Theoretical Computer Science 407 (2008) 134–154

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Random sampling of colourings of sparse random graphs with a constant number of colours

Charilaos Efthymiou *,[1], Paul G. Spirakis [1]

Research Academic Computer Technology Institute, N. Kazantzaki str. Rio Patras, 26500, Greece
Computer Engineering and Informatics Department of the University of Patras, 26500, Greece

ARTICLE INFO

Article history:
Received 11 January 2008
Received in revised form 5 May 2008
Accepted 11 May 2008
Communicated by J. Diaz

Keywords:
Sampling colourings
Sparse random graph
Spin system

ABSTRACT

In this work we present a simple and efficient algorithm which, with high probability, provides an almost uniform sample from the set of proper k-colourings on an instance of sparse random graphs $G_{n,d/n}$, where $k = k(d)$ is a sufficiently large constant. Our algorithm is not based on the Markov Chain Monte Carlo method (M.C.M.C.). Instead, we provide a novel proof of correctness of our algorithm that is based on interesting "spatial mixing" properties of colourings of $G_{n,d/n}$. Our result improves upon previous results (based on M.C.M.C.) that required a number of colours growing unboundedly with n.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

For a graph $G = (V, E)$, a (proper) k-colouring is an assignment $\sigma : V \to [k]$ such that adjacent vertices receive different colours, where for some positive integer k, $[k]$ indicates the set $\{1, \ldots, k\}$. It is well known that it is NP-hard to estimate the minimum number of colours in a proper k-colouring, i.e. estimate the chromatic number of G. However, in many cases there are estimates and upper bounds of the chromatic number e.g. if Δ is the maximum degree of G, then one can k-colour G for $k = \Delta + 1$. Furthermore, for special classes of graphs the chromatic number has been estimated accurately, e.g. in [1], Achlioptas and Naor have found the two possible values of the chromatic number that an instance of a sparse random graph has with high probability (w.h.p.), i.e. with probability that tends to 1 as the size of the graph tends to infinity. All these facts raise the interesting computational challenge of finding the number of proper k-colourings for k greater than the chromatic number.

In [13], Valiant introduced the notion of $\#P$-hardness and proved that counting k-colourings is $\#P$-complete. The existence of a polynomial-time algorithm for exact counting is considered highly unlikely. Thus, we focus on designing polynomial-time algorithms for *approximate counting*. Practically, the closer k gets to the chromatic number of G, the more difficult it becomes to estimate the number of its k-colourings. By [7] and [8] we can reduce the *estimation of the number* of k-colourings of G to *sampling* almost uniformly from the set of all its proper k-colourings. By "almost" we mean with distribution close, in some sense, to the uniform distribution.

In this work, we focus on sampling k-colourings of instances of a sparse random graph, i.e. random graphs with vertices having an expected degree equal to some constant, d, and $k = k(d)$ is a sufficiently large constant which scales as d^{14}, for sufficiently large d.

Definition 1.1. Let n be a positive integer and p, $0 \le p \le 1$. The random graph $G_{n,p}$ is a probability space over the set of graphs on the vertex set $\{1, \ldots, n\}$ determined by

* Corresponding author at: Research Academic Computer Technology Institute, N. Kazantzaki str. Rio Patras, 26500, Greece. Tel.: +30 2610960459.
E-mail addresses: euthimio@ceid.upatras.gr (C. Efthymiou), spirakis@cti.gr (P.G. Spirakis).

1 Partially supported by E.U. within 7th Framework Programme under contract 215270 (FRONTS).

0304-3975/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2008.05.008

$\Pr[\{i, j\} \text{ is an edge of } G] = p$

with these events being mutually independent.

For a sparse random graph the parameter p is of the form $p = d/n$, where d is a positive real constant. We take $d > 1$ (otherwise the problem of sampling is trivial).

The mathematical tool that we use for studying the problem of sampling k-colourings of instances of a sparse random graph is the "spin systems" and more specifically the proper colouring model, also referred to as *antiferromagnetic Potts model at zero temperature* in statistical physics.

Colouring Model on a finite graph. The colouring model on a finite graph $G = (V, E)$ and set of colours $[\mathcal{S}]$, for some positive integer $\mathcal{S}$, is defined as follows. The system consists of a set of *sites*, which correspond to the vertices of G, and each site is assigned a *spin*, i.e. a member of $[\mathcal{S}]$. A *configuration* is an assignment of spins to V. Not all configurations can occur in the colouring model. A configuration that may occur is called a *feasible configuration*. The set of feasible configurations is the set of proper $\mathcal{S}$-colourings of the underlying graph G.

For a colouring model with underlying graph $G = (V, E)$ which uses $\mathcal{S}$ colours, the vertices-sites interact with each other so that the following holds: For any vertex set $V' \subseteq V$, let $\partial V' = \{v \in V \setminus V' | \ \exists u \in V' \ s.t. \ \{u, v\} \in E\}$. Consider the colouring $C(\partial V') \in [\mathcal{S}]^{\partial V'}$, which is such that there is a proper colouring in $[\mathcal{S}]^V$ with the vertices in $\partial V'$ coloured as $C(\partial V')$. In a system where the vertices in $\partial V'$ are coloured as $C(\partial V')$, the colour assignments of the vertices in V' are distributed uniformly over all the colour assignments of the vertices in $V' \cup \partial V'$ that agree with $C(\partial V')$ on the vertices in $\partial V'$.[2]

Frequently, one imposes *boundary conditions* on the model, which correspond to fixing the colour assignment at some "*boundary*" vertex set of G; we use the term "*free boundary*" when there are no boundary conditions specified.

Definition 1.2. For a finite graph $G = (V, E)$ and an integer $\mathcal{S}$, let $PCS(G, \mathcal{S}, C(L))$ be a colouring model with underlying graph G, with feasible configurations containing all the proper $\mathcal{S}$-colourings of G and with the boundary $L \subseteq V$ coloured as $C(L)$. Let $\Omega(G, \mathcal{S}, C(L))$ be the set of feasible configurations of the system.

The omission of the boundary conditions parameter implies *free boundary*. If the first parameter is a class of random graphs, e.g. $G_{n,p}$, then we consider that the underlying graph is an instance of this class.

Clearly, $\Omega(G, \mathcal{S}, C(L))$ is the set of all proper $\mathcal{S}$-colourings of G that have the vertices in the set $L \subseteq V$ coloured as specified by the assignment $C(L)$. For a system $PCS(G, \mathcal{S}, C(L))$ we always assume that the boundary $C(L)$ is such that $\Omega(G, \mathcal{S}, C(L)) \neq \emptyset$.

For convenience, we use the following notation rules throughout this work: In $PCS(G = (V, E), \mathcal{S}, C(\Lambda))$, for $\Lambda \subseteq V$ the colour assignment of each vertex $v \in V$, or the set of vertices $V' \subseteq V$ are considered to be equal to the random variables $X_v^{C_\Lambda} \in [\mathcal{S}]$ and $X_{V'}^{C_\Lambda} \in [\mathcal{S}]^{V'}$, correspondingly.

The probability of finding a colouring model at a specific configuration is the uniform distribution over all proper colourings of the underlying graph. Generally, the probability of finding a system in a specific configuration is given by the *Gibbs* measure specified by this system.

Definition 1.3. For the system $PCS(G = (V, E), S)$, the function $\mu(\cdot) : 2^{[\mathcal{S}]^V} \to [0, 1]$ indicates the Gibbs measure specified by this system.

In the system $PCS(G = (V, E), \mathcal{S}, C(L))$, for $\forall v \in V$, we denote with $\mu(X_v | C(L))$, the marginal Gibbs measure of the random variable X_v.

1.1. Our work and related work

Previous work. The pioneering work of Dyer et.al., in [4], proposes a very interesting Markov Chain Monte Carlo (MCMC) based algorithm, which with high probability (w.h.p.), i.e. with probability that tends to 1 as the size of the graph tends to infinity, provides an almost uniform sample from the set of proper colourings of $G_{n,d/n}$ which uses at least $\Theta(\frac{\log\log n}{\log\log\log n})$ colours. Noting that w.h.p. the maximum degree of a sparse random graph is $\Theta(\log n/\log\log n)$, to our knowledge, this work was the first to present a procedure for sampling colourings that uses fewer colours than the maximum degree.

In parallel and independently, E. Mossel and A. Sly have recently derived essentially the same result as we have here, (i.e. a random sampling k-colourings of a sparse random graph where k is a constant) by using an MCMC approach, [11].

Our work. Consider the system $PCS(G = (V, E), \mathcal{S})$, where G is an instance of $G_{n,d/n}$ and $\mathcal{S} = \mathcal{S}(d)$ is a sufficiently large integer. Here, we present an algorithm which, w.h.p. and in time $O(n^2)$, returns a $\mathcal{S}$-colouring of G[3] according to a probability measure which is *asymptotically* equal the Gibbs measure that the system specifies.

---

2 A rigorous definition of a colouring model involves the definition of a set of functions, the *compatibility functions* (see [15]). However, the definition we give here is a direct consequence of that with the compatibility functions.

3 A configuration of $PCS(G, \mathcal{S})$.

A possible schema of our algorithm is the following: The algorithm assumes an arbitrary permutation of the vertices of the input graph, e.g. $(v_1, v_2, \ldots, v_n)$, and in turn it assigns each of them a colour as follows: For $1 \leq i \leq n$, let $A_i \subseteq V$ be the set of $i-1$ first coloured vertices and let $C(A_i)$ be their colour assignment. Assume that the colouring $C(A_i)$, of the vertices in A_i, is done according to a probability measure which is sufficiently close to $\mu(X_{A_i} = C(A_i))$. Then, the algorithm computes efficiently a "good" estimation of $\mu(X_{v_i}|C(A_i))$ and assigns v_i a colouring according to this probability measure. The notion of "good" in the estimation of $\mu(X_{v_i}|C(A_i))$ implies that this estimation should be so accurate that it will be possible for the algorithm to colour the graph with distribution sufficiently close to Gibbs measure of $PCS(G, \mathcal{S})$.

Note that there is no known procedure that is able to compute exactly each of the measures $\mu(X_{v_i}|C(A_i))$, $i = 1, \ldots, n$, efficiently, i.e. in polynomial-time, for this class of graphs.

So as to provide efficiently "good" approximations of the measures $\mu(X_{v_i}|C(A_i))$, for $i = 1, \ldots, n$, the algorithm exploits ideas which are similar to those presented in [10] and [14], for counting satisfiable truth assignments in a random k-SAT formula and independent sets of general graphs, correspondingly. However, our proof techniques are novel and some technical results are of independent interest. For an earlier version of our result, see [5].

More specifically, our algorithm exploits two properties of the system $PCS(G_{n,d/n}, \mathcal{S})$ which hold w.h.p. The first one is that each vertex v of $G_{n,d/n}$ with all the vertices within graph distance $\epsilon \log n$ from v form an induced subgraph which is either unicyclic or tree, for a sufficiently small constant $\epsilon = \epsilon(d)$. The second one is that for a sufficiently large integer $\mathcal{S} = \mathcal{S}(d)$, the Gibbs measure that $PCS(G_{n,d/n}, \mathcal{S})$ specifies exhibits a specific *spatial mixing property*. In essence, we show that if $\mathcal{S}$ is greater than a specific value, which depends *only* on the expected degree d, then an *asymptotic independence* between the colour assignment of any vertex v and the colour assignment of any subset of vertices, which is at a sufficiently large (graph) distance from v, holds in the system.

Note that showing an asymptotic independence between the colour assignment of any vertex v and the colour assignment of any vertex set at a distance greater than $\epsilon \log n$, for sufficiently small $\epsilon = \epsilon(d)$, implies that when the i-th vertex is to be coloured the following holds: If the colouring $C(A_i)$ is done with probability measure which is sufficiently close to $\mu(X_{A_i} = C(A_i))$, then the algorithm can have a "good" estimation of $\mu(X_{v_i}|C(A_i))$ by just checking only the colourings of vertices that belong to a very simple structured neighborhood of v_i, i.e. a tree with at most one additional edge. The notion of a "good" estimation is the same as the one stated previously. This kind of structure in the neighborhood of the vertex v_i is highly desirable since then it allows us to get a colouring of v_i, which is distributed as this "good" estimation of $\mu(X_{v_i}|C(A_i))$, in time which is upper bounded by a polynomial of n.

The proof of validity of the spatial mixing property of the spin systems we consider here is of independent interest to that of the algorithm. Therefore, it is treated separately in Section 3.

1.2. Further definitions (spatial dependency)

For the graph $G = (V, E)$ and any two vertex sets $V', V'' \subset V$, we denote by $\text{dist}(V', V'')$ the graph distance of the two sets, i.e. the minimum length shortest path between all the pairs of vertices $(v_1, v_2) \in V' \times V''$.

Definition 1.4. Let $G = (V, E)$ be an instance of $G_{n,d/n}$ and let l be a positive real. For the vertex $v \in V$, let $G_{v,d,l}$ be the induced subgraph of G, which contains the vertex v and all the vertices within graph distance $\lfloor l \rfloor$ from v.

For a measure of comparison between probability measures we use the *total variation distance*.

Definition 1.5. For measures μ and ν on the same discrete space Ω, the total variation distance $d_{TV}(\mu, \nu)$ between μ and ν is defined as

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Definition 1.6 (*Spatial Dependency*). Consider the graph $G = (V, E)$, an instance of $G_{n,d/n}$, the positive integers $\mathcal{S}$, l and the positive real s. For each $v \in V$ consider the subgraph $G_{v,d,l}$ with vertex set $V_{v,l}$. For a given $v \in V$, consider also any two $\mathcal{S}$-colourings $C_1(V_1)$ and $C_2(V_1)$ with $V_1 \subset V_{v,l}$ and $\text{dist}(\{v\}, V_1) \geq l$, having $\Omega(G_{v,d,l}, \mathcal{S}, C_1(V_1)), \Omega(G_{v,d,l}, \mathcal{S}, C_2(V_1)) \neq \emptyset$. If for every $v \in V$ it holds

$$d_{TV}\left(\tilde{\mu}(X_v|C_1(V_1)), \tilde{\mu}(X_v|C_2(V_1))\right) \leq s$$

with $\tilde{\mu}(\cdot)$ specified by the system $PCS(G_{v,d,l}, \mathcal{S})$, then we say that "$\forall v \in V$ *the distance* l **Spatial Dependency** *of the* $\mathcal{S}$-*colourings of* G *is* s". This will be denoted as $\forall v \in V\ SD(v, l) = s$.

The above definition can be trivially extended to any system with any type of underlying graph.

It is easy for one to see that if "$\forall v \in V$ the distance l Spatial Dependency of the $\mathcal{S}$-colourings of $G = (V, E)$ is s", then in the system $PCS(G, \mathcal{S})$ and for each vertex $v \in V$ any two colourings $C'_1(V_1)$ and $C'_2(V_1)$ with V_1 as given in Definition 1.6 and C'_1 and C'_2 having $\Omega(G, \mathcal{S}, C'_1(V_1)) \neq \emptyset$ and $\Omega(G, \mathcal{S}, C'_2(V_1)) \neq \emptyset$ it holds:

$$d_{TV}\left(\mu(X_v|C_1(V_1)), \mu(X_v|C_2(V_1))\right) \leq s$$

with $\mu(\cdot)$ specified, now, by the system $PCS(G, \mathcal{S})$. The above holds since for each $v \in V$, if $\Omega_1 = \{C(V_1)|\Omega(G_{v,d,l}, \mathcal{S}, C(V_1) \neq \emptyset\}$ and $\Omega_2 = \{C(V_1)|\Omega(G, \mathcal{S}, C(V_1) \neq \emptyset\}$, then clearly $\Omega_2 \subseteq \Omega_1$.

1.3. Structure of the remaining paper

The remainder of our paper has the following structure. Section 2 is devoted to presenting a detailed description of the sampling algorithm and its properties. More specifically, in Section 2.1 we state the two basic properties of the systems $PCS(G_{n,d/n}, \mathcal{S})$ on which our algorithm is based. These two properties are stated, correspondingly, in Lemma 2.1 and in Theorem 2.2. In Section 2.1 we provide the proof of Lemma 2.1 which is relatively simple, while we postpone until Section 3 the proof of Theorem 2.2, which is more complex and it is somehow of independent interest of the actual algorithm. In Section 2.2 we give a detailed presentation of the sampling algorithm and of its properties regarding accuracy and efficiency without proof. The presentation is accompanied by a discussion on the proof techniques we use.

In Section 3 we provide an analytic discussion regarding the properties of the spin-system that our algorithm considers and we show Theorem 2.2. In Section 4 we present the proof of all the theorems and lemmas that are stated in Section 2.2.

For ease of verification of our proof, we provide here the dependence of each lemma and theorem on lemmas logically preceding it: Lemma 2.1, Lemma 2.3, Lemma 3.9 and Lemma 3.13, have no lemmas preceding them. Lemma 3.15 has predecessors Lemma 2.1, Lemma 3.9 and Lemma 3.13. Lemma 3.16 has predecessor Lemma 3.15. Lemma 3.18 has no predecessors. Lemma 3.20 has predecessors Lemma 3.13 and Lemma 3.9. Lemma 3.21 has predecessors, Lemma 3.18 and Lemma 3.20.

Theorem 2.2 has preceding all the lemmas of this paper except for Lemma 2.3. Theorem 2.4 has predecessors, Theorem 2.2, Lemma 2.3. Theorem 2.5, has predecessor Lemma 2.1. Finally, Theorem 3.5 has no predecessors.

2. Statement of results

2.1. Properties of the spin system

In this section we state two crucial properties that the colouring model has with high probability, when the underlying graph is an instance of a sparse random graph. These properties are stated in the following lemma and theorem. The first one refers to the structure of the neighborhood of each vertex in an instance of a sparse random graph. The second one refers to a property of the colourings (configurations) of such a spin-system. In essence, in a $PCS(G_{n,d/n}, \mathcal{S})$, if $\mathcal{S}$ is greater than a specific value, which depends *only* on the expected degree d, then an *asymptotic independence* between the colour assignment of any vertex v and the colour assignment of any subset of vertices which is at a graph distance, at least, $\left\lfloor \frac{0.9}{4\log(e^2 d/2)} \log n \right\rfloor$ from v holds.

Lemma 2.1. *Let $G = (V, E)$ be an instance of $G_{n,d/n}$, where $d \geq 1$ is a fixed positive real. With high probability (w.h.p.) the graph has* no vertex *v with the following property: The induced subgraph of G that contains v and all vertices within a distance $\epsilon \log n$ from v, contains more than one cycle, for any real $\epsilon > 0$ such that $\epsilon \leq (4\log(e^2 d/2))^{-1}$.*

Proof. To prove the lemma assume the contrary, i.e. there is some vertex $v \in V$ whose corresponding graph $G_{v,d,\epsilon \log n}$ (see Definition 1.4) contains two cycles, i.e. C_1 and C_2 each of length at most $2\epsilon \log n$, the value of ϵ will be determined later. The above assumption implies that there are two pairs of paths starting from v, such that: The paths in each pair do not have all their edges common and there is some vertex in $G_{v,d,\epsilon \log n}$ that can be reached from v by both paths of the pair. The existence of such two pairs of paths implies that in $G_{v,d,\epsilon \log n}$ there is a set of, at most $4\epsilon \log n$, vertices which have among each other a number of edges which exceeds the number of vertices by one.

Thus, the proof of the lemma reduces to showing that in $G_{n,d/n}$, there is no set of, at most, $4\epsilon \log n$ vertices which contains a number of edges that exceed the number of vertices in the set by one, for sufficiently small ϵ. Let D be the event "such a set exists". Setting $r = 4\epsilon \log n$ we have

$$\begin{aligned}
\Pr[D] &\leq \sum_{k=1}^{r} \binom{n}{k} \binom{\binom{k}{2}}{k+1} \left(\frac{d}{n}\right)^{(k+1)} \\
&\leq \sum_{k=1}^{r} \left(\frac{ne}{k}\right)^{k} \left(\frac{ek(k+1)}{2(k+1)}\right)^{(k+1)} \left(\frac{d}{n}\right)^{(k+1)} \\
&\leq \frac{de}{2n} \sum_{k=1}^{r} \left(\frac{e^2}{2}\right)^{k} kd^k \\
&\leq \frac{d^2e^3}{4} \frac{4\epsilon \log n}{n} \sum_{k=0}^{r-1} \left(\frac{de^2}{2}\right)^{k} \\
&\leq \frac{d^2e^3\epsilon \log n}{n} \frac{(de^2/2)^r - 1}{de^2/2 - 1}
\end{aligned}$$

taking ϵ such that $4\epsilon \cdot \log(de^2/2) < 1$ the r.h.s. of the last equation is $o(1)$, as the nominator is $o(n)$. Thus, for sufficiently small ϵ it holds that $\Pr[D] = o(1)$ $\diamond$

Theorem 2.2. *Let $G = (V, E)$ be an instance of $G_{n,d/n}$, where $d > 1$. If $\mathcal{S}$ is a sufficiently large integer, which depends on d, and $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$, then w.h.p., i.e. with probability $1 - O(n^{-0.1})$, for every vertex $v \in V$ the distance $\lfloor \epsilon \log n \rfloor$ Spatial Dependency of the $\mathcal{S}$-colourings of $PCS(G, \mathcal{S})$ is $n^{-1.25}$. For sufficiently large d, we should have $\mathcal{S} \geq d^{14}$.*

It should be mentioned that, if d is relatively small, then with the same probability for $PCS(G_{n,d/n}, \mathcal{S})$ the distance $\lfloor \epsilon \log n \rfloor$ Spatial Dependency of the $\mathcal{S}$-colourings of $PCS(G, \mathcal{S})$ can become $n^{-1.25}$, however, for a number of colours which is a constant greater than d^{14}.

Section 3 is devoted to the proof of Theorem 2.2.

2.2. Our Algorithm

The sampling algorithm we propose here takes as an input the graph $G = (V, E)$, an instance of $G_{n,d/n}$ with $d > 1$, and the positive integer $\mathcal{S}$. Assume that the input graph G has the following three properties: *First*, for each $v \in V$ the subgraph $G_{v,d,\epsilon \log n+2}$, for $\epsilon = \frac{0.9}{4\log(de^2/2)}$, is either unicyclic or tree. *Second*, G is such that the Gibbs measure $\mu(\cdot)$, which $PCS(G, \mathcal{S})$ specifies, exhibits the spatial mixing property that is indicated in the statement of Theorem 2.2. *Third*, G can be coloured with $\mathcal{S}$-3 colours. Under these assumptions and by taking a sufficiently large constant $\mathcal{S}$, the algorithm outputs a $\mathcal{S}$-colouring of G with a probability measure that is within total variation distance $n^{-0.25}$ from the uniform over the set of all $\mathcal{S}$-colourings of the input graph, i.e. the Gibbs measure $\mu(\cdot)$.

By Lemma 2.1 and Theorem 2.2, we know that the first two of the above assumptions about the input graph hold with high probability for an instance of $G_{n,d/n}$. Also, for a number of colours $\mathcal{S}$ as large as is indicated by Theorem 2.2, an instance of $G_{n,d/n}$ is colourable with $\mathcal{S}$-3 colours w.h.p. (see [1,6]). It should be mentioned that our algorithm is based on properties of the spin-system which a priori hold w.h.p. This implies that we can expose the entire instance of the input graph at the beginning and expect these desired properties to hold (which is highly probable).

We continue by giving a description of the sampling algorithm in the form of pseudo code. In what follows, we assume that v_i is the i-th vertex to be coloured by the algorithm and A_i is the set of vertices that have already been coloured before v_i. We also assume that A_i is coloured as $C(A_i)$. We denote with (V_i, E_i) the vertex set and the edge set, correspondingly, of the graph $G_{v_i,d,\epsilon \log n}$, where $\epsilon = \frac{0.9}{4\log(de^2/2)}$. Let $\tilde{\mu}_i(\cdot)$ denote the Gibbs measure specified by the system $PCS(G_{v_i,d,\epsilon \log n}, \mathcal{S})$.

Sampling Algorithm
Input: $G = (V, E)$, instance of $G_{n,d/n}$, number of colours $\mathcal{S}$
Take an arbitrary permutation of the vertices in V, i.e. $(v_1, \ldots, v_n)$
$A_1 = \emptyset$
For $i = 1, \ldots, n$
-Create the subgraph $G_{v_i,d,\epsilon \log n} = (V_i, E_i)$
-**If** $G_{v_i,d,\epsilon \log n} = (V_i, E_i)$ is not a tree or a unicyclic graph
Then Return Failure
-Colour v_i according to $\tilde{\mu}_i(X_{v_i} | C(A_i \cap V_i))$
using dynamic programming
-$A_{i+1} := A_i \cup \{v_i\}$
Return Colouring of G

If the algorithm, at each iteration of the for loop, could assign to the vertex v_i a colouring, according to $\mu(X_{v_i} | C(A_i))$, instead of $\tilde{\mu}_i(X_{v_i} | C(A_i \cap V_i))$, then it would be exact. However, there is no known polynomial-time procedure that is able to compute the exact measure $\mu(X_{v_i} | C(A_i))$ for this class of spin systems. In essence, we use the measure $\tilde{\mu}_i(X_{v_i} | C(A_i \cap V_i))$ as an estimation of $\mu(X_{v_i} | C(A_i))$ in our algorithm.

In what follows, we discuss in detail two major, open issues about the algorithm above. The first one is its **accuracy**, i.e. how close the probability measure of the colouring that is returned, is to the uniform over all proper $\mathcal{S}$-colourings of the input graph. The second one is its **efficiency**, i.e. how much time is needed for the execution of the algorithm with respect to the size of the input graph.

As far as the accuracy of the algorithm is concerned, we see that for a *given* colour assignment $C(A_i)$, the algorithm approximates the measure $\mu(X_{v_i} | C(A_i))$ with the following convex combination of measures:

$$\sum_j a_j \cdot \tilde{\mu}_i(X_{v_i} | C_j(B_i))$$

where $B_i = (A_i \cap V_i) \cup V'$ and $V' \subseteq V$ contain all the vertices at a distance $\lfloor \epsilon \log n \rfloor$ from v_i. Also, for any j, $C_j(B_i)$ is a colouring which agrees with $C(A_i \cap V_i)$ on the colour assignment of the vertices in $A_i \cap V_i$ and the measure $\tilde{\mu}_i()$ assigns it a positive probability a_j.

Furthermore, we note that $\tilde{\mu}_i(X_{v_i} | C_j(B_i)) = \mu(X_{v_i} | C_j(A_i \cup V'))$ as long as the colouring $C_j(B_i)$ is such that $\Omega(G, \mathcal{S}, C_j(B_i)) \neq \emptyset$. However, it is possible that for some j it holds that $\Omega(G, \mathcal{S}, C_j(B_i)) = \emptyset$. If we consider without proof, for the moment, that for all j, $C_j(B_i)$ is such that $\Omega(G, \mathcal{S}, C_j(B_i)) \neq \emptyset$, we can rewrite the above summation as follows:

$$\sum_j a_j \cdot \mu(X_{v_i}|C_j(A_i \cup V')).$$

Since the above combination of measures is convex, the total variation between the estimation of $\mu(X_{v_i}|X_{A_i} = C(A_i))$ and the actual measure is upper bounded by the quantity

$$\max_j \left\{ d_{TV}(\mu(X_{v_i}|C(A_i)), \mu(X_{v_i}|C_j(B_i))) \right\}.$$

To bound the above total variation distance we use Theorem 2.2. More specifically, Theorem 2.2 states that by taking a sufficiently large $\mathcal{S}$, w.h.p. the system that our algorithm considers specifies a Gibbs measure, $\mu(\cdot)$, which exhibits the following property:

$$d_{TV}(\mu(X_v|C_1(V'')), \mu(X_v|C_2(V''))) \leq n^{-1.25} \tag{1}$$

where $V'' \subset V$ is such that $\mathrm{dist}(\{v\}, V'') \geq \lfloor \epsilon \log n \rfloor$ and $C_1(V'')$, $C_2(V'')$ are any colourings such that $\omega(G, \mathcal{S}, C_1(V''))$, $\omega(G, S, C_2(V'')) \neq \emptyset$.

However, for our algorithm we need to get a little further than what is stated in the above equation. When the vertex v_i is to be coloured, it is possible that many vertices at a small distance from v_i have, already, been coloured. Generally, fixing the colour assignments of the vertices in A_i possibly imposes a probability measure for the $\mathcal{S}$-colourings of G which is different from the Gibbs measure $\mu(\cdot)$. Our sampling algorithm, for a given input G and $\mathcal{S}$, specifies a probability measure for the colour assignments of the vertices in A_i and in turn for the $\mathcal{S}$-colourings of G. In order for the algorithm to be accurate, it is essential that at the i-th iteration of the for loop the following relation holds:

$$d_{TV}(\mu_{A_i}(X_{v_i}|C_1(V_i'')), \mu_{A_i}(X_{v_i}|C_2(V_i''))) \leq n^{-1.25}. \tag{2}$$

The measure $\mu_{A_i}(\cdot)$ is the Gibbs measure of the system $PCS(G, \mathcal{S})$ with the vertices in A_i assigned a colouring with the probability measure that is specified by the algorithm.[4] The vertex set $V_i'' \subset V$ contains all vertices at a distance $\lfloor \epsilon \log n \rfloor$ from v_i.

Note that by the law of total probability we can deduce the following.

Remark. If the algorithm assigns the vertices in A_i, for $i = 1, \ldots, n$, a colouring $C(A_i) \in \mathcal{S}^{A_i}$ according to the probability measure $\mu(X_{A_i} = C(A_i))$, then the probability measure of the any configuration of the system $PCS(G, \mathcal{S})$, with the vertices in A_i coloured as above, will be given by $\mu(\cdot)$.

The above remark implies that the statistical properties of the system $PCS(G, \mathcal{S})$ do not change as long as we fix the colour assignments of any vertex set according to the Gibbs measure that this system specifies. This consequently implies that (2) is valid as long as at the i-th iteration of the for loop of the algorithm the vertices in A_i are assigned a colouring according to $\mu(\cdot)$.

Finally, when we argue about the colourings that the measures $\tilde{\mu}_i()$, for $i = 1, \ldots, n$, give positive probability, we use the following lemma:

Lemma 2.3. *Assume that in the execution of the algorithm each of the graphs $G_{v_i,d,\epsilon \log n+2}$, for $i = 1, \ldots, n$, is either unicyclic or tree and $\mathcal{S}$ is as large as specified by Theorem 2.2. With probability, at least, $1-n^{-0.1}$ for all the measures $\tilde{\mu}_i(\cdot)$, where $i = 1, \ldots, n$, the following is valid: For $V_i' \subset V$ such that $\mathrm{dist}(\{v_i\}, V_i') \leq \lfloor \epsilon \log n \rfloor$ and for any colouring $C(V_i')$ for which $\tilde{\mu}_i(X_{V_i'} = C(V_i')) \neq 0$ it holds that $\Omega(G, \mathcal{S}, C(A_i \cup V_i')) \neq \emptyset$, where $i = 1, \ldots, n$.*

The proof of Lemma 2.3 is provided in Section 4. Taking all the above facts into consideration, Theorem 2.4 gives a characterization of the probability measure of the colouring that is returned by the algorithm in terms of its total variation distance from the Gibbs measure $\mu(\cdot)$ that specifies the system $PCS(G, \mathcal{S})$, i.e. the uniform over all the proper $\mathcal{S}$-colourings of the input graph G.

Theorem 2.4. *If $\mathcal{S}$ is a sufficiently large integer constant, then, with probability $1-O(n^{-0.1})$, the sampling algorithm is successful and returns a $\mathcal{S}$-colouring of the input graph G, whose distribution is within total variation distance $n^{-0.25}$ from the uniform over all the proper $\mathcal{S}$-colourings of G.*

The proof of Theorem 2.4 is given in Section 4.

[4] We avoided writing the measures in the above relation by conditioning on the colour assignment of A_i because we want to stress out that the colouring of the vertices in A_i is done according to some probability measure.

As far as the execution time of the algorithm is concerned, we make the following remark. According to Lemma 2.1, the set $\{G_{v_i,d,\epsilon \log n}$, for $i = 1, \ldots, n\}$, for $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$, w.h.p., i.e. with probability $1 - n^{-0.1}$, contains graphs which are either unicyclic or trees. If this is not the case, then we consider that the algorithm fails. As argued in [4], we can have a colouring of the vertex v_i according to $\tilde{\mu}_i(X_{v_i}|C(A_i \cap V_i))$ by generating a random colouring of $G_{v_i,d,\epsilon \log n}$ where the vertices in $A_i \cap V_i$ are colored as $C(A_i \cap V_i)$ in time upper bounded by $l \cdot k^3$, where $l = |V_i|$ and $k = \mathcal{S}$ (for more details see the proof of Theorem 2.5 and [4]).

Theorem 2.5. *The time complexity of the sampling algorithm is w.h.p. asymptotically upper bounded by $O(n^2)$, where n is the number of vertices of the input graph.*

The proof of Theorem 2.5 is given in Section 4.

We note that at the i-th iteration of the for-loop of the algorithm, we can apply the Junction tree algorithm (see [15]) to compute the probability measure $\tilde{\mu}_i(X_u|C(A_i \cap V_i))$. The execution time of the junction tree is asymptotically bounded by $O(n^{2+c})$, where $c < 1$ is a sufficiently large constant.

3. Spatial mixing

This section is devoted to showing Theorem 2.2. For the graph $G = (V, E)$, an instance of $G_{n,d/n}$, we consider the system $PCS(G, \mathcal{S})$, for some positive integer $\mathcal{S}$. Theorem 2.2 states that in $PCS(G, \mathcal{S})$, for every $v \in V$ the distance $\lfloor \epsilon \log n \rfloor$ Spatial Dependency is upper bounded by $n^{-1.25}$ if $\mathcal{S}$ is a sufficiently large constant and $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$.

By definition, for each $v \in V$ the $\lfloor \epsilon \log n \rfloor$ Spatial Dependency in the system $PCS(G, \mathcal{S})$ will be $n^{-1.25}$ with probability equal to the probability of the union of the following set of events: For each $v \in V$, in the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$ for the vertex v the $\lfloor \epsilon \log n \rfloor$ Spatial Dependency is $n^{-1.25}$.

Observe that Theorem 2.2 defines such small ϵ that with probability, at least, $1 - n^{-0.1}$ the set $\{G_{v,d,\epsilon \log n} | v \in V\}$ contains graphs for which each of them is either unicyclic or tree. With this observation, instead of Theorem 2.2 we show the two following lemmas.

Lemma 3.16. *Consider the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$, for $d > 1$, $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$ and for $G_{v,d,\epsilon \log n}$ we condition that it is a tree. If $\mathcal{S}$ is a sufficiently large constant, then with probability, at least, $1 - 2n^{-1.25}$ for the above system it holds that $SD(v, \lfloor \epsilon \log n \rfloor) = n^{-1.25}$. For sufficiently large d, we should have $\mathcal{S} \geq d^{14}$.*

Lemma 3.21. *Consider the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$, for $d > 1$, $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$ and for $G_{v,d,\epsilon \log n}$ we condition that it is a unicyclic graph. If $\mathcal{S}$ is a sufficiently large constant, then with probability, at least, $1 - 2n^{-1.25}$ for the above system it holds that $SD(v, \lfloor \epsilon \log n \rfloor) = n^{-1.25}$. For sufficiently large d, we should have $\mathcal{S} \geq d^{14}$.*

Given the validity of Lemma 3.16 and Lemma 3.21, Theorem 2.2 follows directly, see Corollary 3.1.

Corollary 3.1. *If Lemma 3.16 and Lemma 3.21 are true, then Theorem 2.2 is true, as well.*

Proof. Theorem 2.2 holds for $PCS(G = (V, E), \mathcal{S})$, where G is an instance of $G_{n,d/n}$, if the following event holds with probability at least $1 - O(n^{-0.1})$, for appropriately large $\mathcal{S}$.

$$Event_1 = \text{“for every graph } G_{v,d,\epsilon \log n} \text{ of the set of graphs that } G \text{ specifies, it holds that the } PCS(G_{v,d,\epsilon \log n}, \mathcal{S}) \text{ has the property that } SD(v, \lfloor \epsilon \log n \rfloor) = n^{-1.25}\text{”.}$$

Assume, first, that for each $v \in V$ the subgraph $G_{v,d,\epsilon \log n}$, for $\epsilon = \frac{0.9}{4\log(de^2/2)}$, is either unicyclic or tree. By Lemma 3.16 and Lemma 3.21 for each vertex v in G it holds that for sufficiently large $\mathcal{S}$, which for sufficiently large d becomes $\mathcal{S} \geq d^{14}$, the event

$$Event_v = \text{“ the system } PCS(G_{v,d,\epsilon \log n}, \mathcal{S}) \text{ has the property that } SD(v, \lfloor \epsilon \log n \rfloor) = n^{-1.25}\text{”}$$

holds with probability, at least, $1 - 2n^{-1.25}$. Clearly,

$$\Pr[Event_1] = 1 - \Pr\left[\cup_v \overline{Event_v}\right].$$

By the union bound we get that $\Pr[Event_1] \geq 1 - 2n^{-0.25}$.

By Lemma 2.1, the assumption we have made about the structure of neighborhood of each vertex $v \in V$ is valid with probability at least $1 - n^{-0.1}$. Clearly, for $PCS(G, \mathcal{S})$ the event E_1 holds with probability at least $(1 - n^{-0.1})(1 - 2n^{n^{-0.25}}) = 1 - O(n^{-0.1})$. ◇

The proof of Lemma 3.16 and Lemma 3.21 are based on comparing, in terms of total variation distance, the two measures $\tilde{\mu}(X_v|C(L))$ and $\tilde{\mu}(X_v|C'(L))$ as these are specified by the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$. The vertex set L is assumed to contain all the vertices in $G_{v,d,\epsilon \log n}$ that are at a distance $\lfloor \epsilon \log n \rfloor$ from v and the boundary conditions $C(L)$ and $C'(L)$ are taken so as to maximize the total variation distance of the two measures, while $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L)), \Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C'(L)) \neq \emptyset$.

More specifically, for the case of Lemma 3.16 where the underlying graph of the spin-system is a tree, we introduce a stochastic process which we call ColourRoot. For any tree T rooted at vertex r and some integer k, the process ColourRoot when applied to T and when it uses k colours, it colours in a recursive manner the tree T so that the vertex r is coloured according to the Gibbs measure $\hat{\mu}(X_v|C(V'))$, as this is specified by the system $PCS(T, k, C(V'))$. We consider an appropriate *coupling* for two executions of the ColourRoot, that are applied to $G_{v,d,\epsilon \log n}$ and use $\mathcal{S}$ colours, which assign to v colourings according to $\tilde{\mu}(X_v|X_L = C(L))$ and $\tilde{\mu}(X_v|X_L = C'(L))$, correspondingly. The lemma follows by bounding appropriately the probability for the two processes, in the coupling, to assign a different colour to the vertex v (see [2] for bounding total variation distances by using coupling). For a detailed presentation of the process ColourRoot and the coupling we use, see Section 3.1. For the proof of Lemma 3.16, see Section 3.2.

To prove Lemma 3.21, we provide a lemma which reduces the problem of bounding the total variation distance of the measures $\tilde{\mu}(X_v|C(L))$ and $\tilde{\mu}(X_v|C'(L))$ to bounding appropriately the total variation distance between Gibbs measures defined on systems whose underlying graph is a tree. This tree is constructed subject to the unicyclic graph $G_{v,d,\epsilon \log n}$. Then, Lemma 3.21 follows by using essentially the same approach as we do for proving Lemma 3.16. For a detailed discussion on the reduction and the proof of Lemma 3.21, see Section 3.3.

Remark. Both Lemma 3.16 and Lemma 3.21 are based on the fact that we expect a very large proportion of the vertices of an instance of $G_{n,d/n}$ to have constant degrees. In essence, there is a constant $c_0 = c_0(d)$ such that for any $c > c_0$ the expected proportion of vertices that have a degree less than c tends to 1, exponentially fast with c. This argument is justified by the following corollary, which is proved in [6].

Corollary 3.2 (*Janson et. al. [6], pp 28.*). *If a random variable Z is distributed as in $B(n, q)$, the binomial distribution with parameters n and q, with $\lambda = nq$ then*

$$\Pr[Z \geq x] \leq e^{-x} \quad x \geq 7\lambda.$$

3.1. The process ColourRoot and a coupling

Towards proving Lemma 3.16 and Lemma 3.21 we introduce the stochastic process *ColourRoot* $(T, \mathcal{S}, C(L))$, where $T = (V, E)$ is a tree, $\mathcal{S}$ is a positive integer and the vertices in $L \subset V$ are assigned a colouring $C(L)$ such that $\Omega(T, \mathcal{S}, C(L)) \neq \emptyset$. The process *ColourRoot*$(T, S, C(L))$ assigns a colouring (not necessarily proper) to the vertices in $V \backslash L$ such that $\forall u \in V \backslash L$ the probability measure of its colour assignment is equal to $\mu(X_u|C(L \cap T_u))$, where T_u is the subtree of T rooted at u, while the Gibbs measure $\mu(X_v|C(L \cap T_u))$ is specified by the system $PCS(T_u, \mathcal{S}, C(L \cap T_u))$. When the third parameter of the ColourRoot is omitted, it is implied that there is no fixed colour assignment to any vertex.

The ColourRoot$(T, \mathcal{S}, C(L))$ assigns a colouring to each vertex u in the tree T based on the following observation. For the vertex u in the tree T, consider the vertex set CH_u, which contains the children of u in T_u, and the system $PCS(T_u, \mathcal{S}, C(L \cap T_u))$. For the graph $T_0 = \cup_{w \in CH_u} T_w$ consider the set of $\mathcal{S}$-colorings $\Omega_0 = \Omega(T_0, \mathcal{S}, C(L \cap T_0))$. Assume that each $C \in \Omega_0$ specifies a colouring of the vertices in CH_u that uses all but W_C colours from the set $[\mathcal{S}]$. Note that if the system $PCS(T_u, \mathcal{S}, C(L \cap T_u))$ is in equilibrium, the probability for the vertices in $T_0 = \cup_{w \in CH_u} T_w$ to be coloured as specified by $C \in \Omega_0$ is proportional to the quantity W_C, i.e. $\frac{W_C}{\sum_{C \in \Omega_0} W_C}$.

Definition 3.3. With the above notation, the process ColourRoot$(T, \mathcal{S}, C(L))$ assigns a colour to the vertex u of T as follows:

1. Each $C \in \Omega_0$ is assigned weight W_C which is equal to the number of colours in the set $[\mathcal{S}]$ that do not appear in the colour assignment that C specifies for the vertices in CH_u.
2. Select from Ω_0 so that the probability for each member to be chosen is proportional to the weight it has been assigned to it. Let C' be the chosen member.
3. Assign to the vertex u a colour that is chosen uniformly at random among the colours in the set $[\mathcal{S}]$ that do not appear in the colouring of the vertices in CH_u, as this is specified by C'.

We, also, introduce the notion of *disagreement probability* for a *coupling* of the processes ColourRoot$(T, \mathcal{S}, C(L))$ and ColourRoot $(T, \mathcal{S}, C'(L))$.

Definition 3.4. Consider a coupling of ColourRoot $(T, \mathcal{S}, C(L))$ and ColourRoot$(T, \mathcal{S}, C'(L))$. The disagreement probability for a vertex u in T, denoted by p_u, is equal to the probability of the event that the two processes in the coupling assign different colours to u.

The coupling of ColourRoot $(T, \mathcal{S}, C(L))$ and ColourRoot$(T, \mathcal{S}, C'(L))$ is of our main interests due to the following, very significant, fact.

Theorem 3.5. *Consider the tree $T = (V, E)$ rooted at the vertex r, the integer $\mathcal{S}$, some set $A \subseteq V$ and any two $\mathcal{S}$-colourings $C(A)$ and $C'(A)$ such that $\Omega(T, \mathcal{S}, C(A)), \Omega(T, \mathcal{S}, C'(A)) \neq \emptyset$. Assume that there is a coupling of the ColourRoot$(T, \mathcal{S}, C(A))$ and the ColourRoot$(T, \mathcal{S}, C'(A))$ for which the probability of disagreement for the root r is equal to $\mathrm{p_r}$. Then, it holds that*

$$d_{TV}(\mu(X_r|C(A)), \mu(X_r|C'(A))) \leq \mathrm{p_r}.$$

where $\mu(X_r|C(A))$ and $\mu(X_r|C'(A))$ are specified by the system $PCS(T, \mathcal{S})$.

Proof. The theorem follows directly from the Coupling Lemma (see [2]). ◇

For the system $PCS(T, \mathcal{S})$, where T is a tree rooted at vertex r, one can derive upper bounds for $SD(r, l)$, for some positive integer l, by using the above theorem and the coupling of the ColourRoot which is provided in the following definition.

Definition 3.6. Consider the tree $T = (V, E)$ rooted at vertex r, an integer $\mathcal{S}$ and the set $V_1 \subset V$ such that $\mathrm{dist}(\{r\}, V_1) \geq l$, for some integer l. Let $\mathcal{C}(T, \mathcal{S}, l)$ be the coupling of the processes ColourRoot$(T, \mathcal{S}, C_1(V_1))$ and ColourRoot$(T, \mathcal{S})$ such that the colour assignment $C_1(V_1)$, is taken so as to *maximize* the disagreement probability at the root of T. The coupling $\mathcal{C}(T, \mathcal{S}, l)$ assigns colours to the vertex u of T as follows:

- Couple step 2 of the two processes so as to maximize the probability for the set CH_u to have the same colour assignment in the two processes.
- Conditional on the choices that the two processes have made at their step 2, assign colours to u so as to minimize the disagreement probability p_u.

In the coupling $\mathcal{C}(T, \mathcal{S}, l)$, if the height of T is less than l, then the set V_1, which is given in Definition 3.6, is empty. It is easy for one to see that in that case the disagreement probability for all vertices in T is zero.

Corollary 3.7. *Consider a tree T, rooted at vertex r. If the coupling $\mathcal{C}(T, \mathcal{S}, l)$, for some positive integers $\mathcal{S}$ and l, has disagreement probability $\mathrm{p_r}$ on the root of T, then for $PCS(T, \mathcal{S})$ it holds that $SD(r, l) \leq 2\mathrm{p_r}$.*

Proof. For the tree T, rooted at vertex r, and the integers $\mathcal{S}$, l, assume that in the coupling $\mathcal{C}(T, \mathcal{S}, l)$ the disagreement probability on the root r is $\mathrm{p_r}$. Consider the vertex set L which contains vertices at a distance, at least, l from the root r. Let also $\tilde{C}(L)$ and $\hat{C}(L)$ be the two colourings which maximize the total variation distance of the measures $\mu(X_r|\tilde{C}(L))$ and $\mu(X_r|\hat{C}(L))$, as these are specified by the system $PCS(T, \mathcal{S})$. It holds that

$$\begin{aligned} SD(r, l) &= d_{TV}\left(\mu(X_r|\tilde{C}(L)), \mu(X_r|\hat{C}(L))\right) \\ &\leq d_{TV}\left(\mu(X_r|\tilde{C}(L)), \mu(X_r)\right) + d_{TV}\left(\mu(X_r), \mu(X_r|\hat{C}(L))\right) \\ &\leq 2\mathrm{p_r}. \end{aligned}$$

The second derivation follows by the triangle inequality for measures. The corollary follows. ◇

It should be mentioned that we will not need to give an explicit description of the coupling $\mathcal{C}(T, \mathcal{S}, l)$. It will suffice to show that $\mathcal{C}(T,\mathcal{S},l)$ has two specific properties, i.e. those indicated by Lemmas 3.9 and 3.13.[5]

The remainder of this section contains the statement and the proof of Lemmas 3.9 and 3.13. These two lemmas provide means to derive upper bounds for the probability of disagreement for each vertex u of T, in $\mathcal{C}(T, \mathcal{S}, l)$, by providing an inductive description of the coupling in terms of disagreement probabilities. More specifically, consider some vertex u, in the tree T, and the set CH_u of its children. If the coupling $\mathcal{C}(T, \mathcal{S}, l)$ assigns colours to each vertex $w \in CH_u$ such that the probability of disagreement is equal to p_w, then for the vertex u the probability of disagreement p_u is upper bounded as follows:

$$p_u \leq a(|CH_u|, \mathcal{S}) \cdot \left(\sum_{w \in CH_u} p_w\right)$$

where $a(|CH_u|, \mathcal{S})$ is a quantity of size that depends on the cardinality of CH_u and $\mathcal{S}$.

We distinguish two classes of vertices in T regarding the relation between their number of children and the number of available colours $\mathcal{S}$, i.e. the *mixing* vertices and the *nonmixing* vertices. The *mixing* vertices have a number of children which is smaller than $\mathcal{S}$ and the constant $a(|CH_u|, \mathcal{S})$ is very small, i.e. $<< 1$. The *nonmixing* vertices have high degrees and the constant $a(|CH_u|, \mathcal{S})$ may become very large.

Definition 3.8. Each vertex u of the tree T is "mixing" if, for a given t, the number of its children in T is at most t, otherwise it is "nonmixing".

---

[5] However, if the reader is keen on finding one, then he can deduce it from the proofs of Lemmas 3.9 and 3.13 and the proofs of the claims inside them.

The value of t, the maximum number of children of a mixing vertex in the coupling $\mathcal{C}(T, \mathcal{S}, l)$, is *always less than the number of available colours*. Generally, for a given tree T and a number of colours $\mathcal{S}$, we take t so large as to minimize the probability of disagreement of the root of T.

Lemma 3.9. *Consider the tree T, the integers $\mathcal{S}$, l and the coupling $\mathcal{C}(T, \mathcal{S}, l)$. Let u be a vertex of T which is* mixing. *If for every vertex $w \in CH_u$ the probability of disagreement is p_w, then, for the vertex u, the probability of disagreement p_u is bounded as*

$$p_u \leq \frac{t \cdot \mathcal{S}}{(\mathcal{S} - t)^2} \cdot \left(\sum_{w \in CH_u} p_w \right)$$

where t is the maximum number of children of a mixing vertex.

Proof. When the vertex u is to be coloured in $\mathcal{C}(T, \mathcal{S}, l)$, assume that the processes ColourRoot(T, $\mathcal{S}$, $C_1(V_1)$) and ColourRoot(T, $\mathcal{S}$) during step 2 choose from the set of colourings Ω_C and Ω_F, correspondingly. Let $\mathcal{A}$ be the event that ColourRoot($T, \mathcal{S}, C_1(V_1)$) and ColourRoot($T, \mathcal{S}$) choose colourings from Ω_C and Ω_F, correspondingly, that specify two different colour assignments for the set CH_u. It holds that

$$\Pr[\text{disagreement on } u] = \Pr[\text{disagreement on } u|\mathcal{A}]\Pr[\mathcal{A}] + \Pr[\text{disagreement on } u|\overline{\mathcal{A}}]\Pr[\overline{\mathcal{A}}].$$

It should be mentioned that if the event $\mathcal{A}$ does not hold ($\overline{\mathcal{A}}$ holds), then there is a coupling for step (3) of the two processes ColourRoot that assigns the same colour to the vertex u, i.e. $\Pr[\text{disagreement on } u|\overline{\mathcal{A}}] = 0$. Thus,

$$\Pr[\text{disagreement on } u] = \Pr[\text{disagreement on } u|\mathcal{A}]\Pr[\mathcal{A}]. \tag{3}$$

We will provide appropriate upper bounds for the probability terms in (3) so as to prove the lemma, i.e. the terms $\Pr[\mathcal{A}]$ and $\Pr[\text{disagreement on } u|\mathcal{A}]$. We start with the probability term $\Pr[\mathcal{A}]$. The assumption that for each $w \in CH_u$ the disagreement probability in $\mathcal{C}$ is p_w is equivalent to the following: There is a coupling, call it $\mathcal{K}_1$, which chooses uniformly at random (u.a.r.) from the sets Ω_F and Ω_C and the two chosen elements specify different colour assignments for the vertex set CH_u with probability which is upper bounded by $\sum_{w \in CH_u} p_w$.

Noting that $|\Omega_F| \neq |\Omega_C|$, we create the set Ω'_F such that each element of Ω_F appears $|\Omega_C|$ times in Ω'_F. Similarly, we create the set Ω'_C such that each element of Ω_C appears $|\Omega_F|$ times in Ω'_C. It holds $|\Omega'_C| = |\Omega'_F|$.

Claim 3.10. *We can have a coupling, call it $\mathcal{K}_2$, that chooses uniformly at random (u.a.r.) an element from each of the sets Ω'_C and Ω'_F such that the probability for the two chosen elements to specify different colour assignments for the vertex set CH_u is upper bounded by $\sum_{w \in CH_u} p_w$.*

The proof of Claim 3.10 is given after this proof.

Assume that in $\mathcal{C}$ each of the executions of ColourRoot at step (2), now, considers the sets Ω'_C and Ω'_F, correspondingly, instead of Ω_C and Ω_F. Clearly, the fact that the processes ColourRoot consider the sets Ω'_C and Ω'_F, does not change the marginal probability measure of the colour assignment of the vertex u in each processes.

Claim 3.11. *Consider the coupling $\mathcal{C}(T, \mathcal{S}, l)$ and assume that the number of children of the vertex u is k and the disagreement probability for $w \in CH_u$ is p_w. If at the coupling of step (2) of the processes ColourRoot each $C \in \Omega'_C \cup \Omega'_F$ is assigned weight W_C, then for the event $\mathcal{A}$ it holds that*

$$\Pr[\mathcal{A}] \leq \frac{1}{q_{k,\mathcal{S}}} \frac{\max\limits_{C \in \Omega'_F \cup \Omega'_C} \{W_C\}}{\min\limits_{C \in \Omega'_F \cup \Omega'_C} \{W_C | W_C > 0\}} \sum_{w \in CH_u} p_w. \tag{4}$$

Where $q_{k,\mathcal{S}}$ is the probability of the event that after k trials, not all elements of $[\mathcal{S}]$ have been chosen, when at each trial we choose u.a.r. a member of $[\mathcal{S}]$.

The proof of Claim 3.11 is given after the proof of Claim 3.10.

Since the number of children of a mixing vertex is less than the number of available colours, there is no colouring of the vertices in CH_u that leave no available colour for u. This implies that in our case $q_{k,\mathcal{S}} = 1$. Note, at the coupling of step (2) of the ColourRoot no member of either Ω'_C or Ω'_F is assigned weight which is more than $\mathcal{S}$ and less than $\mathcal{S} - t$, where t is equal to the maximum number of children that a mixing vertex can have. Thus, we conclude

$$\Pr[\mathcal{A}] \leq \frac{\mathcal{S}}{\mathcal{S} - t} \sum_{w \in CH_u} p_w.$$

We proceed to derive a bound for $\Pr[\text{disagreement on } u|\mathcal{A}]$. For this, we use the following claim.

Claim 3.12. *Consider the coupling $\mathcal{C}(T, \mathcal{S}, l)$, when it assigns colourings to the vertex u. The coupled processes chose a member of Ω'_F and Ω'_C, correspondingly, and we assume that each member specifies a list of available colours for the vertex u, i.e. l_1 and l_2. Assuming that $|l_i| > 0$, for $i = 1, 2$, there is a coupling that can choose the same colour for the vertex u with probability at least*

$$1 - \frac{\max\{|l_1 \backslash l_2|, |l_2 \backslash l_1|\}}{\min\{|l_1|, |l_2|\}}.$$

The proof of Claim 3.12 is given after the proof of the two previous claims that appear in this proof.

By Claim 3.12 it holds that

$$\Pr[\text{disagreement on } u|\mathcal{A}] \leq \frac{t}{|\mathcal{S}| - t}$$

since in our case it holds $|l_1 \backslash l_2|, |l_2 \backslash l_1| \leq t$ and $|l_1|, |l_2| \geq \mathcal{S} - t$, where l_1, l_2 are defined as in the statement of Claim 3.12. Combining all the above facts, we get the lemma. $\diamond$

We now proceed to prove the claims stated in the proof of Lemma 3.9.

Proof of Claim 3.10. The coupling $\mathcal{K}_2$ is defined as follows: Choose u.a.r. a member of Ω_C, let C be the chosen element. By using $\mathcal{K}_1$ take the corresponding element of Ω_F, let C' be that element. Then, choose u.a.r. one among the copies of C in Ω'_C and one of the copies of C' in Ω'_F.

The claim follows by noting the following: First, each of the elements of both Ω'_C and Ω'_F is chosen uniformly at random. Second, the chosen elements of Ω'_C and Ω'_F specify different colour assignments for CH_u iff C and C' do the same. $\diamond$

Proof of Claim 3.11. Consider that we choose from Ω'_F such that the element C is chosen with probability proportional to its weight, W_C. Consider the same for the set Ω'_C. If there is a coupling of these two random weighted selections above, such that the probability of the event $\mathcal{A}$ to be upper bounded as in (4), then we are done.

The assumption that for each $w \in CH_u$ the disagreement probability in the coupling $\mathcal{C}(T, \mathcal{S}, l)$ is equal to p_w, is equivalent to the following: There is a mapping, call it $f : \Omega'_F \to \Omega'_C$, which is one to one (and 'onto', since $|\Omega'_F| = |\Omega'_C|$) and for any pair of colourings $(C, f(C)) \in \Omega'_F \times \Omega'_C$, with C chosen u.a.r., the probability to specify different colourings for the vertex set CH_u is upper bounded by $\sum_{w \in CH_u} p_w$.

Clearly, the mapping f defines a coupling for the "nonweighted" joint random selection of the elements the sets Ω'_F and Ω'_C, since the two sets are equal sized. Based on f, we define a coupling for the "weighted" joint random selection of the elements of the sets Ω'_F and Ω'_C.

From Ω'_i, we produce the set Ω_i^W, for $i \in \{C, F\}$ as follows: For each $C \in \Omega'_i$, insert into Ω_i^W, W_C copies of C, i.e. the elements $\{C_1, \ldots, C_{W_C}\}$. The weighted random selection from Ω'_i is equivalent to consider that we have chosen $C \in \Omega'_i$ if a random uniform selection from Ω_i^W have chosen one of $\{C_1, \ldots, C_{W_C}\}$. Thus, the construction of a coupling of the weighted joint selection from the sets Ω'_F and Ω'_C can, equivalently, be reduced to creating a coupling that selects uniformly at random one element from each of the sets Ω_F^W and Ω_C^W. This is what we are doing in the remainder of the proof.

First, we create a mapping $f' : (\Omega_F^W \cup \omega_2) \to (\Omega_C^W \cup \omega_1)$, where $\omega_2 \subset \Omega_F^W$ and $\omega_1 \subset \Omega_C^W$ and they will be defined soon after. We construct the mapping f' based on the mapping f. Then we define the coupling which consists of choosing u.a.r. a member of $\Omega_F^W \cup \omega_2$ and then applying the chosen element to f' so as to get a member of $\Omega_C^W \cup \omega_1$. In this coupling, the marginal probability for each member in Ω_F^W to be chosen will be the same for all members. This will, also, hold for the members of Ω_C^W. The claim will follow by bounding, appropriately, the quantity $\Pr[\mathcal{A}]$ in this coupling.

We define the sets ω_1, ω_2 as we construct f'. The mapping f' is defined as follows: For each $C \in \Omega'_F$, with $f(C) = Q$ and $W_C = W_{f(C)} > 0$, set $f'(C_i) = Q_i$ for $i = 1, \ldots, W_C$. For each $C \in \Omega'_F$, with $W_C > W_{f(C)}$ and $f(C) = Q$, set $f'(C_i) = Q_i$ for $i = 1, \ldots, W_{f(C)}$ and for $i = W_{f(C)} + 1, \ldots, W_C$ set $f'(C_i)$ a u.a.r. chosen member of Ω_C^W. Let ω_1 be the set of all the elements of Ω_C^W that were randomly selected in the manner that is described above. For each $C \in \Omega'_F$, with $W_C < W_{f(C)}$ and $f(C) = Q$, we set $f'(C_i) = Q_i$ for $i = 1, \ldots, W_C$, and for $i = W_C + 1, \ldots, W_{f(C)}$, for the copy Q_i choose u.a.r. a member of Ω'_F to correspond to. Let ω_2 be the set of all the elements of Ω'_F that were randomly selected in the manner that is described above.

If we choose uniformly at random from $\Omega_F^W \cup \omega_2$, each element of Ω_F^W appears equiprobably. Similarly, if we choose u.a.r. from $\Omega_C^W \cup \omega_1$, each element of Ω_C^W appears equiprobably. Furthermore, if we choose u.a.r. from $\Omega_F^W \cup \omega_2$, and apply f' to get a member from $\Omega_C^W \cup \omega_1$, all the members of $\Omega_C^W \cup \omega_1$ have the same probability to be chosen, since the mapping f' is one to one and onto ($|\Omega_F^W \cup \omega_2| = |\Omega_C^W \cup \omega_1|$). Thus, in the coupling where we choose u.a.r. $\Omega_F^W \cup \omega_2$ and apply the mapping f' and get a member of Ω_C^W, the marginal probability for all the members of Ω_F^W (and Ω_C^W) to be chosen is the same.

What remains to be shown is that in the coupling, above, the event $\mathcal{A}$ occurs with probability $\Pr[\mathcal{A}]$ which is upper bounded as in (4).

Clearly, for the colourings in the pairs $(C, f(C)) \in \Omega'_F \times \Omega'_C$ that define the same colour assignment for the vertex set CH_u, we have $W_C = W_{f(C)}$. For $C \in \Omega'_F$ in such a pair of colourings, it holds that the copy C_i, that C has in Ω_F^W, is corresponded through f' to the copy Q_i, that $Q = f(C)$ has in Ω_C^W, i.e. $Q_i = f(C_i)$, for $i = 1, \ldots, W_C$. Note that the event $\mathcal{A}$ does not hold for these pairs, $(C_i, f'(C_i))$ for $i = 1, \ldots, W_C$.

For each pair $(C, f(C)) \in \Omega'_F \times \Omega'_C$ that define a different colour assignment for the vertex set CH_u, it does not necessarily hold $W_C = W_{f(C)}$. Consider, first, the case where $W_C = W_{f(C)}$. Then, for $C \in \Omega'_F$ in such a pair of colourings, it holds that each copy C_i, that C has in Ω_F^W, is corresponded through f' to the copy Q_i, that $Q = f(C)$ has in Ω_C^W, for $i = 1, \ldots, W_C$. The event $\mathcal{A}$ does hold for these pairs, $(C_i, f'(C_i))$ for $i = 1, \ldots, W_C$.

Finally, we consider the case where the pair $(C, f(C)) \in \Omega'_F \times \Omega'_C$ defines a different colour assignment for the vertex set CH_u and $W_C \neq W_{f(C)}$. W.l.o.g. we assume that $W_C > W_{f(C)}$. Then, for $C \in \Omega'_F$ in such a pair of colourings, it holds that each copy C_i, that C has in Ω_F^W, is corresponded through f' to the copy of Q_i, that $Q = f(C)$ has in Ω_C^W, for $i = 1, \ldots, W_{f(C)}$. The event $\mathcal{A}$ does hold for the pairs $(C_i, f'(C_i))$, $i = 1, \ldots, W_{f(C)}$. The copy C_i, for $i = W_{f(C)} + 1, \ldots, W_C$, is mapped through f' to a u.a.r. chosen member of ω_C^W. Note that the event $\mathcal{A}$ does not necessarily hold for these pair. However, we assume that it does, which is, clearly, an overestimate for the probability $\Pr[\mathcal{A}]$.

Let, $\Omega_A^W \subset \Omega_F^W \cup \omega_2$ be such that $\Omega_A^W = \{C \in \Omega_F^W \cup \omega_2 | \text{ for } (C, f'(C)) \text{ the event } \mathcal{A} \text{ holds}\}$ and $\Omega_A \subset \Omega'_F$ be $\Omega_A = \{C \in \Omega'_F | \text{ for } (C, f(C)) \text{ the event } \mathcal{A} \text{ holds}\}$. Clearly, $\Pr[\mathcal{A}] = \frac{|\Omega_A^W|}{|\Omega_F^W \cup \omega_2|}$.

Let $\Omega_i^{(>0)} \subset \Omega'_i$ be such that $\Omega_i^{(>0)} = \{C \in \Omega'_i | W_C > 0\}$ and $q_i = \frac{|\Omega_i^{(>0)}|}{|\Omega'_i|}$, for $i \in \{C, F\}$. One can see that $|\Omega_A^W| \leq |\Omega_A| \max_{C \in \Omega'_C \cup \Omega'_F}\{W_C\}$, and $|\Omega_F^W \cup \omega_2| \geq |\Omega_F^{(>0)}| \cdot \min_{C \in \Omega_C^{>0} \cup \Omega_F^{>0}}\{W_C\}$. From the fact that $|\Omega_F^{>0}| = q_F |\Omega_F|$ we get that

$$\Pr[\mathcal{A}] \leq \frac{\max_{C \in \Omega'_C \cup \Omega'_F}\{W_C\}}{q_F \cdot \min_{C \in \Omega_C^{>0} \cup \Omega_F^{>0}}\{W_C\}} \frac{|\Omega_A|}{|\Omega_F|}.$$

Clearly, $q_F = q_{k,\mathcal{S}}$, where $q_{k,\mathcal{S}}$ is defined in the statement of the claim. The claim follows by noting that $\frac{|\Omega_A|}{|\Omega_F|} \leq \sum_{w \in CH_u} p_w$. ◇

Proof of Claim 3.12. The coupling with which we can choose the same colour for the vertex u with probability indicated in the statement of the claim is the *maximal coupling* (see [9]).

More specifically, we assume, w.l.o.g., that $|l_1| \geq |l_2|$. Let U be a random variable uniformly distributed in (0, 1). We assume that if $\frac{i-1}{|l_1|} < U < \frac{i}{|l_1|}$, we choose the color $i \in l_1$, for $= 1, \ldots, |l_1|$. Also, $\forall i \in l_1 \cap l_2$ assume that if $\frac{i-1}{|l_1|} < U < \frac{i}{|l_1|}$ we choose i in l_2. For U everywhere else in (0, 1) make an arbitrary arrangement so as each element of l_2 is chosen with probability $1/l_2$. By the assumption that $|l_1| \geq |l_2|$, all members $i \in l_1 \cap l_2$ we have been assigned intervals which correspond to probability $1/|l_1| \leq 1/|l_2|$.

Clearly, the interval in (0, 1) that corresponds to choosing different colourings from l_1 and l_2 is of length $\frac{|l_1 \backslash l_2|}{|l_1|}$. The claim follows by the fact that

$$\frac{|l_1 \backslash l_2|}{|l_1|} \leq \frac{\max\{|l_1 \backslash l_2|, |l_2 \backslash l_1|\}}{\min\{|l_1|, |l_2|\}}. \quad \diamond$$

Lemma 3.13. *Consider the tree T, the integers $\mathcal{S}$, l and the coupling $\mathcal{C}(T, \mathcal{S}, l)$. Let u be a vertex of T which is* nonmixing *and has k children. If for every $w \in CH_u$ the probability of disagreement is p_w, then, for the vertex u, the probability of disagreement p_u is bounded as*

$$p_u \leq S \frac{1}{q_{k,\mathcal{S}}} \left(\sum_{w \in CH_u} p_w \right) \tag{5}$$

where $q_{k,\mathcal{S}}$ is the probability of the event that after k trials, not all elements of the set $[\mathcal{S}]$ have been chosen, when at each trial we choose uniformly at random a member of $[\mathcal{S}]$.

Proof. When the vertex u is to be coloured in $\mathcal{C}(T, \mathcal{S}, l)$, assume that the processes ColourRoot(T, $\mathcal{S}$, $C_1(V_1)$) and ColourRoot(T, $\mathcal{S}$) during step 2 choose from the set of colourings Ω_C and Ω_F, correspondingly. Let $\mathcal{A}$ be the event that ColourRoot($T, \mathcal{S}, C_1(V_1)$) and ColourRoot($T, \mathcal{S}$) choose colourings from Ω_C and Ω_F, correspondingly, that specify two different colour assignments for the set CH_u. Then it holds

$$\Pr[\text{disagreement on } u] = \Pr[\text{disagreement on } u | \mathcal{A}] \Pr[\mathcal{A}] + \Pr[\text{disagreement on } u | \overline{\mathcal{A}}] \Pr[\overline{\mathcal{A}}].$$

It should be mentioned that, if the event $\mathcal{A}$ does not hold ($\overline{\mathcal{A}}$ holds), then there is a coupling for step (3) of the ColourRoot that assigns the same colour to the vertex u in $\mathcal{C}$, i.e. $\Pr[\text{disagreement on } u | \overline{\mathcal{A}}] = 0$. Thus,

$$\Pr[\text{disagreement on } u] = \Pr[\text{disagreement on } u | \mathcal{A}] \Pr[\mathcal{A}]. \tag{6}$$

We derive appropriate upper bounds for the probability terms in (6) so as to prove the lemma, i.e. the terms $\Pr[\mathcal{A}]$ and $\Pr[\text{disagreement on } u|\mathcal{A}]$. We work exactly in the same manner as in the proof of Lemma 3.9 so as to get an upper bound for the term $\Pr[\mathcal{A}]$, i.e. we have

$$\Pr[\mathcal{A}] \leq \frac{1}{q_{k,\mathcal{S}}} \frac{\max_{C\in\Omega'_F\cup\Omega'_C}\{W_C\}}{\min_{C\in\Omega'_F\cup\Omega'_C}\{W_C|W_C>0\}} \sum_{w\in CH_u} p_w.$$

Note that at the coupling, of the step (2) of the ColourRoot, for colouring u, no member of either Ω'_C or Ω'_F is assigned weight more than $\mathcal{S}$ and the minimum non zero weight is 1. Furthermore, for a nonmixing vertex u of sufficiently high degree, there are colourings of its children that use every colour in $[\mathcal{S}]$, these colourings are assigned weight zero, in this case we have $q_{k,\mathcal{S}} \leq 1$.

The lemma follows by assuming that $\Pr[\text{disagreement on } u|\mathcal{A}] = 1$ which is, clearly, an overestimate for this probability term. $\diamond$

3.2. *The case of a tree — The proof of Lemma 3.16*

Consider an instance of $G_{n,d/n}$, where $d > 1$, and for each vertex v consider the graph $G_{v,d,\epsilon \log n}$, where $\epsilon = \frac{0.9}{\log(e^2d/2)}$. By Lemma 2.1 it holds that w.h.p. $G_{v,d,\epsilon \log n}$ is either a unicyclic graph or a tree. Here, we condition that the graph $G_{v,d,\epsilon \log n}$ is a tree.

Definition 3.14. The graph $G_{v,d,\epsilon \log n}$ when we condition that it is a tree, defines a probability space over the trees which we denote by T_d.

Note that each nonleaf vertex of an instance of T_d has a number of children whose distribution is dominated by $B(n, d/n)$, i.e. the binomial distribution with parameters n and d/n.

Consider the coupling $\mathcal{C}(T, \mathcal{S}, \lfloor\epsilon \log n\rfloor)$, where T is an instance of an instance of T_d rooted at the vertex r and ϵ, $\mathcal{S}$ are as large as is indicated in Lemma 3.16. Lemma 3.16 will follow by showing that with probability, at least, $1 - 2n^{-1.25}$ it holds $\mathrm{p_r} \leq n^{-1.25}/2$ (see Corollary 3.7).

In $\mathcal{C}(T, \mathcal{S}, l)$, where T is an instance of T_d rooted at r and a given $\mathcal{S}$, the disagreement probability $\mathrm{p_r}$ depends only on the structure of T. We remind the reader that in $\mathcal{C}$, we assume that the boundary conditions are set so as to *maximize* the disagreement probability $\mathrm{p_r}$. Clearly, $\mathrm{p_r}$ is a *random variable*. We use the Lemma 3.9 and Lemma 3.13 to derive an upper bound for the expectation of $\mathrm{p_r}$ which depends on l, $\mathcal{S}$ and t, the maximum number of children of a mixing vertex. Let $q(t)$ be the probability for a random variable, distributed as in $B(n-1, d/n)$, for fixed d, to be less than t.

Lemma 3.15. *For positive integers $\mathcal{S}$, l, real $d > 1$ and T, an instance of T_d rooted at the vertex r, the expectation of the disagreement probability $\mathrm{p_r}$, in the coupling $\mathcal{C}(T, \mathcal{S}, l)$, is bounded as*

$$E[p_r] \leq \left(d\frac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2}q(t) + 2d\left(\mathcal{S}(1-q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1}\left(\exp\left\{\frac{d}{(\mathcal{S}-1)}\right\} - q(t)\right)\right)\right)^l. \tag{7}$$

Proof. We remind the reader that t stands for the maximum number of children of a mixing vertex. Let $q(t)$ be the probability for a random variable, distributed as in $B(n-1, d/n)$, for fixed d, to be less than t. Let

$$a(i) = \begin{cases} \dfrac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2} & \text{if } i \leq t \\ \dfrac{\mathcal{S}}{q_{i,\mathcal{S}}} & \text{otherwise} \end{cases}$$

where $q_{i,\mathcal{S}}$, is as defined in the statement of Lemma 3.13.

Consider the coupling $\mathcal{C}(T, \mathcal{S}, l)$, where T is an instance of T_d rooted at the vertex r. Let $E[\mathrm{p_r}]$ be the expectation of the disagreement probability on the root r. Conditioning on the number of children of r and the disagreement probability p_w, $\forall w \in CH_r$ in $\mathcal{C}(T, \mathcal{S}, l)$, by Lemma 3.9 and Lemma 3.13 we get

$$E[\mathrm{p_r}|p_w,\ \forall w \in CH_r] \leq a(|CH_r|) \cdot \sum_{w\in CH_r} p_w.$$

By definition, $\forall w \in CH_r$, p_w is upper bounded by the disagreement probability on the vertex w in the coupling $\mathcal{C}(T_w, \mathcal{S}, l-1)$ where T_w is the subtree of T rooted at vertex w. Call this disagreement probability p^*_w. It should be clear to the reader that p_w refers to the coupling $\mathcal{C}(T, \mathcal{S}, l)$, while p^*_w refers to $\mathcal{C}(T_w, \mathcal{S}, l-1)$. It is direct that

$$E[\mathrm{p_r}] \leq \sum_{i=0}^{n} ia(i)\Pr[|CH_r| = i]E[p^*_w] \quad \text{for } w \in CH_r.$$

Note that the random variables p^*_w for $w \in CH_r$ are identically distributed and independent of the number of children of r. Also, noting that the function $f(i) = i \cdot a(i)$ is increasing for $t << \mathcal{S}$ and by the fact that the distribution of the number of children of r is dominated by the $B(n, d/n)$, (by proposition 9.1.2. of [12]), it holds that

$$E[\mathrm{p_r}] \leq \sum_{i=0}^{n} i \cdot a(i) \binom{n}{i} p^i (1-p)^{n-i} E[p^*_w] \quad \text{for } w \in CH_r \tag{8}$$

where $p = d/n$. Let $S_1 = \sum_{i=0}^{t} i \cdot a(i) \binom{n}{i} p^i (1-p)^{n-i}$ and $S_2 = \sum_{i=t+1}^{n} i \cdot a(i) \binom{n}{i} p^i (1-p)^{n-i}$.

$$\begin{aligned} S_1 &\leq \frac{t \cdot \mathcal{S}}{(\mathcal{S}-t)^2} \sum_{i=0}^{t} i \binom{n}{i} p^i (1-p)^{n-i} \\ &= \frac{t \cdot \mathcal{S}}{(\mathcal{S}-t)^2} np \sum_{i=0}^{t-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} \\ &= \frac{t \cdot \mathcal{S}}{(\mathcal{S}-t)^2} q(t) d. \end{aligned}$$

Before calculating S_2, we eliminate the probability term $q_{i,\mathcal{S}}$ from the $a(i)$ for $i > t$. For $q_{i,\mathcal{S}}$ it holds that

$$q_{i,\mathcal{S}} \geq \mathcal{S} \left(1 - \frac{1}{\mathcal{S}}\right)^i \left(1 - q_{i,(\mathcal{S}-1)}\right)$$

i.e. the probability of the event "not choosing *some* element of $[\mathcal{S}]$ after i trials" is greater than, or equal to the probability of the event "not choosing *exactly* one element of $[\mathcal{S}]$", since the second event is a special case of the first one. Furthermore, since $q_{k,(\mathcal{S}-1)} \leq q_{k,\mathcal{S}}$ we get that

$$q_{i,\mathcal{S}} \geq \mathcal{S} \left(1 - \frac{1}{\mathcal{S}}\right)^i \left(1 - q_{i,\mathcal{S}}\right).$$

Let $\Omega = \{1, \ldots, n\}$ and let $t_0 = \sup\{t \in \Omega |\ q_{t,\mathcal{S}} \geq 1/2\}$. Instead of using $q_{i,\mathcal{S}}$ we make the following simplification. For $i > t_0$ we bound the quantity $1/q_{i,\mathcal{S}}$ as

$$\frac{1}{q_{i,\mathcal{S}}} \leq \frac{2}{\mathcal{S}\left(1 - \frac{1}{\mathcal{S}}\right)^i} = \frac{2}{\mathcal{S}} \left(\frac{\mathcal{S}}{\mathcal{S}-1}\right)^i.$$

Also, for $i \leq t_0$, clearly, $1/q_{i,\mathcal{S}} \leq 2$.

$$\begin{aligned} S_2 &\leq 2\mathcal{S} \sum_{i=t+1}^{t_0} i \binom{n}{i} p^i (1-p)^{n-i} + 2 \sum_{i=t_0+1}^{n} i \binom{n}{i} \left(\frac{\mathcal{S}}{\mathcal{S}-1}\right)^i p^i (1-p)^{n-i} \\ &\leq 2\mathcal{S} \sum_{i=t+1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} + 2 \sum_{i=t+1}^{n} i \binom{n}{i} \left(\frac{\mathcal{S}}{\mathcal{S}-1}\right)^i p^i (1-p)^{n-i} \\ &\leq 2\mathcal{S} np \sum_{i=t}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} + 2np \frac{\mathcal{S}}{\mathcal{S}-1} \sum_{i=t}^{n-1} \binom{n-1}{i} \left(\frac{\mathcal{S}}{\mathcal{S}-1}\right)^i p^i (1-p)^{n-1-i} \\ &\leq 2\mathcal{S} d \left(1 - \sum_{i=0}^{t-1} \binom{n-1}{i} p^i (1-p)^{n-1-i}\right) \\ &\quad + 2\frac{\mathcal{S}}{\mathcal{S}-1} d \left(\left(1 - p + \frac{\mathcal{S}}{\mathcal{S}-1} p\right)^{n-1} - \sum_{i=0}^{t-1} \binom{n-1}{i} \left(\frac{\mathcal{S}}{\mathcal{S}-1}\right)^i p^i (1-p)^{n-1-i}\right) \\ &\leq 2\mathcal{S} d\,(1 - q(t)) + 2\frac{\mathcal{S}}{\mathcal{S}-1} d \left(\left(1 + \frac{1}{\mathcal{S}-1} p\right)^{n-1} - \sum_{i=0}^{t-1} \binom{n-1}{i} p^i (1-p)^{n-1-i}\right) \\ &\leq 2\mathcal{S} d\,(1 - q(t)) + 2\frac{\mathcal{S}}{\mathcal{S}-1} d \left(\left(1 + \frac{1}{\mathcal{S}-1} p\right)^{n-1} - q(t)\right) \\ &\leq 2d \left(\mathcal{S}(1 - q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1} \left(\exp\{d/(\mathcal{S}-1)\} - q(t)\right)\right). \end{aligned}$$

Substituting the bounds for S_1 and S_2 in (8) we get

$$E[\mathrm{p_r}] \leq \left(d \frac{t\mathcal{S}}{(\mathcal{S}-t)^2} q(t) + 2d \left(\mathcal{S}(1 - q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1} \left(\exp\{d/(\mathcal{S}-1)\} - q(t)\right)\right)\right) E[p^*_w]$$

for $w \in CH_r$. We can substitute $E[p^*_w]$ in the same manner as $E[\mathrm{p_r}]$, i.e. by using induction and assuming that for the vertices at distance l from the root the expectation of the probability of disagreement is 1. Then, the lemma follows. $\diamond$

Finally, Lemma 3.16 follows by setting appropriate quantities for $\mathcal{S}$ and l in (7) and then by applying the Markov inequality. Here it is crucial to remark that if d is sufficiently large, then for $t \geq 7d$ it holds $q(t) \geq 1 - d^{-28}$.

Lemma 3.16. *Consider the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$, for $d > 1$, $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$ and for $G_{v,d,\epsilon \log n}$ we assume that it is a tree. If the $\mathcal{S}$ is a sufficiently large constant, then with probability, at least, $1 - 2n^{-1.25}$ for the above system it holds that $SD(v, \lfloor \epsilon \log n \rfloor) = n^{-1.25}$. For sufficiently large d, we should have $\mathcal{S} \geq d^{14}$.*

Proof. In the coupling $\mathcal{C}(G_{v,d,\epsilon \log n}, \mathcal{S}, l)$ it holds that the expectation of p_v is bounded as

$$E[p_v] \leq \left(d\frac{t \cdot \mathcal{S}}{(\mathcal{S} - t)^2} q(t) + 2d\left(\mathcal{S}(1 - q(t)) + \frac{\mathcal{S}}{\mathcal{S} - 1}\left(\exp\{d/(\mathcal{S} - 1)\} - q(t)\right)\right)\right)^l \tag{9}$$

where l is the minimum distance of v and the boundary set L, $q(t)$ is equal to the probability for a random variable, distributed as in $B(n-1, d/n)$, for fixed d, to be less than t, the maximum number of children of a mixing vertex.

Set $l = \epsilon \log n$ in (9), where $\epsilon = \frac{0.9}{\log(e^2 d/2)}$. So as to prove the lemma, it suffices to show that for $\mathcal{S}$ as described in the statement (of the lemma) and appropriately large t we get $E[p_v] \leq n^{-2.5}$. Clearly, for $E[p_v] \leq n^{-2.5}$ and by using the Markov Inequality (see [3]) we can get that

$$\Pr[p_v \geq n^{-1.25}/2] \leq 2\frac{E[p_v]}{n^{-1.25}} \leq 2n^{-1.25}.$$

If $E[p_v] \leq n^{-2.5}$, then with probability at least $1 - \Pr[p_v \geq 2n^{-1.25}] \geq 1 - 2n^{-1.25}$ for the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$ it holds that $SD(v, \lfloor \epsilon \log n \rfloor) \leq n^{-1.25}$, which proves the lemma. This statement follows by Definition 1.6 and Theorem 3.5. Thus, what remains to be shown is that there are appropriate values for t and $\mathcal{S}$ such that $E[p_v] \leq n^{-2.5}$.

First, we show that if d is a sufficiently large constant, then for $\mathcal{S} \geq d^{14}$ and t such that $q(t) \geq 1 - d^{-28}$ we get $E[p_v] \leq n^{-2.5}$. Using Corollary 3.2 we see that when $t \geq \max\{7d, \ 28\log d + 1\}$ it holds $q(t) \geq 1 - d^{-28}$. Assuming that d is a sufficiently large constant, we substitute the values of $\mathcal{S}$ with d^{14} and $t = 7d$ in (9) and get

$$E[p_v] \leq \left(\frac{7d^{16}}{(d^{14} - 7d)^2} + 2d\left(d^{14}d^{-28} + \frac{d^{14}}{d^{14} - 1}\left(1 + \frac{d}{d^{14} - 1} + \frac{e^{\xi}}{2!}\frac{d^2}{(d^{14} - 1)^2} - 1 + d^{-28}\right)\right)\right)^{\epsilon \log n}$$

where $0 < \xi < d/(d^{14} - 1)$. In the above inequality we used the fact that $1 - d^{-28} \leq q(t) \leq 1$ and we substituted $e^{d/(\mathcal{S}-1)}$ by an appropriate polynomial, which is derived by MacLaurin series of the function $f(x) = e^x$, for real x. Thus, we get

$$\begin{aligned} E[p_v] &\leq \left(\frac{7d^{-12}}{(1 - 7d^{-13})^2} + 2d\left(d^{-14} + \frac{1}{1 - d^{-14}}\left(\frac{d^{-13}}{1 - d^{-14}} + \frac{e}{2}\frac{d^{-26}}{(1 - d^{-14})^2} + d^{-28}\right)\right)\right)^{\epsilon \log n} \\ &\leq \left(d^{-12}\left(\frac{7}{(1 - 7d^{-13})^2} + 2d^{-1} + \frac{2}{(1 - d^{-14})^2} + \frac{ed^{-13}}{(1 - d^{-14})^3} + \frac{2d^{-15}}{1 - d^{-14}}\right)\right)^{\epsilon \log n}. \end{aligned}$$

Taking d at least 20, we get that

$$E[p_v] \leq n^{\epsilon \log(9.2d^{-12})}.$$

Replacing ϵ, we see that it suffices to hold $0.9\log(9.2d^{-12}) \leq -10\log(e^2 d/2)$, or $9.2d^{-12} \leq (e^2 d/2)^{-11.11}$ which clearly holds for sufficiently large constant d.

For relatively smaller d, one can easily see that setting $\mathcal{S} = d^x$, for an appropriately large constant exponent x and arranging the quantity t so as $q(t) \geq 1 - d^{-2x}$ can get

$$E[p_v] \leq \left(d^{-x+2}\left(\frac{t/d}{(1 - d^{-x}t)^2} + 2d^{-1} + \frac{2d^{-x-1}}{1 - d^{-x}} + \frac{2}{(1 - d^{-x})^2} + e\frac{d^{-x+1}}{(1 - d^{-x})^3}\right)\right)^{\epsilon \log n}.$$

We take x sufficiently large so as to have $1 - d^{-x} \geq 1 - 10^{-3}$ and $xd^{-x} \leq 10^{-3}$.

If $t = 7d$, then, with the above assumptions, we can easily derive that $E[p_v] \leq (d^{-x+2}16)^{\epsilon \log n}$. If $E[p_v] \leq n^{-2.5}$, then we should have $16d^{-x+2} \leq (e^2 d/2)^{11.11}$, which clearly holds for sufficiently large x.

If $2x\log d + 1 > 7d$, then by Corollary 3.2 we should have $t = 2x\log d + 1$. With the assumptions we have made for x we get that $E[p_v] \leq (d^{-x+2}(2.1x\frac{\log d}{d} + 9))^{\epsilon \log n}$. If $E[p_v] \leq n^{-2.5}$, then it should hold $(d^{-x+2}(2.1x\frac{\log d}{d} + 9)) \leq (e^2 d/2)^{11.11}$, which clearly holds for sufficiently large x. The lemma follows. ◇

3.3. The case of a unicyclic graph — The proof of Lemma 3.21

Consider an instance of $G_{n,d/n}$, and the set of its subgraphs $G_{v,d,\epsilon \log n}$, as in Section 3.2. By Lemma 2.1, it holds that w.h.p. $G_{v,d,\epsilon \log n}$ is either a unicyclic graph or a tree. Now we condition that $G_{v,d,\epsilon \log n}$ is a unicyclic graph.

First, we show how can we extend the techniques for proving Lemma 3.16, i.e. proving spatial mixing properties of system with an underlying graph which is a tree, and prove Lemma 3.21, which refers to systems with a unicyclic underlying graph.

Consider the depth first search in $G_{v,d,\epsilon \log n}$ that starts from the vertex v and let u be the first vertex of the unique cycle that is reached by the search. Clearly, there are two possible choices for this search to explore the vertices of the cycle that u belongs to. If w_1 and w_2 are the vertices on the cycle that are also adjacent to u, then let T^1 and T^2 be the two depth-first search trees of $G_{v,d,\epsilon \log n}$, rooted at v, with the first tree having u adjacent only to w_1 and the second having u adjacent only to w_2.

Definition 3.17. With the above notation, the tree $T_{r,d,\epsilon \log n}$ is isomorphic to the tree that comes up from the union of T^1 and $T^2_{w_2}$ plus an edge connecting the vertices u in T^1 with w_2 in $T^2_{w_2}$. The root r of $T_{r,d,\epsilon \log n}$ corresponds to the vertex v in T^1.

Note that the number of children of a nonleaf vertex of $T_{r,d,\epsilon \log n}$ has distribution which is dominated by $B(n, d/n)$ with the condition that it is at least 2.

Each of the trees T^1 and $T^2_{w_2}$, in the definition of $T_{r,d,\epsilon \log n}$, are isomorphic to some subgraph of $G_{v,d,\epsilon \log n}$, i.e. there is a correspondence between the vertices in T^1 and $T^2_{w_2}$ with the vertices in $G_{v,d,\epsilon \log n}$. Based on this correspondence, we can define a (surjective) function $h : V_T \to V_G$, where V_T is the set of vertices of $T_{r,d,\epsilon \log n}$ and V_G the set of vertices in $G_{v,d,\epsilon \log n}$.

Let L be the set that contains all the vertices in $G_{v,d,\epsilon \log n}$ that are at a graph distance, at least, $\lfloor \epsilon \log n \rfloor$ from v. Consider the $\mathcal{S}$-colouring $C_1(L)$ which is such that the set the total variation distance of the Gibbs measures $\mu(X_v|C_1(L))$ and $\mu(X_v)$, as these are specified by the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$, is maximized.

For the tree $T_{r,d,\epsilon \log n}$ derived by $G_{v,d,\epsilon \log n}$, the integers $\mathcal{S}$ and l, let $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, l)$, be a coupling of the stochastic processes ColourRoot$(T_{r,d,\epsilon \log n}, \mathcal{S}, C_1^T(A))$ and ColourRoot$(T_{r,d,\epsilon \log n}, \mathcal{S})$. The set A is such that $\forall \hat{v} \in A\ \exists \hat{u} \in (V_1 \cup L)$ such that $h(\hat{v}) = \hat{u}$ and $C_1^T(\hat{v}) = C_1(\hat{u})$. The coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, l)$ is, in essence, the same as the coupling $\mathcal{C}(T_{r,d,\epsilon \log n}, \mathcal{S}, l)$ with only one difference: Consider $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, l)$ when it assigns colourings to a *nonmixing* vertex u which has i children and assume that for each $w \in CH_u$ the disagreement probability is p_w. Then, the disagreement probability p_u, for the vertex u, is bounded as

$$p_u \leq \frac{\mathcal{S}}{q_{i,S,2}} \sum_{w \in CH_u} p_w. \tag{10}$$

where $q_{i,S,2}$ is the probability of the event that after k trials not all elements of the set $[\mathcal{S}]$ have been chosen, when at each trial we choose u.a.r. a member of $[\mathcal{S}]$ and conditioning that the first two trials chose different elements of $[\mathcal{S}]$.

Comparing the bound in (10) with that was given in (5), in the statement of Lemma 3.13, we see that $q_{i,S,2} \leq q_{i,S}$. This fact implies that the coupling $\mathcal{C}'$ exists as, on the same input, it gives either the same, or worse bounds than $\mathcal{C}$ for the disagreement probabilities.

Lemma 3.18. *Consider the graph $G_{v,d,\epsilon \log n}$, the corresponding tree $T_{r,d,\epsilon \log n}$, with $d > 1$ and $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$, and the positive integer $\mathcal{S}$. If $\mathrm{p_r}$ is the bound for the disagreement probability that we derive from Lemma 3.9 and (10) for the coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$, then for the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$ it holds that $SD(v, \lfloor \epsilon \log n \rfloor) \leq 2\mathrm{p_r}$.*

Proof. Let L be the set that contains all the vertices in $G_{v,d,\epsilon \log n}$ that are at graph distance, at least, $\lfloor \epsilon \log n \rfloor$ from v. Consider the $\mathcal{S}$-colouring $C(L)$ which is such that the total variation distance of the Gibbs measures $\mu(X_v|C(L))$ and $\mu(X_v)$, as these are specified by the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$, is maximized.

Let u be the vertex in $G_{v,d,\epsilon \log n}$ which belongs to the unique cycle of $G_{v,d,\epsilon \log n}$ and among all the vertices on the cycle it has the smallest distance from v. Let G^u be the *connected* subgraph of $G_{v,d,\epsilon \log n}$ that contains the vertex u and the vertices whose distance from v is greater than that of the vertex u from v. It is easy to see that G^u is a unicyclic graph and $G_{v,d,\epsilon \log n} \backslash G^u$ is a tree.

Assume that there is a coupling such that choosing uniformly at random a member from each of the sets of $\mathcal{S}$-colourings $\Omega(G^u, \mathcal{S}, C(L \cap G^u))$ and $\Omega(G^u, \mathcal{S})$ the probability for the two chosen colourings to specify different colour assignments for the vertex u is Q. Also, let p_v be the upper bound of the disagreement probability in the coupling of the processes ColourRoot$(T, \mathcal{S}, C(T \cap L))$ and ColourRoot$(T, \mathcal{S})$, where $T = G_{v,d,\epsilon \log n} \backslash (G^u \backslash \{u\})$, that we derive by using Lemma 3.9 and (10) and assuming that the disagreement probability of the vertex u is set a priori to Q. Note that the graph $G_{v,d,\epsilon \log n} \backslash (G^u \backslash \{u\})$ is a tree. It is easy for one to see that p_v is, also, an upper bound for the total variation distance of the Gibbs measures $\mu(X_v|C(L))$ and $\mu(X_v)$, as these are specified by the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$.

For an appropriately constructed tree T^u, with respect to G^u, and appropriate boundary condition C_1', if the coupling of the processes ColourRoot$(T^u, \mathcal{S}, C_1')$ and ColourRoot$(T^u, \mathcal{S})$ has disagreement probability P at the vertex u, for which $Q \leq P$,[6]

[6] The probability Q is at the beginning of the previous paragraph.

then the lemma will follow. W.l.o.g. assume that the vertex v belongs to the unique cycle of $G_{v,d,\epsilon \log n}$, i.e. the G^u and $G_{v,d,\epsilon \log n}$ are identical.

Consider the function $h : V_T \to V_G$, where V_T is the set of vertices of $T_{r,d,\epsilon \log n}$ and V_G the set of vertices in $G_{v,d,\epsilon \log n}$, $h(\cdot)$ is defined in the paragraph after Definition 3.17. Let L' be the set of vertices in $T_{r,d,\epsilon \log n}$ such that $L' = \{u \in V_T | h(u) \in L\}$. It is direct that the vertex set L' is at a distance, at least, $\lfloor \epsilon \log n \rfloor$ from r, in $T_{r,d,\epsilon \log n}$. Let, also, $C_T(L')$ be a colouring which assigns each $w \in L'$ the same colour as $C(L)$ specifies for $w_0 = h(w)$.

If N_v is the vertex set that contains all the adjacent vertices of v in $G_{v,d,\epsilon \log n}$, then for each $w_0 \in N_v$ we define G^{w_0} to be the *connected* component of $G_{v,d,\epsilon \log n} \backslash \{v\}$ that contains w_0. It is straightforward that the component G^{w_0} is a tree which is isomorphic to the subtree T_w of $T_{r,d,\epsilon \log n}$, where $h(w) = w_0$. This isomorphism implies that there is a correspondence between the elements of the sets $\Omega(G^{w_0}, \mathcal{S}, C(L))$ and $\Omega(T_w, \mathcal{S}, C_T(L'))$ such that any two corresponding colourings C^1 and C^2 have the property that $\forall u_1 \in T_w\ C^1(h(u_1)) = C^2(u_1)$. Clearly, there is a similar correspondence between the members of the sets of $\Omega(G^{w_0}, \mathcal{S})$ and $\Omega(T_w, \mathcal{S})$, for $w \in CH_r$ and $w_0 = h(w)$.

Assume that in $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$ for each $w \in CH_r$ the disagreement probability is p_w. Then, with the above correspondence between the pairs of sets $\Omega(G^{w_0}, \mathcal{S}, C(L))$ with $\Omega(T_w, \mathcal{S}, C_T(L'))$ and $\Omega(G^{w_0}, \mathcal{S})$ with $\Omega(T_w, \mathcal{S})$ we conclude that there is a coupling such that choosing u.a.r. from the sets $\Omega(G^{w_0}, \mathcal{S}, C(L))$ and $\Omega(G^{w_0}, \mathcal{S})$ the probability for the two chosen elements to assign different colour to the vertex w_0 is upper bounded by p_w, where $h(w) = w_0$. With these facts, we prove the following claim.

Claim 3.19. *If the coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$ has disagreement probability on the vertex r upper bounded by* $\mathrm{p_r}$, *then we can have a coupling of the uniform random choices from the sets $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S})$ such that the two chosen elements specify different colour assignments for the vertex v with probability upper bounded by* $\mathrm{p_r}$.

The proof of Claim 3.19 is given after the end of this proof.

With the above claim and what follows, we get the proof of the lemma. Let $\tilde{C}(L)$ and $\hat{C}(L)$ be the two colourings which maximize the total variation distance of the measures $\mu(X_v|\tilde{C}(L)$ and $\mu(X_v|\hat{C}(L))$, as these are specified by the system $PCS(G_{v,d,\epsilon \log n}, \mathcal{S})$.

$$\begin{aligned}
SD(v, \epsilon \log n) &= d_{TV}\left(\mu(X_v|\tilde{C}(L), \mu(X_v|\hat{C}(L))\right) \\
&\le d_{TV}\left(\mu(X_v|\tilde{C}(L), \mu(X_v)\right) + d_{TV}\left(\mu(X_v), \mu(X_v|\hat{C}(L))\right) \\
&\le 2d_{TV}\left(\mu(X_v|C(L), \mu(X_v)\right) \\
&\le 2\mathrm{p_r}
\end{aligned}$$

where $\mathrm{p_r}$ is the bound of the disagreement probability on the vertex r in $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$. ◇

We now prove the claim that appears in the proof of Lemma 3.18.

Proof of Claim 3.19. The claim will follow by proving that there exists a coupling of uniform random choices from $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S})$ such that the two chosen elements specify different colour assignments for v with probability upper bounded by the disagreement probability $\mathrm{p_r}$ in the coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \epsilon \log n)$.

Consider the coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$. Firstly, we assume that t' is the maximum number of children of a *mixing* vertex. Secondly, we assume that for each vertex $w \in CH_r$, the probability of disagreement is p_w.

By the above assumption and by what is stated in the proof of Lemma 3.18, it holds that there is a coupling of the uniform random choices from the sets $\Omega(G^{w_0}, \mathcal{S}, C(L))$ and $\Omega(G^{w_0}, \mathcal{S})$ such that the probability for the two chosen elements to assign different colour to the vertex w_0 is upper bounded by p_w, where $h(w) = w_0$. Note also that the vertices v and r have the same degree in $G_{v,d,\epsilon \log n}$ and $T_{r,d,\epsilon \log n}$, correspondingly. However, it is not direct that we can apply the Lemma 3.9 and (10) for bounding the probability of interest in the coupling of the uniform random choices from $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S})$. The difference between the coupling $\mathcal{C}'$ and that of the uniform random choices of $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S})$ is that in the second case the colour assignments of two adjacent vertices of v are correlated, while in the case of $\mathcal{C}'$ they are not.

We distinguish two cases for the degrees of the vertices r and v. In the first case we assume that the degrees are at most t' (r is mixing) and in the second that the degrees are more than t' (r is nonmixing).

In the first case, where the degree of v is at most t', despite the fact that the colour assignments of two adjacent vertices of v are correlated, we can still apply Lemma 3.9. Using the notation of the statement of Lemma 3.9 it holds the following: If there is a disagreement in the vertices in CH_u, then for bounding p_u we assume the worst case of disagreement in the colourings of the vertices in CH_u, i.e. it is assumed that all the vertices have different colour assignments in the coupling. Clearly, this leads to an overestimate for p_u, even for the case where the colourings are correlated. Therefore, this bound is, also, an overestimate for bounding the probability that v is assigned two different colourings in the coupling of the uniform random choices of $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon \log n}, \mathcal{S})$. Clearly, we get the same bound for the probability of disagreement $\mathrm{p_r}$ in the coupling $\mathcal{C}'(T_{r,d,\epsilon \log n}, \mathcal{S}, \lfloor \epsilon \log n \rfloor)$.

If the degree of the vertex v is i, which is greater than t', then we use Lemma 3.13, with a little modification. One can see that the term $1/q_{i,\mathcal{S}}$ in (5) of the statement of Lemma 3.13 is not exact for our case. More specifically, in the last paragraph of the proof of Claim 3.11, for our case the quantity q_F is not equal to $q_{i,\mathcal{S}}$ due to the fact that two vertices do not choose independently their colour assignments. However, it is direct that for this case q_F is lower bounded by the probability of the event that after i trials, not all the elements of $[\mathcal{S}]$ have been chosen, when at each trial we choose u.a.r. a member of $[\mathcal{S}]$ and conditioning that the first two trials choose different elements of $[\mathcal{S}]$. With this modification we can see that in the coupling of the uniform random choices from the sets $\Omega(G_{v,d,\epsilon\log n}, \mathcal{S}, C(L))$ and $\Omega(G_{v,d,\epsilon\log n}, \mathcal{S})$ the two chosen element specify different colour assignments for the vertex v with probability bounded by the quantity

$$\frac{\mathcal{S}}{q_{i,S,2}} \sum_{w \in CH_r} p_w.$$

Where $q_{i,S,2}$ is the probability of the event that after k trials, not all elements of $[\mathcal{S}]$ have been chosen, when at each trial we choose u.a.r. a member of $[\mathcal{S}]$ and conditioning that the first two trials chose different elements of $[\mathcal{S}]$. Clearly, this is the same bound we derive for $\mathrm{p_r}$ in the coupling $\mathcal{C}'$. The claim follows.

Towards proving Lemma 3.21, we use Lemma 3.18 which allows us to consider the tree $T_{r,d,\epsilon\log n}$ derived by unicyclic graph $G_{v,d,\epsilon\log n}$, instead of $G_{v,d,\epsilon\log n}$. We can follow the same approach as that in Section 3.2 for showing the desired spatial mixing properties for systems with underlying graph the tree $T_{r,d,\epsilon\log n}$. Note that now the coupling is $\mathcal{C}'$. Let $q(t)$ be equal to the probability for a random variable, distributed as in $B(n-1, d/n)$, for fixed d, to be less than t.

Lemma 3.20. *For positive integers $\mathcal{S}$, l, real $d > 1$ in the coupling $\mathcal{C}'(T_{r,d,\epsilon\log n}, \mathcal{S}, l)$ the expectation of the disagreement probability* $\mathrm{p_r}$ *is bounded as*

$$E[p_r] \le \left(\frac{1}{1-(d+1)e^{-d}}\left(d\frac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2}q(t) + 2d\left(\mathcal{S}(1-q(t)) + \exp\left\{\frac{d}{\mathcal{S}-2}\right\} - q(t)\right)\right)\right)^l.$$

Proof. We remind the reader that t stands for the maximum number of children of a mixing vertex. Let $q(t)$ be the probability for a random variable, distributed as in $B(n-1, d/n)$, for fixed d, to be less than t. ⋄

In $T_{r,d,\epsilon\log n}$ the number of children of a nonleaf vertex has distribution which is dominated by the $B(n, d/n)$ with the condition that there are at least two children. Let Z be a random variable distributed as in $B(n, d/n)$, clearly

$$\Pr[Z \ge 2] = 1 - \left(1-\frac{d}{n}\right)^n - n\frac{d}{n}\left(1-\frac{d}{n}\right)^{n-1} \ge 1-(d+1)e^{-d}.$$

Let

$$a(i) = \begin{cases} \dfrac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2} & \text{if } i \le t \\ \dfrac{\mathcal{S}}{q_{i,\mathcal{S},2}} & \text{otherwise} \end{cases}$$

where $q_{i,\mathcal{S},2}$, is as defined in (10). Consider the coupling $\mathcal{C}'(T, \mathcal{S}, l)$, where T is an instance of $T_{r,d,\epsilon\log n}$ rooted at the vertex r. Let $E[\mathrm{p_r}]$ be the expectation of the disagreement probability on the root r. Conditioning on the number of children of r and the disagreement probability $p_w, \forall w \in CH_r$ in $\mathcal{C}'(T, \mathcal{S}, l)$, by Lemma 3.9 and (10) we have

$$E[\mathrm{p_r}|p_w,\ \forall w \in CH_r] \le a(|CH_r|)\sum_{w\in CH_r} p_w.$$

By definition, $\forall w \in CH_r$, p_w is upper bounded by the disagreement probability on the vertex w in the coupling $\mathcal{C}'(T_w, \mathcal{S}, l-1)$ where T_w is the subtree of T rooted at vertex w. Call this disagreement probability p^*_w. It should be clear to the reader that p_w refers to the coupling $\mathcal{C}'(T, \mathcal{S}, l)$, while p^*_w to $\mathcal{C}'(T_w, \mathcal{S}, l-1)$. It is direct that

$$E[\mathrm{p_r}] \le \sum_{i=0}^{n} ia(i)\Pr[|CH_r| = i]E[p^*_w] \quad \text{for } w \in CH_r. \tag{11}$$

Also, noting that the function $f(i) = i\cdot a(i)$ is increasing for $t << \mathcal{S}$ and by the fact that the distribution of the number of children of r is dominated by the $B(n, d/n)$, with the condition that it is greater than 1, (by proposition 9.1.2. of [12]) it holds that

$$E[\mathrm{p_r}] \le \frac{1}{1-(d+1)e^{-d}}\sum_{i=0}^{n} a(i)\binom{n}{i}p^i(1-p)^{n-i}E[p^*_w] \quad \text{for } w \in CH_r$$

where $p = d/n$.

Let $S_1 = \sum_{i=0}^{t} i \cdot a(i)\binom{n}{i}p^i(1-p)^{n-i}$ and $S_2 = \sum_{i=t+1}^{n} i \cdot a(i)\binom{n}{i}p^i(1-p)^{n-i}$. Using the derivations of Lemma 3.15 we have that

$$S_1 = \frac{t \cdot \mathcal{S}}{(\mathcal{S}-t)^2} q(t)d.$$

Before calculating S_2, we eliminate the probability term $q_{i,\mathcal{S},2}$ from $a(i)$ for $i > t$. Note that $q_{i,\mathcal{S},2} > q_{i-1,\mathcal{S}-1}$, i.e. $q_{i-1,\mathcal{S}-1}$ is the probability for not choosing all the elements of a set of cardinality $\mathcal{S}-1$ after $i-1$ trials when at each trial we choose u.a.r. a member of the set. For $q_{i,\mathcal{S},2}$ it holds that

$$q_{i,\mathcal{S},2} \geq q_{i-1,\mathcal{S}-1} \geq (\mathcal{S}-1)\left(1-\frac{1}{\mathcal{S}-1}\right)^{i-1}\left(1-q_{i-1,(\mathcal{S}-2)}\right)$$

i.e. the probability of the event "not choosing *some* of the $\mathcal{S}-1$ elements after $i-1$ trials" is greater than, or equal to the probability of the event "not choosing *exactly* one of $\mathcal{S}-1$ elements after $i-1$ trials", since the second event is a special case of the first one. Furthermore, since $q_{i-1,(\mathcal{S}-2)} \leq q_{i-1,\mathcal{S}-1}$ we get that

$$q_{i-1,\mathcal{S}-1} \geq (\mathcal{S}-1)\left(1-\frac{1}{\mathcal{S}-1}\right)^{i-1}\left(1-q_{i-1,\mathcal{S}-1}\right). \tag{12}$$

Let $\Omega = \{1,\ldots,n\}$ and let $t_0 = \sup\{t \in \Omega \mid q_{t-1,\mathcal{S}-1} \geq 1/2\}$. Instead of using $q_{i-1,\mathcal{S}-1}$ we make the following simplification. For $i > t_0$ we bound the quantity $1/q_{i-1,\mathcal{S}-1}$ as

$$\frac{1}{q_{i-1,\mathcal{S}-1}} \leq \frac{1}{(\mathcal{S}-1)\left(1-\frac{1}{\mathcal{S}-1}\right)^{i-1}} = \frac{2}{\mathcal{S}-1}\left(\frac{\mathcal{S}-1}{\mathcal{S}-2}\right)^{i-1}.$$

Also, for $i \leq t_0$, clearly, $1/q_{i-1,\mathcal{S}-1} \leq 2$. With derivations similar to those in the proof of Lemma 3.15 for S_2 we get that

$$S_2 \leq 2d\left(\mathcal{S}(1-q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1}\left(\exp\{d/(\mathcal{S}-2)\} - q(t)\right)\right).$$

Substituting the bounds for S_1 and S_2 in (11) we get

$$E[\mathrm{p_r}] \leq \frac{1}{1-(d+1)e^{-d}}\left(d\frac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2}q(t) + 2d\left(\mathcal{S}(1-q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1}\left(\exp\{d/(\mathcal{S}-2)\} - q(t)\right)\right)\right)E[p_w^*]$$

for $w \in CH_r$. We can substitute $E[p_w^*]$ in the same manner as $E[\mathrm{p_r}]$, Using induction and assuming that for the vertices at a distance l from the root the expectation of the probability of disagreement is 1, the lemma follows. ◇

Lemma 3.21, follows by combining the Lemmas 3.18 and 3.20.

Lemma 3.21. *Consider a system $PCS(G_{v,d,\epsilon\log n},\mathcal{S})$, for $d > 1$, $\epsilon = \frac{0.9}{4\log(e^2 d/2)}$ and for $G_{v,d,\epsilon\log n}$ we assume that it is a unicyclic graph. If the $\mathcal{S}$ is a sufficiently large constant, then with probability, at least, $1-2n^{-1.25}$ for the above system it holds that $SD(v,\lfloor\epsilon\log n\rfloor) = n^{-1.25}$. For sufficiently large d, we should have $\mathcal{S} \geq d^{14}$.*

Proof. To prove the lemma we first show that using the coupling $\mathcal{C}'(T_{r,d,\epsilon\log n},\lfloor\epsilon\log n\rfloor)$ for the system $PCS(T_{r,d,\epsilon\log n},\mathcal{S})$ it holds that $SD(r,\lfloor\epsilon\log n\rfloor) = n^{-1.25}$ with probability at least $1-2n^{-1.25}$, when $\mathcal{S}$ is a sufficiently large constant and for sufficiently large d, we should have $\mathcal{S} \geq d^{14}$. Then, the lemma will follow by using Lemma 3.18.

In the coupling $\mathcal{C}'(T_{r,d,\epsilon\log n},\mathcal{S},l)$ it holds that the expectation of p_v is bounded as

$$E[\mathrm{p_r}] \leq \left(\frac{1}{1-(d+1)e^{-d}}\left(d\frac{t\cdot\mathcal{S}}{(\mathcal{S}-t)^2}q(t) + 2d\left(\mathcal{S}(1-q(t)) + \frac{\mathcal{S}}{\mathcal{S}-1}\left(\exp\{d/(\mathcal{S}-2)\} - q(t)\right)\right)\right)\right)^l. \tag{13}$$

The quantity $q(t)$ is equal to the probability for a random variable, distributed as in $B(n-1,d/n)$, for fixed d, to be less than t, the maximum number of children of a mixing vertex.

Set $l = \epsilon\log n$ in (9), with $\epsilon = \frac{0.9}{\log(e^2 d/2)}$. So as to prove the lemma, it suffices to show that for $\mathcal{S}$ as described in the statement (of the lemma) and appropriately large t we get $E[\mathrm{p_r}] \leq n^{-2.5}$. Clearly, for $E[\mathrm{p_r}] \leq n^{-2.5}$ and by using the Markov Inequality (see [3]) we can get that

$$\Pr[\mathrm{p_r} \geq n^{-1.25}/2] \leq 2\frac{E[\mathrm{p_r}]}{n^{-1.25}} \geq 2n^{-1.25}.$$

If $E[\mathrm{p_r}] \leq n^{-2.5}$, then with probability at least $1-\Pr[\mathrm{p_r} \geq 2n^{-1.25}] \geq 1-2n^{-1.25}$ for the system $PCS(T_{r,d,\epsilon\log n},\mathcal{S})$ it holds that $SD(r,\lfloor\epsilon\log n\rfloor) \leq n^{-1.25}$, which proves the lemma. This statement follows by Definition 1.6 and Theorem 3.5 and Lemma 3.18.

First we show that for sufficiently large d, for $\mathcal{S} \geq d^{14}$ and t such that $q(t) \geq 1-d^{-28}$ we get $E[p_r] \leq n^{-2.5}$. By Corollary 3.2 we derive that for $t = \max\{7d, 28\log d + 1\}$ it holds $q(t) \geq 1-d^{-28}$.

Assuming that d is a sufficiently large constant, we substitute $\mathcal{S}$ and t in (13) and we get

$$E[\mathrm{p_r}] \leq \left(\frac{1}{1-(d+1)e^{-d}}\left(\frac{7d^{16}}{(d^{14}-7d)^2}+ + \frac{2d}{1-d^{-14}}\left(d^{-14}+1+\frac{d}{d^{14}-1}+\frac{e^{\xi}}{2!}\frac{d^2}{(d^{14}-1)^2}-1+d^{-28}\right)\right)\right)^{\epsilon\log n}$$

where $0<\xi<d/(d^{14}-1)$. In the above inequality we used the fact that $1-d^{-28}\leq q(t)\leq 1$, and we substituted $e^{d/(\mathcal{S}-1)}$ by an appropriate polynomial, which is derived by the MacLaurin series of the function $f(x)=e^x$, for x real. Thus, we get

$$E[\mathrm{p_r}] \leq \left(\frac{d^{-12}}{1-(d+1)e^{-d}}\left(\frac{7}{(1-7d^{-13})^2}+\frac{2d^{-1}}{1-d^{-14}}+\frac{2}{(1-d^{-14})^2}+\frac{ed^{-13}}{(1-d^{-14})^3}\frac{+2d^{-15}}{1-d^{-14}}\right)\right)^{\epsilon\log n}.$$

Taking d at least 20, we get that

$$E[\mathrm{p_r}] \leq n^{\epsilon\log(9.2d^{-12})}.$$

Replacing ϵ, we see that it suffices to hold $0.9\log(9.2d^{-12})\leq -10\log(e^2d/2)$, or $9.2d^{-12}\leq (e^2d/2)^{-11.11}$ which clearly holds for sufficiently large constant d.

For relatively smaller d, one can easily see that setting $\mathcal{S}=d^x$, for an appropriately large constant exponent x and arranging the quantity t so as $q(t)\geq 1-d^{-2x}$ can get

$$E[\mathrm{p_r}] \leq \left(\frac{d^{-x+2}}{1-(d+1)e^{-d}}\left(\frac{t/d}{(1-d^{-x}t)^2}+\frac{2d^{-1}}{1-d^{-x}}+\frac{2d^{-x-1}}{1-d^{-x}}+\frac{2}{(1-d^{-x})^2}+e\frac{d^{-x+2}}{(1-d^{-x})^3}\right)\right)^{\epsilon\log n}.$$

We take x sufficiently large so as to have $1-d^{-x}\geq 1-10^{-3}$ and $xd^{-x}\leq 10^{-3}$.

If $t=7d$, then, with the above assumptions, about x, we can easily derive that $E[\mathrm{p_r}]\leq (d^{-x+2}44)^{\epsilon\log n}$. For this case to have $E[\mathrm{p_r}]\leq n^{-2.5}$ we should have $44d^{-x+2}\leq (e^2d/2)^{11.11}$, which clearly holds for sufficiently large x.

If $2x\log d+1>7d$, then by Corollary 3.2 we should have $t=2x\log d+1$. With the assumptions we have made for x we get that $E[p_v]\leq (d^{-x+2}(8x\frac{\log d}{d}+16))^{\epsilon\log n}$. Thus, if $E[\mathrm{p_r}]\leq n^{-2.5}$, then we should have $(d^{-x+2}(8x\frac{\log d}{d}+16))\leq (e^2d/2)^{11.11}$, which clearly holds for sufficiently large x.

The lemma follows. ◇

4. Properties of the algorithm

We close this work by providing the proofs of the lemma and the two theorems that appear in Section 2.2.

Proof of Lemma 2.3. By Theorem 2.2, the sampling algorithm will need, at least, d^x colours, where $x=x(d)$ is a decreasing function of d with minimum value equal to 14. It is easy for one to see that with probability, at least, $1-n^{-0.1}$ the input graph of the algorithm is colourable with at least d^x-3 colours. More specifically, the input graph is colourable with probability at least $1-n^{-0.1}$ with less colours (see Janson et.al. in [6], Section 7.2).

From now on, we assume that the input is colourable with d^x-3 colours. We remind the reader that we have, already, assumed that each of the graphs $G_{v_i,d,\epsilon\log n+2}$, for $i=1,\ldots,n$, is either unicyclic or tree. W.l.o.g. we assume that V_i' contains all the vertices at a distance at most $\lfloor\epsilon\log n\rfloor$ from v_i.

The lemma will follow by showing, that, under the above assumptions, for any colouring $C(V_i')$ such that $\tilde{\mu}_i(X_{V_i'}=C(V_i'))\neq 0$, it also holds $\Omega(G,\mathcal{S},C(V_i'))\neq\emptyset$. To show this, we are going to construct a colouring of the input graph which has the vertices in V_i' coloured as $C(V_i')$.

Since the input is colourable with at least d^x-3 colours, we can colour all the vertices at a distance at least $\lfloor\epsilon\log n\rfloor+2$ from v_i, by using the d^x-3 colours. Then, we assign the vertices in V_i' the colouring $C(V_i')$. Since $\tilde{\mu}_i(C(V_i'))>0$ it clearly holds $\Omega(G_{v_i,d,\epsilon\log n},\mathcal{S},C(V_i'))\neq\emptyset$.

The only vertices that are not coloured yet, are the vertices that are at a distance $\lfloor\epsilon\log n\rfloor+1$ from v_i. By the assumption we have made for the structure of $G_{v_i,d,\epsilon\log n+2}$ it holds that each of the uncoloured vertices either has at most two neighbors in V_i' and no uncoloured neighbor, or it has one neighbor in V_i' and one uncoloured neighbor.

For an uncoloured vertex v of the first case, we note that the neighbors in V_i use at most 2 colours while the remaining neighbors use all but 3 colours. Clearly, there exists at least one colour that does not appear in the neighborhood of v. We colour v with this colour.

In the same manner, for an uncoloured vertex v of the second case, we see that there exist, at least, 2 colours that do not appear in its neighborhood and there exists a neighbor u which is not coloured yet. We can assign u a colour that does not appear in its colored neighbors. For the vertex v there remains, at least, one available colour which is assigned to it.

The lemma follows. ◇

Proof of Theorem 2.4. The algorithm is considered successful if the input graph $G = (V, E)$ has the following properties: *First*, for the iteration i of the for-loop of the algorithm, the induced subgraph that contains v_i and all the vertices that are within a graph distance $\lfloor \epsilon \log n \rfloor + 2$ from v_i, with $\epsilon = \frac{0.9}{4\log(de^2/2)}$, is either unicyclic or a tree. According to Lemma 2.1 this holds with probability, at least, $1 - n^{-0.1}$. *Second*, for a number of colours, $\mathcal{S}$, as large as indicated by Theorem 2.2 the Gibbs measure of $PCS(G, \mathcal{S})$ exhibits the spatial mixing property stated in Theorem 2.2. *Third*, the measures $\tilde{\mu}_i$, for $i = 1, \ldots, n$, have the property indicated in Lemma 2.3.

Consider that the input graph of the algorithm G is an instance of $G_{n,d/n}$ and we take $\mathcal{S}$ as large as indicated in Theorem 2.2. Then, the algorithm is successful with probability at least $1 - O(n^{-0.1})$.

From now on, we assume that the input G belongs to this set of instances of $G_{n,d/n}$ that the algorithm is successful (which includes almost all instances for sufficiently large $\mathcal{S}$).

What remains to be shown is the bound for the total variation distance between the probability measure of the colouring that is returned by the algorithm and the uniform over all the proper $\mathcal{S}$-colouring of the input graph G.

Consider the following coupling of our algorithm and an ideal algorithm that gives a perfect uniform sample by colouring vertices one by one in some way. At each repetition, both algorithms assign a colour to some (the same) vertex in $G_{n,d/n}$. Consider a specific step of the coupling where the vertex v is to be coloured. By Theorem 2.2 and (2), we can have a sufficiently large $\mathcal{S}$ such that, conditioning on the fact that all vertices until now are identically coloured by the two algorithms, the probability for v to have a different colour assignment in the coupling is at most $n^{-1.25}$. Thus, the probability for the coupling to end with no disagreement is at least $(1 - n^{-1.25})^n > 1 - n^{-0.25}$. The theorem follows. ◇

Proof of Theorem 2.5. First, we note that the algorithm will return failure if any of the graphs $G_{v_i,d,\epsilon \log n}$ is neither unicyclic nor a tree. The number of steps the algorithm needs in the case of failure is at most equal to the number of steps that it will need in the case of a nonfail. Thus, time complexity of the nonfailing execution is an upper bound of the time complexity of the algorithm.

The algorithm needs $O(n)$ steps to create the graph $G_{v_i,d,\epsilon \log n}$ at the i-th iteration of its for loop. The graph $G_{v_i,d,\epsilon \log n}$ can be created by using any traversal algorithm, e.g. depth first search. This time bound follows by the fact that the number of vertices and the number of vertices in $G_{v_i,d,\epsilon \log n}$ are upper bounded by the number of vertices and the edges of the input graph. Using the Chernoff bounds (see [6]) it is direct to show that w.h.p. the number of edges in an instance of $G_{n,d/n}$ is $O(n)$.

Implementing a colouring of v_i according to $\tilde{\mu}_i(X_{v_i}|C(A_i \cap V_i))$ is equivalent to generating a random list colouring of $G_{v_i,d,\epsilon \log n}$ and keeping only the colour assignment of v_i from this colouring. In the list colouring problem every vertex $u \in V_i$ has a set $L(u)$ of valid colours, where $L(u) \subseteq [\mathcal{S}]$ and u can only receive a colour in $L(u)$. As argued in [4], for a tree on l vertices we can exactly compute the number of k-colourings in time $l \cdot k$. Therefore we can also generate a random list colouring of the tree. Also, for a unicyclic component we can simply consider all the $\leq k^2$ colourings of the endpoints of the extra edge and then recurse the remaining tree. In essence, the time we need to colour v_i according to $\tilde{\mu}_i(X_{v_i}|C(A_i \cap V_i))$ is at most $O(n)$. The theorem follows by noting that we need to colour n vertices. ◇

Acknowledgements

We wish to thank D. Achlioptas for comments on our work and also E. Mossel and A. Sly for the exchange of comments about our and their proving methodologies. We wish to thank previous anonymous referees for comments that allowed us to structure the proof arguments better.

References

[1] D. Achlioptas, A. Naor, The two possible values of the chromatic number of a random graph, in: Proc of 36th Annual ACM STOC, 2004, pp. 587–593.
[2] D. Aldous, Random walks on finite groups and rapidly mixing Markov chains, in: Séminaire de Probabilités XVII 1981/82, Springer-Verlag, Berlin, 1983, pp. 243–297.
[3] N. Alon, J.H. Spencer, The Probabilistic Method, 2nd edition, Wiley & Sons, Inc., 2000.
[4] M. Dyer, A. Flaxman, A.M. Frieze, E. Vigoda, Randomly coloring sparse random graphs with fewer colors than the maximum degree, Random Structures and Algorithms 29 (2006) 450–465.
[5] C. Efthymiou, P.G. Spirakis, Randomly colouring sparse random graphs using a constant number of colours, DELIS-TR-0501. http://delis.upb.de/docs/tr-abst-0501.html.
[6] S. Janson, T. Luczak, A. Ruciński, Random Graphs, Wiley & Sons Inc., 2000.
[7] M.R. Jerrum, L.G. Valiant, V.V. Vazirani, Random generation of combinatorial strucutres from a uniform distribution, Theoretical Computer Science 43 (1986) 169–188.
[8] M. Jerrum, A very simple algorithm for estimating the number of k-colourings of a low-degree graph, Random Structures and Algorithms 7 (1995) 157–165.
[9] M. Molloy, Glauber dynamics on colorings of a graph with high girth and maximum degree, SIAM Journal on Computing 33 (2004) 721–737.
[10] A. Montanari, D. Shah, Counting good truth assignments for random k-SAT formulae, in: Proc. the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, pp. 1255–1264.
[11] E. Mossel, A. Sly, Gibbs rapidly samples colourings of $g_{n,d/n}$, personal communication. Also in http://arxiv.org/abs/0707.3241.
[12] S. Ross, Stochastic Processes, 2nd edition, John Wiley and Sons, Inc., 1995.
[13] L. Valiant, The complexity of enumeration and reliability problems, SIAM Journal on Computing 8 (1979) 410–421.
[14] D. Weitz, Counting independent sets up to the tree threshold, in: Proc of 38th ACM Symposium on Theory of Computing, STOC 06, 2006, pp. 140–149.
[15] M. Wainwright, M. Jordan, Graphical models, exponental families and variational inference, Technical Report 649, Dept. of Statistics U.C. Berkeley 2003.