreements propagates into $B$ along paths from $w$. A disagreement at vertex $u' \in B$ at distance $r$ from $w$ propagates to a neighbour $u$ at distance $r+1$ if $X'(u) \neq Y'(u)$.

LEMMA 6.2. *There is a coupling such that the following is true: In the coupling above, it holds that* $Pr[X'(u) \neq Y'(u)] \leq \rho_u$, *where*

$$\rho_u = \begin{cases} \frac{2}{k-(1+\alpha)d} & \text{if } \Delta(u) \leq (1+\alpha)d \\ 1 & \text{otherwise.} \end{cases}$$

A proof of Lemma 6.2 appears in [4]. However, in Section 6.3 we provide a different proof of this lemma.

The disagreement percolation is dominated by an independent process, i.e. the disagreement propagates over the path $L$ that start from $w$ with probability at most $\prod_{u \in L \setminus \{w\}} \varrho_u$.

Consider a configuration of the vertices in $B$ such that each vertex $u \in B$ is either "disagreeing" or "agreeing". Let the product measure $\mathcal{P}$ which specifies that the vertex $u$ is disagreeing with probability $\rho_u$. In this context we call a "path of disagreement" every self avoiding path [14] that has all of its vertices disagreeing.

From the construction of the blocks it is direct that the disagreeing vertex $w$ has only one neighbour in $B$. Let $v_0$ be the only vertex in $B$ which is adjacent to $w$. Consider a "agreeing-disagreeing" configuration of the vertices in $B$ acquired according to the probability measure $\mathcal{P}$. In this configuration, let $Z_i$ be the number of paths of disagreements of length $i$ that start from the vertex $v_0$. It holds that

$$E[R_B] \leq \sum_{i \geq 0} E_{\mathcal{P}}[Z_i],$$

where $E_{\mathcal{P}}[\cdot]$ denotes the expectation taken w.r.t. the measure $\mathcal{P}$.

Instead of bounding $E_{\mathcal{P}}[Z_i]$ w.r.t $B$, it is equivalent to study the same quantity on the tree $T$, the tree of self-avoiding walks defined as follows: $T$ is rooted at vertex $v_0$. At level $i$, $T$ contains all the vertices in $B$ which are reachable from $v_0$ through a self-avoiding path inside $B$ of length exactly $i$. At this point we use the following lemma.

LEMMA 6.3. *Consider a tree* $H = (V, E)$ *containing vertices of different degrees. Let* $l_i$ *denote the set of vertices at the level* $i$ *in* $H$. *For a vertex* $v \in H$ *let* $L_v$ *denote the simple path connecting it to the root* $r$. *For* $L_v$ *we define the following weight:*

$$\mathcal{C}_{p,s}(L_v) = \prod_{u \in L_v} \left( \mathbb{I}_{\{\Delta(u) \leq s\}} \cdot p + \mathbb{I}_{\{\Delta(u) > s\}} \right),$$

*where* $p \in [0,1]$ *and* $s > 0$ *is an integer. Assume that for any vertex* $v$ *such that* $\Delta(v) > s$ *the following condition holds:*

$$\frac{\prod_{u \in L_v} \left( \frac{\mathbb{I}_{\{\Delta(u) \leq s\}}}{(1+\zeta)} + \delta^{10} \Delta(u) \mathbb{I}_{\{\Delta(u) > s\}} \right)}{(1+\zeta)} \leq 1, \tag{6.11}$$

*for some* $\zeta > 0$. *Assume that* $\delta, s$ *are sufficiently large, also* $\delta, s \gg (1+\zeta)$ *while both* $(s \cdot p), (\delta \cdot p) \in \left[ \frac{1}{100(1+\zeta)}, \frac{1}{1+\zeta} \right]$. *Then (6.11) implies that*

$$\sum_{v \in l_i} \mathcal{C}_{p,s}(L_v) \leq p \cdot (1-\theta)^i \qquad \forall i \geq 0, \tag{6.12}$$

*for any* $\theta \leq \min\{1 - ps(1+\zeta), 1 - (ps)^{9/10}\}$.

For the proof of Lemma 6.3 see in the full version of this work in [9]. The basic idea of the proof was presented at the final part of Section 3.1.

[13] It will be conceivable from the analysis that this is the worst case assumption.

[14] A path $L$ is called *self-avoiding* if there are no two $u_j, u_{j'} \in L$ such that $u_j = u_{j'}$.

At this point, observe the following: Let $P_v$ be a path in $T$ from the root $v_0$ to some vertex $v$. The probability that $P_v$ is a path of disagreement is upper bounded by the quantity $\mathcal{C}_{p,s}(P_v)$ where $p = \frac{2}{\frac{11}{2}d-(1+\alpha)d} = \frac{1}{2.245d}$, $s = (1+\alpha)d = 1.01d$. Thus, it holds that

$$E_{\mathcal{P}}[Z_i] \leq \sum_{v \in l_i(T)} \mathcal{C}_{p,s}(P_v),$$

where $l_i(T)$ is the set of vertices at level $i$ in $T$. If we set $\delta = d$ and $\zeta = \gamma$, each $P_v$ in $T$ satisfies the condition in (6.11) as the disagreeing vertex $w$ is a break-point and the root of $T$ is adjacent to $w$. Also, it holds that $sp, \delta p \in [\frac{1}{100(1+\gamma)}, \frac{1}{1+\gamma}]$. Thus, Lemma 6.3 implies that $E_{\mathcal{P}}[Z_i] \leq p(1-\theta)^i$. So as to bound $E_{\mathcal{P}}[Z_i]$ we take the maximum possible value for $\theta$, subject to the parameters $\delta, s, p, \zeta$. It is direct that $\theta \leq 0.5127$, that is

$$E_{\mathcal{P}}[Z_i] \leq \frac{0.44543}{d} (0.4873)^i .$$

From the above we get that $E[R_B] \leq \frac{0.44543}{d} \cdot \frac{1}{0.5127} = \frac{0.8688}{d}$, as promised.

**6.3 Proof of Lemma 6.2** For the moment assume that the block $B$ is a tree. The case where $B$ is a tree with an extra edge will follow directly. Assume, also, that the vertex $u$ is the first vertex that is going to be coloured in the coupling. Let

$$S_X = \left\{\sigma \in [k]^B \middle| \, Pr[X'(B) = \sigma | X'(V \backslash B)] > 0, \; \sigma_u \notin \{X'(w), Y'(w)\}\right\}.$$

That is, $S_X$ contains the colourings of $B$ which can be assigned to $B$ in the first chain and at the same time they do not use for $u$ a colour that appears in the disagreement. We define $S_Y$ in the same manner w.r.t. the second chain. It is trivial to verify that $S_Y$ and $S_X$ are identical. This implies that we can have a coupling such that

$$Pr[X'(u) = Y'(u) | X'(u), Y'(u) \notin \{X'(w), Y'(w)\}] = 1.$$

That is,

$$\begin{aligned} Pr[X'(u) \neq Y'(u)] &= Pr[X'(u), Y'(u) \in \{X'(w), Y'(w)\}] \\ &\leq \max\{Pr[X'(u) = Y'(w)], Pr[Y'(u) = X'(w)]\}. \end{aligned}$$

The last inequality follows from the maximal coupling of $X'(u), Y'(u)$. It is not hard to see that the maximum, above, is upper bounded by $\frac{1}{k-\Delta(u)}$. If $k \geq \Delta(u)$, then the maximum is upper bounded by 1.

If $u$ is not the first vertex to be coloured in the coupling, the situation is essentially the same. I.e. there is a set of vertices $C$ which is already coloured. $C$ is a connected subtree of $B$. Then vertex $u$ is the root at of an uncolored subtree of $B$, e.g. $T_u$. The interesting case is when $u$ is connected to a vertex in $C$ which is disagreeing. Then it is direct to see that the previous arguments apply directly to this case and we get the same bounds for $Pr[X'(u) \neq Y'(u)]$.

For the case where the block $B$ is a unicyclic graph we work similarly. The only difference, now, is that there are two paths from which the disagreement can reach the vertex $u$, as opposed to one path in the case of trees. This implies that in the worst case, $X'(u)$ (or $Y'(u)$ ) should avoid two colours so as there is no disagreement. I.e. for appropriate $c_1, c_2, c_3, c_4 \in [k]$, it holds that

$$\begin{aligned} Pr[X'(u) \neq Y'(u)] &\leq \\ &\leq \max\{Pr[X'(u) \in \{c_1, c_2\}], Pr[Y'(u) \in c_3, c_4]\}. \end{aligned}$$

If $\Delta(u) < k-1$, then the r.h.s. of the inequality above is upper bounded by $\frac{2}{k-\Delta(u)}$. Otherwise, i.e. $\Delta(u) \geq k-1$, we get the trivial bound 1. The lemma follows.

**Acknowledgement.** The author of this work would like to thank Prof. Amin Coja-Oghlan, Prof. Mike Molloy and Prof. Elchanan Mossel for their comments on the work, their remarks and suggestions for improvements. Also, I would like to thank the anonymous reviewers of SODA 2014 for their corrections and their helpful remarks.

## References

[1] D. Achlioptas, and A. Naor. *The two possible values of the chromatic number of a random graph.* In proc. of 36th STOC, pp. 587-593, 2004.

[2] J. van den Berg and C. Maes. *Disagreement percolation in the study of Markov fields.* Annals of Probability **22**, pp 749-763, 1994.

[3] R. Bubley and M. Dyer. *Path Coupling: a Technique for Proving Rapid Mixing in Markov Chains.* In proc. of 38th FOCS, pp 223-231, 1997

[4] M. Dyer, A. Flaxman, A. M. Frieze and E. Vigoda. *Random colouring sparse random graphs with fewer colours than the maximum degree.* Random Struct. and Algorithms **29**, pp. 450-465, 2006.

[5] M. Dyer, A. M. Frieze, A. Hayes, E. Vigoda. *Randomly colouring constant degree graphs.* In proc. of 45th FOCS, pp 582-589, 2004.

[6] M. Dyer, A. M. Frieze and M. Jerrum. *On counting independent sets in sparse graphs.* In proc. of 40th FOCS, pp 210-217, 1999.

[7] M. Dyer, C. S. Greenhill. *On Markov chains for independent sets.* In Journal of Algorithms **35** (1), pp 17-49, 2000.

[8] C. Efthymiou. *A simple algorithm for random colouring $G(n, d/n)$ using $(2+\epsilon)d$ colours.* In proc. of SODA 2012, pp 272-280

[9] C. Efthymiou. *MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold.* In http://arxiv.org/abs/1304.6666

[10] A. Frieze and E. Vigoda. *A Survey on the use of Markov Chains to Randomly Sample Colorings.* In Combinatorics, Complexity and Chance, Oxford University Press, 2007.

[11] A. M. Frieze, J. Vera. *On randomly colouring locally sparse graphs.* Discrete Mathematics & Theoretical Computer Science **8** (1), pp 121-128 2006

[12] L. A. Goldberg, R. A. Martin, M. Paterson. *Strong Spatial Mixing with Fewer Colors for Lattice Graphs.* SIAM J. Comput. **35** (2), pp 486-517 2005

[13] G. R. Grimmett and C. J. H. McDiarmid. *On colouring random graphs.* Math. Proc. Cambridge Philos. Soc., **77**, pp 313 - 324, 1975

[14] T. Hayes, J. Vera and E. Vigoda. *Randomly coloring planar graphs with fewer colors than the maximum degree.* In proc. of 39th STOC, pp 450-458, 2007.

[15] S. Janson, T. Luczak and A. Rucinski. *Random graphs.* J. Wiley and Sons, Inc. 2000.

[16] M. Jerrum. *Counting, Sampling and Integrating: Algorithms and Complexity,* Lectures in Mathematics, ETH Zurich. Birkhäuser-Verlag, 2003.

[17] M. Jerrum and A. Sinclair. *The Markov chain Monte Carlo method:an approach to approximate counting and integration.* In Approximation Algorithms for NP-hard problems, PWS 1996. (Dorit, Hochbaum ed.)

[18] F. Kelly. *Stochastic models of computer communication systems.* J. R. Statist. Soc. B **47**, (3), pp 379 - 395, 1985.

[19] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems.* In PNAS **104**, pp 10318-10323, 2007.

[20] D. Levin, Y. Peres and E. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society (2008).

[21] F. Martinelli, A. Sinclair and D. Weitz. Fast mixing for independent sets, colorings and other models on trees. In proc. of 15th SODA, pp 456-465, 2004.

[22] M. Molloy. *The Glauber dynamics on the colourings of a graph with large girth and maximum degree.* SIAM J. Computing **33**, pp 721-737, 2004.

[23] E. Mossel and A. Sly. *Rapid mixing of Gibbs sampling on graphs that are sparse on average..* Random Struct. Algorithms **35** (2), pp 250-270, 2009

[24] E. Mossel and A. Sly. *Gibbs Rapidly Samples Colorings of $G_{n,d/n}$.* In journal Probability Theory and Related Fields, **148**, (1-2), 2010.

[25] A. Sinclair, P. Srivastava and Y. Yin. *Spatial mixing and approximation algorithms for graphs with bounded connective constant.* To appear appear in FOCS 2013. Full version http://arxiv.org/abs/1308.1762

[26] B. G. Pittel, J. Spencer and N. C .Wormald. *Sudden emergence of a giant $k$-core in a random graph.* Journal of Combinatoral Theorey B **67**, pp 111-151, 1996.

[27] E. Vigoda. *Improved bounds for sampling colorings.* Journal of Mathematical Physics, **41** (3), pp 1555-1569, 2000. A preliminary version appears in FOCS 1999.

[28] E. Vigoda. *A note on the glauber dynamics for sampling independent sets.* Elctron. J. Comb. **8** (1), pp 1-8, 2001.

[29] D. Weitz. *Counting independent sets up to the tree threshold.* In Proc of STOC 2006.

[30] L. Zdeborová, F. Krzakala. *Phase transition in the Coloring of Random Graphs* Phys. Rev. E **76**, 031131 (2007). preprint: arxiv:0704.1269 .