Deterministic Counting of Graph Colourings Using Sequences of Subgraphs

Charilaos Efthymiou
Goethe University, Mathematics Institute, Frankfurt 60054, Germarny
efthymiou@gmail.com

November 15, 2018

Abstract

In this paper we propose a deterministic algorithm for approximately counting the k-colourings of sparse random graphs $G(n,d/n)$. In particular, our algorithm computes in polynomial time a $(1 \pm n^{-\Omega(1)})$-approximation of the logarithm of the number of k-colourings of $G(n,d/n)$ for $k \geq (2+\epsilon)d$ with high probability over the graph instances.

Our algorithm is related to the algorithms of A. Bandyopadhyay et al. in SODA '06, and A. Montanari et al. in SODA '06, i.e. it uses *spatial correlation decay* to compute *deterministically* marginals of *Gibbs distribution*. We develop a scheme whose accuracy depends on *non-reconstruction* of the colourings of $G(n,d/n)$, rather than *uniqueness* that are required in previous works. This leaves open the possibility for our schema to be sufficiently accurate even for $k < d$.

The set up for establishing correlation decay is as follows: Given $G(n,d/n)$, we alter the graph structure in some specific region Λ of the graph by deleting edges between vertices of Λ. Then we show that the effect of this change on the marginals of Gibbs distribution, diminishes as we move away from Λ. Our approach is novel and suggests a new context for the study of deterministic counting algorithms.

1 Introduction

For a graph $G=(V,E)$ and a positive integer k, a proper k-colouring is an assignment $\sigma : V \to [k]$ (we use $[k]$ to denote $\{1,\dots,k\}$), such that adjacent vertices receive different members of $[k]$, i.e. different "colours". Here we focus on the problem of *counting* the k-colourings of G. In particular, we consider the cases where the underlying graph is an instance of Erdős-Rényi random graph $G(n,p)$, where $p=d/n$ and d is 'large' but remains bounded as $n \to \infty$. We say that an event occurs *with high probability (w.h.p.)* if the probability of the event to occur tends to 1 as $n \to \infty$.

Usually, a counting problem is reduced to computing marginal probabilities of *Gibbs distribution*, see [19]. Typically, we estimate these marginals by using a *sampling* algorithm. The most powerful method for sampling is the Markov Chain Monte Carlo (MCMC). There the main technical challenge is to establish that the underlying Markov chain mixes in polynomial time (see [18, 17]). The MCMC method gives *probabilistic* approximation guarantees.

Recently, new approaches were proposed for *deterministic* counting algorithms in [3] and [29]. The work in [3] is for counting colourings and independent sets, while [29] is for independent sets. These new approaches link the correlation decay to *computing efficiently* marginals of Gibbs distributions. The two algorithms in [3, 29] suggest two different approaches for computing marginals. The one in [3] applies mainly to locally tree graphs. Spatial correlation decay is exploited so as to restrict the computations of marginals and consider only small areas of the graph. The accuracy of the computations there relies on establishing the so-called *uniqueness conditions* on trees. On the other hand, the algorithm in [29] applies to a wider family of graphs, i.e. not necessarily locally treelike ones. It uses a more elaborate technique which somehow handles the existence of cycles in the computation of marginals, mainly by fixing the spins of certain sites appropriately. The approximation guarantees for the second algorithm are stronger than those of the first one. However, the stronger results do not come for free. The spatial mixing assumptions there are stronger, e.g. for the case of counting independent sets it requires *strong spatial mixing conditions*.

Our approach for computing Gibbs marginals is closer to [3] as w.h.p. the instance of $G(n, d/n)$ is locally tree like. However, this is not just an extension of [3] to random graphs. First we express the bounds for k in terms of the expected degree of the graph, rather than the maximum degree which is the case in [3]. Furthermore, we relate the computation of Gibbs marginals to weaker notions of spatial mixing, namely the so-called *non reconstruction conditions*. Compared to Gibbs uniqueness condition, which is required in [3], non-reconstruction is weaker and holds for a wider range of k. This leaves open the possibility for our schema to be sufficiently accurate for counting k-colourings of $G(n, d/n)$ even for $k < d$, i.e. when uniqueness condition is not expected to hold.

Further Motivation. Apart from its use for counting algorithms, the problem of computing efficiently good approximations of Gibbs marginals is a very interesting problem on its own. It is related to the empirical success of heuristics suggested by statistical physicists such as *Belief Propagation* and *Survey Propagation* (see e.g. [20]). In theoretical computer science, these heuristics are studied in the context of finding solution of random instances of Constraint Satisfaction Problems, e.g. random graph colouring, random k-SAT, etc. Similar ideas for computing marginals were also suggested in *coding theory* and *artificial intelligence* (see in [21]).

Related Work. Algorithms that follow a similar approach as the one in [3], appear in [23, 10]. The one in [23] is for computing Gibbs marginals for random instances of k-SAT. The one in [10] is for random colouring of $G(n, d/n)$. The algorithm in [10] does not compute the log partition function, however, it can be altered so as to do so. Then, it is not hard to show that it requires at least $d^{7/2}$ colours.

On the other hand, counting algorithms as the one in [29] give better polynomial time approximations, compared to the ones referred in the previous paragraph. However, they require stronger correlation decay conditions. Attempts to establish such strong conditions were successful for two spin cases, e.g. independent sets, matchings, Ising spins (see [29, 4, 24]]). For the multi-spin cases, such as colourings, things seem harder. The best algorithm of this category for counting k-colourings requires $k > 2.8\Delta$ and girth at least 4 (see [11]), where Δ is the maximum degree of the underlying graph.

The author of this work, in a subsequent paper [9], uses some of the ideas that appear here in an algorithm for approximate random colouring $G(n, d/n)$. The algorithm there yields similar results as here but the approximation guarantees are probabilistic ones, i.e. the same as the Monte Carlo algorithms.

1.1 Results

Let $Z(G, k)$ denote the number of k-colourings of the graph G. In statistic physics literature the quantity $Z(G, k)$ is also known as the *partition function*. Our algorithm computes an approximation for the log-partition function $\log Z(G, k)$.

Definition 1.1. *Ψ is defined to be an ϵ-approximation of the log-partition function $\log Z(G, k)$ if*

$$(1-\epsilon)\frac{\log Z(G,k)}{n} \leq \Psi \leq (1+\epsilon)\frac{\log Z(G,k)}{n}.$$

The results of our work are the following ones:

Theorem 1.1. *Let $\epsilon > 0$ be a fixed number and let d be sufficiently large. For $k \geq (2+\epsilon)d$ and with probability at least $1 - n^{-a}$, over the graph instances, our algorithm computes an n^{-b}-approximation of $\log Z(G_{n,d/n}, k)$, in time $O(n^s)$, where a, b and s are positive real numbers which depend on k.*

Roughly speaking the above theorem implies that for typical instances of $G(n, d/n)$ and $k \geq (2+\epsilon)d$ our algorithm is able to compute Gibbs marginals of the k-colourings of $G(n, d/n)$ within error $o(n^{-c})$, where $c > 0$ is fixed. Furthermore, the fact that the Gibbs distribution of k-colourings is symmetric and the fact that w.h.p. all but a vanishing fraction of the edges in $G(n, d/n)$ do not belong to cycles shorter than $\Theta(\ln n)$ implies the following result.

Corollary 1.1. *For sufficiently large d and $k \geq (2+\epsilon)d$, w.h.p. it holds that*

$$\left| \frac{\log Z(G_{n,d/n}, k)}{n} - \left(\log k + \frac{d}{2} \cdot \log\left(1 - \frac{1}{k}\right)\right)\right| \leq n^{-c},$$

for fixed $c > 0$.

Observe that the concentration result in Corollary 1.1, for the number of k-colouring of $G(n, d/n)$, is derived by using correlation decay arguments. In the literature of random structures such results are typically derived by using the so-called "Second moment method". A less accurate result can be derived from the work of Achlioptas and Naor in [1] with some extra work, i.e. the error there is $O(\log^{-1} n)$.

Finally, a related question and somehow a natural one is whether we can distinguish efficiently the instances of $G(n, d/n)$ that have their log-partition function concentrated. That is, for a sufficiently large function $h(n, d, k)$ we can answer whether a given instance $G(n, d/n)$ is such that

$$\left| \frac{\log Z(G(n, d/n), k)}{n} - \left(\log k + \frac{d}{2} \cdot \log\left(1 - \frac{1}{k}\right)\right)\right| \leq h(n, k, d),$$

or not. This goes beyond what we can get from the second moment method, as the later uses non-constructive arguments. We show that such distinction of instances is possible. The reason is that our arguments for correlation decay are tightly related to the degrees of vertices. That is, *examining the degrees of the vertices* we can infer whether the number of colourings of $G(n, d/n)$ is concentrated.

Let $S(n, d)$ denote the set of graphs on n vertices which have the following properties: Their number of edges is at most $3dn/4$. There are at most $n^{0.3}$ cycles, each of them, of length at most $\frac{\log n}{10 \log d}$. Finally, for each vertex v in the graph, the induced subgraph that contains v and all vertices within distance $\frac{\log n}{4\log(e^2 d/2)}$ is either tree or a unicyclic graph. In the following result, we show that for the graphs in $S(n, d/n)$ it is possible to verify whether the log-partition function is concentrated or not.

Corollary 1.2. *Let $\epsilon > 0$ be a fixed number and let d be sufficiently large. For $k \geq (2+\epsilon)d$, there exists a set of graphs $S(n, d)$ such that the following holds: For any sufficiently large real function $h(n, d, k) \geq n^{-O(1)}$ it can be verified in polynomial time whether the property*

$$\left| \frac{\log Z(G, k)}{n} - \left(\log k + \frac{d}{2} \cdot \log\left(1 - \frac{1}{k}\right)\right)\right| \leq h(n, k, d). \tag{1}$$

holds or not, for any $G \in S(n, d)$. Furthermore, $Pr[G_{n,d/n} \in S(n, d)] = 1 - n^{-0.1}$ and deciding whether $G_{n,d/n} \in S(n, d)$ can be made in polynomial time.

1.2 Contribution

We could partition the contribution of our work into two parts. The first part includes a new approximation-schema for computing deterministically Gibbs marginals. In the second part we present the tool for bounding correlation decay quantities that arise in the schema.

Approximating Gibbs Marginals. The problem of counting k-colourings of a graph $G = (V, E)$ reduces to the problem of estimating Gibbs marginals which can be formulated as follows:

Problem 1. Consider the graph $G = (V, E)$ and let $\mu(\cdot)$ denote the Gibbs distribution over the proper k-colourings of G. For the small (fixed sized) set of vertices $\Lambda \subset V$ and for $\sigma_\Lambda \in [k]^\Lambda$, compute the probability $\mu(\sigma_\Lambda)$.

In the general case computing $\mu(\sigma_\Lambda)$ exactly requires superpolynomial time. So the focus is on approximating it. One possible approach for computing an approximation of the marginal in Problem 1 was suggested in [3] for locally tree graphs. Roughly speaking the idea can be described as follows: The Gibbs marginal on Λ can be expressed as a convex combination of boundary conditions on $L_{t,\Lambda}$, the vertices at distance t from Λ, as follows

$$\mu(\sigma_\Lambda) = \sum_{\tau \in [k]^{L_{t,\Lambda}}} \mu(\sigma_\Lambda | \tau)\mu(\tau). \tag{2}$$

Pick t such that we can compute in polynomial time each of the marginals $\mu(\sigma_\Lambda|\tau)$. The problem, then, reduces to the not easier task of computing the coefficients $\mu(\tau)$. The authors in [3] noticed that the problem of estimating these coefficients somehow "degenerates" if k is so large that the marginals $\mu(\sigma_\Lambda|\tau)$ and $\mu(\sigma_\Lambda|\tau')$ are sufficiently close to each other, for any $\tau, \tau' \in [k]^{L_{t,\Lambda}}$ in the support of μ. In this case, the convexity implies that $\mu(\sigma_\Lambda)$ is sufficiently close to any of the conditional marginals in the r.h.s. of (2). Using this observation and the fact that we have chosen t such that the conditional marginals can be computed in polynomial time, it is direct that the above schema gives in polynomial time an approximation of $\mu(\sigma_\Lambda)$.

We should remark that the conditional marginals above are close to each other if a certain kind of independence hold, between the colourings of Λ and the colourings of $L_{t,\Lambda}$. Establishing such a kind of independence is related to what is known in statistical physics as establishing "Dobrushin Uniqueness Condition" (see [12]).

Our approach, here, is in a similar spirit. However, it amounts to substituting the coefficients $\mu(\tau)$ with new, different, ones. The aim is not to bypass the estimation of coefficients but somehow to *approximate* them. So instead of G we consider the graph $G_{t,\Lambda}$, the induced subgraph of G that contains the set Λ and all its neighbours within graph distance t. We denote with $\hat{\mu}(\sigma_\Lambda)$ the new Gibbs marginal of the event σ_Λ in the k-colourings of $G_{t,\Lambda}$. We will use $\hat{\mu}(\sigma_\Lambda)$ to approximate $\mu(\sigma_\Lambda)$. Note that we have chosen t so as the computation of $\hat{\mu}(\sigma_\Lambda)$ can be carried out efficiently. Writing the corresponding of (2) for the graph $G_{t,\Lambda}$ we get that

$$\hat{\mu}(\sigma_\Lambda) = \sum_{\tau \in [k]^{L_t(\Lambda)}} \hat{\mu}(\sigma_\Lambda|\tau)\hat{\mu}(\tau).$$

Remark 1. Someone could use uniqueness condition here as well, i.e. work as in [3]. However, here we make a more detailed comparison of $\hat{\mu}(\sigma_\Lambda)$ and $\mu(\sigma_\Lambda)$. As a matter of fact, our analysis gives rise to non-reconstruction spatial mixing conditions.

The key observation to compare $\hat{\mu}$ and μ is the following one: The distribution $\hat{\mu}(\cdot)$ can be seen as being induced by the deletion of the edges that connect the neighbourhood $G_{t,\Lambda}$ with the rest of the graph G. We require that the deletion of these edges does not have great effect on the marginals on Λ. It turns out that this is equivalent to requiring *non-reconstructibility* condition [1] with (sufficiently fast) exponential decay. That is, let G' be either G (the graph in Problem 1) or any of its subgraph. Let μ' be the Gibbs

[1] Non-reconstructibility is equivalent to extremality of Gibbs measure for infinite graphs, see e.g. [12].

distribution of the colourings of G'. Then, *non-reconstructibility* condition with exponential decay can be expressed as follows:

$$\max_{\mathcal{C}\in[k]^x} ||\mu'(\cdot) - \mu'(\cdot|\mathcal{C})||_{L_{x,t}} \leq \exp(-at), \tag{3}$$

where x is a vertex in G', $L_{t,x}$ contains all the vertices which are at distance t from x and $\alpha > 0$ is a fixed number.

For the distributions ν_a, ν_b on $[k]^V$, we let $||\nu_a - \nu_b||$ denote their *total variation distance*, i.e.

$$||\nu_a - \nu_b|| = \max_{\Omega'\subseteq[k]^V} |\nu_a(\Omega') - \nu_b(\Omega')|. \tag{4}$$

For $\Lambda \subseteq V$ let $||\nu_a - \nu_b||_\Lambda$ denote the total variation distance between the projections of ν_a and ν_b on $[k]^\Lambda$.

Bounds for Spatial Correlation Decay. We complement the new approach for estimating Gibbs marginals, by providing a general tool for bounding correlation decay conditions as in (3). We bound the correlation between some vertex x and the vertices at distance t from x by studying the probability of the following event: Choose u.a.r. a k-colouring of G'. Let ρ be the probability that there are two colour classes that specify a *connected* subgraph of G' that contains both x and some vertices at distance t. Then we show that $\max_{\mathcal{C}\in[k]^x} ||\mu'(\cdot) - \mu'(\cdot|\mathcal{C})||_{L_{x,t}} \leq \rho$.

We derive bounds for the quantity ρ by using the well-known technique from statistical physics called "disagreement percolation" coupling construction [6]. It turns out that using the disagreement percolation we express the decay of correlation as in (3) in terms of percolation-probabilities on the graph. Our technique is general and simple, e.g. there is no need for restrictions on the graph structure which was the case in [3, 10, 23]. Furthermore, it allows expressing the corresponding bounds in terms of the degree of each vertex, not the maximum degree.

Remark 2. "Disagreement Percolation" has been used for bounding different kinds of correlation decay in works for MCMC sampling colouring, e.g. [14, 7]. Also, disagreement percolation appears (implicitly) in [5] as part of a more general technique for showing non-reconstruction for colourings on trees. Our setting here is more general than [5] as it considers graphs with cycles. i.e. there are technical issues that need to be addressed.

Remark 3. For the sparse random graphs with bounded expected degree d there is a work by Montanari et al. in [22] that shows non-reconstructibility for k smaller than what we derive here. Unfortunately, we cannot use this result here, mainly, because it does not imply that the corresponding spatial mixing conditions are monotone in the graph structure. Note that if we could use the non-reconstructibility bounds from [22], then our results for counting would be even better.

1.3 Structure of the paper

The rest of the paper is organized as follows: In Section 2 we present some basic concepts and describe the counting to marginal estimation reduction. In Section 3 we give a general description of our counting algorithm and relate its accuracy with certain kind of spatial correlation decay conditions. Then, we provide the results which are used for bounding spatial correlation decay (in Section 3.2).

In Section 4 we discuss the technical details for applying the counting algorithm on $G_{n,d/n}$. We prove Theorem 1.1, Corollary 1.1 and Corollary 1.2. In Section 5 we prove the results that appear in Section 3.2, for bounding spatial correlation decay. Finally, in Section 6 we provide the proofs of some technical results we use.

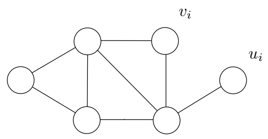

Figure 1: Graph G_i.

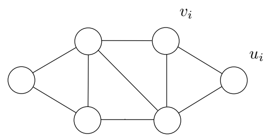

Figure 2: Graph G_{i+1}.

2 Basics and Problem Formulation

Our algorithm is studied in the context of finite spin-systems, a concept that originates in statistical physics. In particular, we use the finite *colouring model*.

The **Finite Colouring Model** with underlying graph $G = (V, E)$ that uses k colours is specified by a set of "*sites*", which correspond to the vertices of G, a set of "*spins*", i.e. the set $[k]$, and a symmetric function $U : [k] \times [k] \to \{0, 1\}$ such that for $i, j \in [k]$

$$U(i,j) = \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

We always assume that k is such that $Z(G, k) \neq \emptyset$.

A *configuration* $\sigma \in [k]^V$ of the system assigns each vertex ("site") $x \in V$ the colour ("spin value") $\sigma_x \in [k]$. The probability to find the system in configuration σ is determined by the *Gibbs distribution*, which is defined as

$$\mu(\sigma) = \frac{\prod_{\{x,y\}\in E} U(\sigma_x, \sigma_y)}{Z(G,k)}.$$

It is direct that the Gibbs distribution corresponds to the uniform distribution over the set of k-colouring of the underlying graph G. A *boundary condition* corresponds to fixing the colour assignment of a specific "*boundary*" vertex set of G.

Another concept we will need is that of the *sequence of subgraphs*.

Definition 2.1 (Sequence of subgraphs)**.** *For the graph $G = (V, E)$, let $\mathcal{G}(G) = \{G_i = (V, E_i)\}_{i=0}^{r}$ denote a sequence of subgraphs of G which has the following properties:*

- *G_0 is a spanning subgraph of G*
- *$E_i \subset E_{i+1}$ for $0 \leq i < r$ and $E_r = E$*
- *the term G_{i+1} compared to G_i has an additional edge, the edge $\Psi_i = \{v_i, u_i\}$.*

When we refer to $\mathcal{G}(G)$ we specify the graph G_0 while we, usually, assume that there is some arbitrary rule which gives the terms $G_1, \ldots, G_r$. In Figures 1 and 2 there is an example of two consecutive terms of a sequence $\mathcal{G}(G)$, for some graph G. Observe that in G_i the vertices v_i and u_i are not adjacent, while in G_{i+1} we add the edge $\Psi_i = \{v_i, u_i\}$.

Lemma 2.1. *For the graph $G = (V, E)$ consider a sequence of subgraphs $\mathcal{G}(G)$ where G_0 is edgeless. Let X_i be a random colouring of $G_i \in \mathcal{G}(G)$. For some integer $k > 0$, we have that*

$$|Z(G,k)| = k^n \cdot \prod_{i=1}^{|E|-1} Pr[X_i(v_i) \neq X_i(u_i)],$$

where the vertices v_i and u_i are incident to Ψ_i.

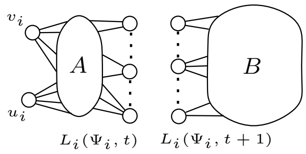

Figure 3: Graph G_{i0}.

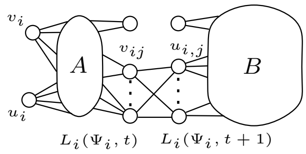

Figure 4: Graph $G_{i,j+1}$.

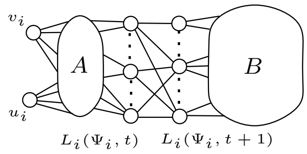

Figure 5: Graph G_{i,r_i}.

The proof of the above lemma is standard and can be found in various places (e.g. [19, 8, 16]), for completeness we present it in Section 6.3.

We close this section with some additional notation. For $\Lambda \subseteq V$ and some integer $t > 0$, we let $L(\Lambda, t)$ denote the set of vertices at graph distance *exactly* t from Λ. Also, we let $B(\Lambda, t)$ denote the set of vertices *within* graph distance t from Λ.

3 Counting Schema

For clarity reasons, we present the counting schema by assuming that we are given a fixed graph $G = (V, E)$ and some integer k such that $Z(G, k) > 0$.

The schema is based on computing Gibbs marginals as it is described in Lemma 2.1. That is, given G, we consider a sequence of subgraphs $\mathcal{G}(G) = G_0, \dots, G_r$ with G_0 being edgeless. For each $G_i \in \mathcal{G}(G)$ let X_i be a random colouring. In our schema we compute an *approximation* of each probability term $Pr[X_i(v_i) \neq X_i(u_i)]$ by working as follows: We consider a new sequence of subgraphs $\mathcal{G}(G_i) = G_{i,0}, \dots, G_{i,r_i}$ defined as follows: G_{i,r_i}, is the graph G_i while $G_{i,0}$ is derived from G_i by removing all the edges between the sets $L(\Psi_i, t)$ and $L(\Psi_i, t+1)$[2], where $t > 0$ is some appropriate integer. We consider Y_i a random colouring of the graph $G_{i,0} \in \mathcal{G}(G_i)$. Our schema approximates $Pr[X(v_i) \neq X(u_i)]$ with $Pr[Y_i(v_i) \neq Y_i(u_i)]$.

Observe that the computation of $Pr[Y_i(v_i) \neq Y_i(u_i)]$ depends on the induced subgraph of G_i which contains only vertices within graph distance t from $\Psi_i = \{v_i, u_i\}$. Taking sufficiently small t it makes it possible to compute $Pr[Y_i(v_i) \neq Y_i(u_i)]$ in polynomial time.

Figures 3, 4 and 5 illustrate some members of $\mathcal{G}(G_i)$. That is, Figure 3 shows the first term of the sequence. Figure 4 shows the graph $G_{i,j+1}$, i.e. the edge $\Psi_{i,j} = \{u_{i,j}, v_{u,i}\}$ has just been inserted. In Figure 5 we have the final term of $\mathcal{G}(G_i)$, the graph G_{i,r_i}.

In what follows we provide the pseudocode of the counting algorithm.

Counting Schema

Input: G, k, t.
Set $\mathcal{Z} = k^n$.
Compute $\mathcal{G}(G) = \{G_0, \dots, G_r\}$.
For $0 \leq i \leq r-1$ do

- Compute $\mathcal{G}(G_i)$.
- Compute the exact value of $Pr[Y_i(v_i) \neq Y_i(u_i)]$.
- Set $\mathcal{Z} = \mathcal{Z} \cdot Pr[Y_i(v_i) \neq Y_i(u_i)]$.

End For.
Output: $\log(\mathcal{Z})/n$.

[2]Both $L(\Psi_i, t)$ and $L(\Psi_i, t+1)$ are considered w.r.t. graph G_i.

Two natural questions arise for the counting algorithm. The first one is its *accuracy*, i.e. how close $\frac{1}{n}\log \mathcal{Z}$ and $\frac{1}{n}\log Z(G,k)$ are. The second one is about the *time complexity*.

As far as the time complexity is regarded, typically, the execution time is dominated by the computations for $Pr[Y_i(v_i) \neq Y_i(u_i)]$. Let us remark, here, that there is no standard way of computing $Pr[Y_i(v_i) \neq Y_i(u_i)]$. In the next section where we study the application of the above schema on $G(n, d/n)$ we choose t such that the computation of the marginal $Pr[Y_i(v_i) \neq Y_i(u_i)]$ can be carried out efficiently by using a *dynamic programming algorithm*.

As far as the accuracy is concerned we have the following results.

Proposition 3.1. *For the counting schema it holds that*

$$\frac{1}{n}|\log \mathcal{Z} - \log Z(G,k)| \leq \frac{2}{n}\sum_{i=0}^{r-1} \frac{|Pr[X_i(v_i) \neq X_i(u_i)] - Pr[Y_i(v_i) \neq Y_i(u_i)]|}{Pr[X_i(v_i) \neq X_i(u_i)]},$$

when each of the summands on the r.h.s. is sufficiently small.

The proof of Proposition 3.1 appears in Section 6.1.

So as to show that the estimation $\log \mathcal{Z}$ is accurate, we work as follows: We derive a constant lower bound for $Pr[X_i(v_i) \neq X_i(u_i)]$, which is used to for the denominator in Proposition 3.1. Then, we show that $Pr[X_i(v_i) \neq X_i(u_i)]$ and $Pr[Y_i(v_i) \neq Y_i(u_i)]$ are asymptotically equal. There, we use the following proposition.

Proposition 3.2. *For* $0 \leq i \leq r-1$ *it holds that*

$$\begin{aligned}&|Pr[X_i(v_i) \neq X_i(u_i)] - Pr[Y_i(v_i) \neq Y_i(u_i)]| \leq \\ &\quad\leq \sum_{j=0}^{r_i-1} C_{ij} \max_{\sigma,\tau,\in\Omega(G_{ij},k)} \left\{||\mu_{i,j}(\cdot|\sigma_{v_{ij}}) - \mu_{i,j}(\cdot|\tau_{v_{ij}})||_{\Psi_i\cup\{u_{ij}\}} + ||\mu_{i,j}(\cdot|\sigma_{v_{ij}}) - \mu_{i,j}(\cdot|\tau_{v_{ij}})||_{\{u_{ij}\}}\right\},\end{aligned}$$

where $C_{ij} = \max_{s,t\in[k]}\left\{(Pr[X_{i,j}(u_{i,j}) = s, X_{i,j}(v_{i,j}) = t])^{-2}\right\}$ *and* r_i *is the number of terms in the sequence* $\mathcal{G}(G_i)$.

The proof of Proposition 3.2 is given in Section 6.2.

3.1 Remarks on the Spatial Conditions

It is interesting to discuss the implications of the spatial mixing conditions required by Proposition 3.1 and Proposition 3.2. If every C_{ij} in Proposition 3.2 is a sufficiently small constant, which will be the case here, then the spatial mixing condition can be summarized as follows:

$$\frac{1}{n}|\log \mathcal{Z} - \log Z(G,k)| \leq f(G,t)\cdot \max_{i,j,x,\sigma,\tau} ||\mu_{ij}(\cdot|\sigma_x) - \mu_{ij}(\cdot|\tau_x)||_{\Lambda},$$

where $f(G,t)$ is a quantity that grows linearly with the number of terms in both sequences $\mathcal{G}(G)$ and $\mathcal{G}(G_i)$ and $\Lambda \subset V$ is an appropriate defined region in G. Then, a sufficient condition for the counting schema to be accurate is that, for every $0 \leq i \leq r$ and $0 \leq j \leq r_i$ we have

$$\max_{x\in V} \max_{\sigma_x,\tau_x\in[k]^{\{x\}}} ||\mu_{ij}(\cdot|\sigma_x) - \mu_{ij}(\cdot|\tau_x)||_{L(\{x\},t)} \leq \exp\left(-a\cdot t\right) \tag{5}$$

for sufficiently large $a > 0$. Another expression for the condition in (5) can be derived by using the following (standard) lemma.

Lemma 3.1. *For any graph $G = (V, E)$ and k, let μ be the Gibbs distribution of its k-colourings. For every $x \in V$ and $\Lambda \subseteq V$ it holds*

$$\max_{\sigma_x,\tau_x \in [k]^{\{x\}}} ||\mu(\cdot|\sigma_x) - \mu(\cdot|\tau_x)||_\Lambda \leq 2k \cdot \sum_{A \in [k]^\Lambda} \mu(A) \cdot ||\mu(\cdot|A) - \mu(\cdot)||_x.$$

For a proof Lemma 3.1 see in Section 6.4.

In the light of the above lemma and for k constant the condition in (5) is equivalent to the following one: For $0 \leq i \leq r$ and $0 \leq i \leq r_i$

$$\max_{x \in V} \max_{\sigma_x,\tau_x \in [k]^{\{x\}}} \sum_{A \in [k]^{L(\{x\},t)}} \mu_{ij}(A) \cdot ||\mu_{ij}(\cdot|A) - \mu_{ij}(\cdot)||_x \leq \exp(-a' \cdot t), \tag{6}$$

for appropriate $a' > 0$. What the condition in (6) implies is that a "typical" colouring of $L(\{x\}, t)$ in G_{ij} should have small impact on the Gibbs marginal on x.

3.2 Bounds for Spatial Correlation decay

In this section, we provide the method that we use to derive an upper bound for the quantities that express spatial correlation decay in Proposition 3.2, i.e. $||\mu_{ij}(\cdot|\sigma_x) - \mu_{ij}(\cdot|\tau_x)||_\Lambda$, for $x \in V$ and $\Lambda \subset V$. The derivation of these bounds are of independent interest from the discussion in the Section 3.1. The method is based on the well-known "*disagreement percolation*" coupling construction, from [6].

Consider a configuration space on the vertices of G such that each vertex $v \in V$ is set either *disagreeing* or *non-disagreeing*. In such a configuration, we call *path of disagreement* any simple path which has all its vertices disagreeing. Given an integer s and $w \in V$ we let $\mathcal{P}_{s,w}$ be the *product measure* under which each vertex $v \in V \backslash \{w\}$ of degree $\Delta(v) < s$ is disagreeing with probability $\frac{1}{s-\Delta(v)}$ and non-disagreeing with the remaining probability. If $s \leq \Delta(v)$, then v is disagreeing with probability 1. The vertex w is set disagreeing with probability 1, regardless of its degree. Using the above concepts we show the following result.

Theorem 3.1. *Consider the graph $G = (V, E)$, $v \in V$, $\Lambda \subseteq V$ and an integer $k > 0$. Let μ denote the Gibbs distribution of the k-colourings of G. Also, let $\mathcal{P}_{k,v}$ denote the product measure defined above. It holds that*

$$\max_{\sigma_v,\eta_v \in [k]^{\{v\}}} ||\mu(\cdot|\sigma_v) - \mu(\cdot|\eta_v)||_\Lambda \leq \mathcal{P}_{s,v}[\,\exists \textit{ path of disagreement connecting } \{v\} \textit{ and } \Lambda].$$

The proof of Theorem 3.1 is given in Section 5.

Roughly speaking, we bound $||\mu(\cdot|\sigma_v) - \mu(\cdot|\eta_v)||_\Lambda$, in Theorem 3.1, by working as follows: We use coupling, i.e. we couple X, Y two random colourings of G that assign the vertex x colour σ_v and η_v, respectively. Then, by Coupling Lemma [2] we have that

$$||\mu(\cdot|\sigma_v) - \mu(\cdot|\eta_v)||_\Lambda \leq Pr[X(\Lambda) \neq Y(\Lambda)].$$

The coupling of X, Y is done by specifying what Y is, given X. In particular, given X, we let G_X denote the maximal *connected* subgraph of G which contains the vertex v and vertices from the colour classes specified by σ_v and η_v in the colouring X. Then, we derive Y as follows: For every vertex $u \notin G_X$ it holds that $Y(u) = X(u)$. For $u \in G_X$ if $X(u) = \sigma_x$, then $Y(u) = \tau_x$ and the other way around[3]. In Figures 6 and 7 we illustrate this coupling, e.g. $\sigma_v =$"Blue" and $\eta_v =$"Green".

[3] I.e. if $X(u) = \tau_x$, then $Y(u) = \sigma_x$.

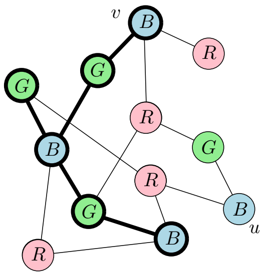

Figure 6: Colouring X.

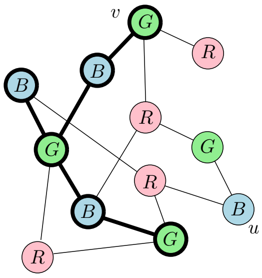

Figure 7: Colouring Y.

It is not hard to see that in the above coupling X, Y disagree only on the colour assignments for the vertices in G_X. That is

$$Pr[X(\Lambda) \neq Y(\Lambda)] = Pr[\exists \Lambda' \subseteq \Lambda : \Lambda' \subseteq G_X \text{ in the coupling}].$$

Of course, bounding the probability term on the r.h.s. of the inequality above is not a trivial task. However, we show that the above process (of getting G_X) is stochastically dominated by an independent process, i.e. disagreement percolation. That is, we show that

$$Pr[\exists \Lambda' \subseteq \Lambda : \Lambda' \subseteq G_X \text{ in the coupling}] \leq \mathcal{P}_{s,v}[\,\exists \text{ path of disagreement connecting } \{v\} \text{ and } \Lambda].$$

4 Application to $G(n, d/n)$

In this section we show Theorem 1.1, Corollary 1.1 and Corollary 1.2. For technical reasons, which we discuss later, we require the following sequence of subgraphs.

Sequence of subgraphs $\mathcal{G}(G_{n,d/n})$: Let r be the greatest index in $\mathcal{G}(G_{n,d/n})$, e.g. $\mathcal{G}(G_{n,d/n}) = G_0, \ldots, G_r$. The term G_0 is an edgeless graph. Let R be the set of all edges in $G_{n,d/n}$ that do not belong to a cycle of length smaller than $\frac{\log n}{10 \log d}$ but they are incident to some vertex that belongs to such a cycle. There is an index i_0 such that for every $i \geq i_0$, G_i differs from G_{i-1} in some edge from R while for $i < i_0$ no edge from the set R appears in G_i.

For $0 \leq i \leq r$ consider that the sequence of subgraphs $\mathcal{G}(G_i)$ defined as follows: $G_{i,0}$ is derived by G_i by deleting all the edges that connect the sets of vertices $L(\Psi_i, t)$ and $L(\Psi_i, t+1)$ where $t = \frac{\log n}{2 \log d}$.

Typically we are in the case where k, the number of colours, is smaller than the maximum degree of $G(n, d/n)$[4]. Then, there can be situations where $(C_{i,j})^{-1}$ (defined in Proposition 3.2) and $Pr[X_i(v_i) \neq X_i(u_i)]$ are very small. According to Proposition 3.2, this can increase the error dramatically. The analysis implies that these situations arise when the vertices that are involved, i.e. v_i, u_i, or v_{ij}, u_{ij}, have large degrees and belong to small cycles at the same time. It is easy to see that choosing $\mathcal{G}(n, d/n)$ as we describe above, we avoid such undesirable situations for any $i < i_0$. Furthermore, the terms $Pr[X_{i,0}(v_i) \neq X_{i,0}(u_i)]$ for $i \geq i_0$ are too few, i.e. $O(n^{0.3})$, and it turns out that each of them is bounded away from zero. This implies that their contribution to $\log(Z(G(n, d/n)))$ is negligible.

Setting the parameter $t = \frac{\log n}{2 \log d}$, the component in $G_{i,0}$ which contains $\{v_i, u_i\}$ is w.h.p. a tree with $O(\log n)$ extra edges, for every $0 \leq i < i_0$. This allows the computation of every Gibbs marginal in polynomial-time. To be more specific we work as follows:

[4]The maximum degree in $G_{n,d/n}$ is $\Theta\left(\frac{\log n}{\log \log n}\right)$ w.h.p. (see [15])

Computing Probabilities . The probability term $Pr[Y_i(v_i) \neq Y_i(u_i)]$, for $0 \leq i < i_0$, can be computed by using Dynamic Programming (D.P.). More specifically, using DP we can compute exactly the number of list colourings of a tree T. In the list colouring problem every vertex $v \in T$ has a set $List(v)$ of valid colours, where $List(v) \subseteq [k]$ and v only receives a colour in $List(v)$. For a tree on l vertices, using dynamic programming we can compute exactly the number of list colourings in time lk.

For $0 \leq i < i_0$, the connected component in $G_{i,0}$ that contains $\{v_i, u_i\}$ is a tree with at most $\Theta(\log n)$ extra edges w.h.p. For such component we can consider all the $k^{O(\log n)}$ colourings of the endpoints of the extra edges and for each of these colourings recurse on the remaining tree. Since in our case k is constant, $k^{O(\log n)} = n^{O(1)}$. It follows that the number of list colourings of the connected component, in $G_{i,0}$, that contains $\{v_i, u_i\}$ can be counted in polynomial time for every i. This is sufficient for computing $Pr[Y_i(v_i) \neq Y_i(u_i)]$ efficiently[5].

The pseudocode of the counting schema for the case of $G(n, d/n)$ follows.

Counting Schema $G(n, d/n)$

Input: $G(n, d/n)$, k

Compute the set of edges R.

If $|R| > n^{0.3}$, compute $\log(Z(G_{n,d/n}, k))$ by *exhaustive enumeration.*

Compute the sequence of subgraphs $\mathcal{G}(G_{n,d/n})$.

Set $\mathcal{Z} = 1$

For $0 < i < r - |R|$ do

- Compute the exact value of $Pr[Y_i(v_i) \neq Y_i(u_i)]$.
- Set $\mathcal{Z} = \mathcal{Z} \cdot Pr[Y_i(v_i) \neq Y_i(u_i)]$.

End for.

Set $\mathcal{Z} = \mathcal{Z} \cdot k^n$.

Output: $\log(\mathcal{Z})/n$.

Observe that, above, implicitly we set $Pr[Y_i(v_i) \neq Y_i(u_i)] = 1$ for $i \geq i_0$. It turns out that the error introduced by working this way is negligible. Theorem 1.1 follows as a corollary of the following two propositions.

Proposition 4.1. *Let $\epsilon > 0$ be a fixed number and let d be sufficiently large. For $k \geq (2+\epsilon)d$ the counting schema computes an n^{-b}-approximation of $\log Z(G(n, d/n), k)$, with probability at least $1 - n^{-a}$, over the graph instances and $a, b > 0$ depend on k.*

The proof of Proposition 4.1 appears in Section 4.1 and makes a heavy use of Theorem 3.1.

Proposition 4.2. *There are real constants $h, s > 0$ such that the time complexity for the counting schema to compute $\log Z(G(n, d/n), k)$ is $O(n^s)$, with probability at least $1 - n^{-h}$, over the graph instances.*

Proof: The theorem follows directly from the paragraph, "Computing Probabilities", above. $\Diamond$

4.1 Proof of Proposition 4.1

First we present a series of results that will be useful for the proof of Proposition 4.1. In all our results that follow we assume that $\epsilon > 0$ is a fixed number and $d > 0$ is sufficiently large, i.e. $d > d_0(\epsilon)$.

[5]A similar DP approach is also used in [7] and [10].

Proposition 4.3. *Consider the measure $\mathcal{P}_{k,x}$ w.r.t. $G(n, d/n)$, for $k \geq (2+\epsilon)d$ and some vertex x in the graph. For a set of vertices Ψ, let $D^{(l)}$ denote the number of paths of disagreement between x and Ψ, of length at least l, for any integer $l = O(\log n)$. Then, there exists a real $\gamma = \gamma(k) > 1$ such that*

$$Pr[D^{(l)} > 0] \leq \frac{8}{\epsilon} \cdot \frac{|\Psi|}{n} \gamma^{-l}, \tag{7}$$

where $|\Psi|$ is the cardinality of Ψ. The probability term above, is w.r.t $\mathcal{P}_{k,x}$ and the graph instances.

The proof of Proposition 4.3 appears in Section 4.2. Also, from the proof of Proposition 4.3 it is direct to deduce the following corollary.

Corollary 4.1. *The bound for the probability in (7) holds even if we remove an arbitrary set of edges of $G(n, d/n)$.*

The following lemma is standard. We denote by C_l the number of cycles of length at most l. Also, we remind the reader that the set R is the set of edges of $G(n, d/n)$ that do not belong to a cycle of length smaller than $\frac{\log n}{10 \log d}$ but they are incident to a vertex that belongs to such a cycle.

Lemma 4.1. *With probability at least $1 - n^{-0.19}$, the following holds: (A) $|R| \leq n^{0.3}$. (B) $C_l \leq n^{0.3}$, for $l = \frac{\log n}{10 \log d}$. (C) After removing the edges in R from $G_{n,d/n}$, each of the cycles of length less than $\frac{\log n}{10 \log d}$ becomes isolated from the rest of the graph.*

For completeness we present the proof of Lemma 4.1 in Section 6.5.

Lemma 4.2. *For $\mathcal{G}(G(n, d/n))$, $\mathcal{G}(G_i)$ as defined in Section 4 and for constant $k \geq (2+\epsilon)d$, the following holds:*

$$Pr[C_{i,j} < 2k^4, \text{ for } 0 \leq i < i_0, 0 \leq j \leq r_i] \geq 1 - n^{-\frac{\log \gamma}{11 \log d}}, \tag{8}$$

where C_{ij}, γ are defined in the statements of Proposition 3.2 and Proposition 4.3, respectively.

Proof: Let $X_{i,j}$ be a random colouring of $G_{i,j}$. We remind the reader that

$$C_{ij} = \max_{s,t \in [k]} \left\{ (Pr[X_{i,j}(u_{i,j}) = s, X_{i,j}(v_{i,j}) = t])^{-2} \right\}.$$

We show that $C_{i,j}$ is reasonably small by comparing $Pr[X_{i,j}(u_{i,j}) = s | X_{i,j}(v_{i,j}) = t]$ with $Pr[X_{i,j}(u_{i,j}) = s] = 1/k$ and by showing that these two probability terms do not differ much. In particular, we have

$$|Pr[X_{i,j}(u_{i,j}) = s | X_{i,j}(v_{i,j}) = t] - Pr[X_{i,j}(u_{i,j}) = s]| \leq \max_{\sigma, \eta \in [k]^{\{v_{i,j}\}}} ||\mu_{ij}(\cdot|\sigma) - \mu_{ij}(\cdot|\eta)||_{u_{ij}}. \tag{9}$$

Then, we show that with probability at least $1 - n^{-\frac{\log \gamma}{11 \log d}}$ for $0 \leq i < i_0$ and $0 \leq j \leq r_i$ it holds that

$$\max_{\sigma, \eta \in [k]^{\{v_{i,j}\}}} ||\mu_{ij}(\cdot|\sigma) - \mu_{ij}(\cdot|\eta)||_{u_{ij}} \leq \frac{1}{10k}. \tag{10}$$

Given the above, it is straightforward to verify (8) by using (9) and (10). Then, the lemma follows.

We are going to use Theorem 3.1 to prove (10). For a pair of adjacent vertices x, y in the graph let $D_{x,y}$ denote the number of paths of disagreement that start from x and end in y but they do not use the edge $\{x, y\}$. Also, we let $\varrho_{x,y} = \mathcal{P}_{k,x}[D_{x,y} > 0]$. Finally, given some integer $s > 1$ we let $D^{(s)}_{x,y}$ denote the number of paths of disagreement that start form x, end in y and their length is at least s. Similarly, let $\varrho^{(s)}_{x,w} = \mathcal{P}_{k,x}[D^{(s)}_{x,y} > 0]$.

Let $e = \{x, y\}$ be a random edge in $G(n, d/n)$ conditional that the shorter cycle that contains it is of length at least $\frac{\log n}{10 \log d}$. Let $e' = \{x', y'\}$ be a randomly chosen edge in $G(n, d/n)$. It holds that

$$E[\varrho_{x,y}] \leq \frac{1}{\psi} E[\varrho^{(l)}_{x',y'}], \tag{11}$$

where l denotes the distance between the vertices x and y. Also, ψ is the probability that a randomly chosen edge in $G(n, d/n)$ does not belong to a cycle shorter than $\frac{\log n}{10 \log d}$. It is straightforward to show that $\psi = 1 - o(1)$. Using Proposition 4.3 and the fact that $l \geq \frac{\log n}{10 \log d}$ we have that

$$E[\varrho^{(l)}_{x',y'}] \leq \frac{8}{\epsilon} n^{-\left(1+\frac{\log \gamma}{10 \log d}\right)}. \tag{12}$$

From (11) and (12) we get that $E[\varrho_{x,y}] \leq \frac{10}{\epsilon} n^{-\left(1+\frac{\log \gamma}{10 \log d}\right)}$. From Markov's inequality we get that

$$Pr\left[\varrho_{x,y} \geq \frac{1}{10k}\right] \leq \frac{100k}{\epsilon} n^{-\left(1+\frac{\log \gamma}{10 \log d}\right)}.$$

Let L be number of edges $\{x, y\}$ in $G(n, d/n)$ such that the shortest cycle that contains each of them is of length at least $\frac{\log n}{10 \log d}$ and $\varrho_{x,y} \geq \frac{1}{10k}$. Using the linearity of expectation, it is straightforward to show that $E[L] \leq \frac{60dk}{\epsilon} n^{-\frac{\log \gamma}{10 \log d}}$. Applying, Markov's inequality we get that

$$Pr[L > 0] \leq \frac{60dk}{\epsilon} n^{-\frac{\log \gamma}{10 \log d}}. \tag{13}$$

Observe that the probability for path between two vertices to be a path of disagreement is an increasing function of the degrees of its vertices (when k is fixed). From this observation and (13) we have that for every $v_{i,j}$ and $u_{i,j}$ it holds that $\varrho_{v_{i,j},u_{ij}} \leq 1/(10k)$ with probability at least $1 - \frac{60dk}{\epsilon} n^{-\frac{\log \gamma}{10 \log d}}$. The lemma follows by using Theorem 3.1, i.e. it holds that

$$\max_{\sigma,\eta \in [k]^{\{v_{i,j}\}}} ||\mu_{ij}(\cdot|\sigma) - \mu_{ij}(\cdot|\eta)||_{u_{ij}} \leq \mathcal{P}_{k,v_{i,j}}[D_{v_{i,j},u_{i,j}} > 0] = \varrho_{v_{i,j},u_{ij}}.$$

$\Diamond$

Lemma 4.3. *Let γ be as in the statement of Proposition 4.3. For $\mathcal{G}(G_{n,d/n})$ as defined in Section 4 and for $k \geq (2 + \epsilon)$ the following holds:*

- *Let I be the set such that $i \in I$, iff the edge $\{v_i, u_i\}$ does not belong to any cycle of length less than $\frac{\log n}{10 \log d}$. With probability at least $1 - n^{-\frac{\log \gamma}{22 \log d}}$ over the instances $G(n, d/n)$ it holds that*

$$\left| Pr[X_i(u_i) \neq X_i(v_i)] - \left(1 - \frac{1}{k}\right) \right| \leq n^{-\frac{\log \gamma}{21 \log d}}, \quad \forall i \in I. \tag{14}$$

- *Let I' be the set such that $i \in I'$, iff the edge $\{v_i, u_i\}$ belongs to cycle of length less than $\frac{\log n}{10 \log d}$. With probability at least $1 - n^{-0.19}$ over the instances $G(n, d/n)$ it holds that*

$$Pr[X_i(u_i) \neq X_i(v_i)] = \Theta(1).$$

Proof: First we consider the edges $\{v_i, u_i\}$ such that $i \in I$. There, we use the following fact.

$$\left| Pr[X_i(u_i) \neq X_i(v_i] - \left(1 - \frac{1}{k}\right)\right| \leq \max_{\sigma,\eta \in [k]^{\{v_i\}}} ||\mu_i(\cdot|\sigma) - \mu_i(\cdot|\eta)||_{u_i} \leq \mathcal{P}_{k,v_i}[D_{v_i,u_i} > 0],$$

where D_{v_i,u_i} is the number of paths of disagreement in $G(n, d/n)$ that connect v_i and u_i but they do not use the edge $\{v_i, u_i\}$.

As in the proof of Lemma 4.2, for the vertices x', y' we let $\varrho_{x',y'} = \mathcal{P}_{k,x'}[D_{x',y'} > 0]$. We work in the same manner as in the proof of Lemma 4.2 to get tail bounds for $\varrho_{x',y'}$, i.e. we get the following: For a random edge $\{x, y\}$ such that the shortest cycle that contains it is of length at least $\frac{\log n}{10 \log d}$, it holds that

$$Pr\left[\varrho_{x,y} \geq n^{-\frac{\log \gamma}{20 \log d}}\right] \leq \frac{10}{\epsilon} n^{-\left(1 + \frac{\log \gamma}{20 \log d}\right)}. \tag{15}$$

Let L be number of edges in $G(n, d/n)$ such that the shortest cycle that contains each of them is of length at least $\frac{\log n}{10 \log d}$ and $\varrho_{x,y} \geq n^{-\frac{\log \gamma}{20 \log d}}$. Using the linearity of expectation it is straightforward to show that $E[L] \leq \frac{6d}{\epsilon} n^{-\frac{\log \gamma}{20 \log d}}$. Applying, Markov's inequality we get that

$$Pr[L > 0] \leq \frac{6d}{\epsilon} n^{-\frac{\log \gamma}{20 \log d}}. \tag{16}$$

It is immediate that (14) holds.

In the latter case, we consider v_i and u_i which belong to small cycle, i.e. of length at most $\frac{\log n}{10 \log d}$. Such a pair of vertices appears in the schema only when we have removed from $G_{n,d/n}$ all the edges in R. By Lemma 4.1 we have that with probability at least $1 - n^{-0.19}$ the removal of the edges in R disconnects every small cycle from the rest of $G_{n,d/n}$. Thus, for the second case, where v_i, u_i belong to a small, isolated cycle, $Pr[X_i(u_i) \neq X_i(v_i]$ is trivially lower bounded by some constant, since $k \gg 2$. The lemma follows. ♢

Using Lemma 2.1 and the previous lemmas, in this section, we get the following corollary.

Corollary 4.2. *For $k \geq (2+\epsilon)d$, the log-partition function of the k-colourings of $G_{n,d/n}$ is $\Theta(n)$, w.h.p.*

We have all the lemmas we need to show Proposition 4.1.

Proof of Proposition 4.1: Let $\mathcal{D}$ be the event that " (a) $r \leq \rho = \frac{dn}{2}(1 + n^{-1/3})$, (b) $\max_i\{r_i\} \leq 10dn^{1/2} \log n$, (c) $|R| \leq n^{0.3}$, (d) $\min_i\{Pr[X_i(v_i) \neq X_i(u_i)]\} = \Theta(1)$, (e) $\max_{i,j}(C_{i,j}) \leq 2k^4$".

We remind the reader that we denote with r the number of terms in $\mathcal{G}(G(n, d/n))$, r_i the number of terms in $\mathcal{G}(G_i)$, for every $G_i \in \mathcal{G}(G(n, d/n))$.

Claim 4.1. *It holds that $Pr[\mathcal{D}] \geq 1 - n^{-\beta}$, for some fixed $\beta > 0$.*

Proof: From all the previous results in Section 4.1, it suffices to show that $\max_i\{r_i\} \leq 5dn^{1/2} \log n$ with sufficiently large probability.

Clearly, r_i is equal to the number of edges between $L\left(\Psi_i, \frac{\log n}{2 \log d}\right)$ and $L\left(\Psi_i, \frac{\log n}{2 \log d} + 1\right)$ in G_i. The number of vertices at distance $\frac{\log n}{2 \log d}$ from Ψ is dominated by a Galton-Watson tree of $\frac{\log n}{2 \log d}$ levels, with a number of offspring per individual distributed as in $\mathcal{B}(n, d/n)$ and the initial population being 2. With standard arguments (e.g. see Theorem 6 in [24]), it holds that with probability at least $1 - n^{-3}$, the number of vertices at level $\frac{\log n}{2 \log d}$ is at most $9n^{1/2} \log n$. Clearly r_i is at most the sum of degrees of these vertices. In turn, this sum is dominated by a sum of $9n^{1/2}\log n$ independent $\mathcal{B}(n,d/n)$. It is direct to derive that $r_i = 10dn^{1/2}\log n$ with probability at least $1-n^{-3}$, by using Chernoff bounds. The claim follows. $\Diamond$

By Proposition 3.1 we have that

$$E\left[\frac{1}{n}|\log \mathcal{Z} - \log Z(G(n,d/n))||\mathcal{D}\right] \leq \frac{2}{n}\sum_{i=0}^{\rho} E\left[\frac{|Pr[X_i(v_i)\neq X_i(u_i)] - Pr[X_{i,0}(v_i)\neq X_{i,0}(u_i)]|}{Pr[X_i(v_i)\neq X_i(u_i)]}|\mathcal{D}\right], \quad (17)$$

where the expectation is over the graph instances $G(n,d/n)$. Using Proposition 3.2, we have that

$$E\left[|\frac{Pr[X_i(v_i)\neq X_i(u_i)] - Pr[Y_i(v_i)\neq Y_i(u_i)]}{Pr[X_i(v_i)\neq X_i(u_i)]}||\mathcal{D}\right] \leq C\cdot E\left[\sum_{j=0}^{r_i-1} C_{i,j}\cdot Q_{ij}|\mathcal{D}\right], \quad (18)$$

where $C>0$ is a fixed number and

$$Q_{i,j} = \max_{\sigma,\tau\in[k]^{\{v_{ij}\}}}\left\{||\mu_{i,j}(\cdot|\sigma) - \mu_{ij}(\cdot|\tau)||_{\Psi_i\cup\{u_{i,j}\}} + ||\mu_{i,j}(\cdot|\sigma) - \mu_{ij}(\cdot|\tau)||_{u_{ij}}\right\}.$$

Clearly (18) holds since, conditioning on event $\mathcal{D}$, we have a constant lower bound on $Pr[X_i(v_i)\neq X_i(u_i)]$, for every i. Also, the following holds: For any $i\leq i_0$ we have that

$$E\left[\sum_{j=0}^{r_i-1} C_{i,j}\cdot Q_{ij}|\mathcal{D}\right] \leq 2k^4 \sum_{j=0}^{5dn^{1/2}\log n} E[Q_{i,j}|\mathcal{D}], \quad (19)$$

since from conditioning on $\mathcal{D}$, it holds that $r_i \leq 10dn^{1/2}\log n$ and $C_{ij} < 2k^4$. Also, we have the following,

$$E[Q_{ij}|\mathcal{D}] \leq \frac{E[Q_{ij}]}{Pr[\mathcal{D}]} \leq \frac{35}{\epsilon} n^{-\left(1+\frac{\log\gamma}{10\log(d)}\right)} \qquad [\text{as } Pr[\mathcal{D}]>3/4], \quad (20)$$

where the bound for $E[Q_{i,j}]$ in the last inequality follows by working exactly as in Lemma 4.2. The quantity γ is defined in Proposition 4.3. We remind the reader than for $i<i_0$ the distance between $v_{i,j}$ and $u_{i,j}$ is at least $\frac{\log n}{10\log d}$.

Plugging into (18) the inequalities in (20) and (19), we get the following: For sufficiently large n and for any $i\leq i_0$ we have that

$$E\left[|\frac{Pr[X_i(v_i)\neq X_i(u_i)] - Pr[Y_i(v_i)\neq Y_i(u_i)]}{Pr[X_i(v_i)\neq X_i(u_i)]}||\mathcal{D}\right] \leq n^{-\frac{1}{2}-\frac{\log\gamma}{10\log(d)}}. \quad (21)$$

From the pseudocode of the schema for $G(n,d/n)$ we have that for $i\geq i_0$ the schema estimates $Pr[X_i(v_i)\neq X_i(u_i)]$ by assuming that they are 1. Assuming that the event $\mathcal{D}$ holds, then, it is not hard to show that

$$\frac{|Pr[X_i(v_i)\neq X_i(u_i)] - 1|}{Pr[X_i(v_i)\neq X_i(u_i)]} = \Theta(1) \qquad \text{for } i\geq i_0. \quad (22)$$

Plugging (21) and (22) into (17) we get that

$$E\left[\frac{1}{n}|\log \mathcal{Z} - \log Z(G(n,d/n))||\mathcal{D}\right] \leq 2n^{-\left(1/2+\frac{\log\gamma}{11\log d}\right)}.$$

Using Markov's inequality we get that

$$Pr\left[\frac{1}{n}|\log \mathcal{Z} - \log Z(G(n,d/n),k)| \geq n^{-1/4}|\mathcal{D}\right] \leq 2n^{-\left(1/4+\frac{\log\gamma}{11\log d}\right)}.$$

The proposition follows from the above inequality and the fact that $Pr[\mathcal{D}] \geq 1-n^{-\beta}$, for fixed $\beta > 0$. $\Diamond$

4.2 Proof of Proposition 4.3

For the proof of Proposition 4.3, we need the following result.

Lemma 4.4. *Consider the graph $G(n,d/n)$ and let π be a permutation of $l+1$ vertices of $G_{n,d/n}$, for $0 \leq l \leq \Theta(\log^6 n)$. Consider, also, the product measure $\mathcal{P}_{k,x_1}$ w.r.t. the graph $G(n,d/n)$, where $x_1 = \pi(1)$ and $k \geq (2+\epsilon)d$. Setting $\Gamma = 1$ if π is a path of disagreement, otherwise $\Gamma = 0$, it holds that*

$$E[\Gamma] \leq \left(\frac{d}{n}\right)^l \cdot \left(\left(\frac{1}{(1+\epsilon/2)d} + d^{-20}\right)^l + 2n^{-\log^4 n}\right),$$

where the expectation is taken w.r.t. both $\mathcal{P}_{k,x_1}$ and $G(n,d/n)$.

Proof: Call π the path that corresponds to the permutation π, e.g. $\pi = (x_1,\ldots x_{l+1})$. Let I_π be the event that there exists the path $(x_1,\ldots,x_{l+1})$ in $G_{n,d/n}$. It holds that

$$E[\Gamma] = \left(\frac{d}{n}\right)^l \cdot E[\Gamma|I_\pi],$$

Let Q_π denote the event that the vertices in π have degree less than $\log^6 n$. Using Chernoff bounds it is easy to show that $Pr[Q_\pi|I_\pi] \geq 1-n^{-\log^4(n)}$. Also, it holds that

$$\begin{aligned} E[\Gamma|I_\pi] &= E[\Gamma|I_\pi,Q_\pi]Pr[Q_\pi|I_\pi] + E[\Gamma|I_\pi,\bar{Q}_\pi]Pr[\bar{Q}_\pi|I_\pi] \\ &\leq E[\Gamma|I_\pi,Q_\pi] + n^{-\log^4(n)}. \end{aligned}$$

It suffices to show that for $0 \leq l \leq \Theta(\log^6 n)$ and sufficiently large n it holds that

$$E[\Gamma|I_\pi,Q_\pi] \leq \left(\frac{1}{(1+\epsilon/2)d} + d^{-20}\right)^l. \tag{23}$$

We show (23) by using induction on l. Clearly for $l=0$ the inequality in (23) is true. Assuming that (23) holds for $l=l_0$, we will show that it holds for $l=l_0+1$, as well.

Let D_i, denote the event that the vertex x_i is disagreeing. It suffices to show that

$$Pr[D_{l_0+1}|\wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] \leq \frac{1}{(1+\epsilon/2)d} + d^{-20}. \tag{24}$$

Using the law of total probability, we have that

$$\begin{aligned} Pr[D_{l_0+1}|\wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] &\leq Pr[D_{l_0+1}|\wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi, \Delta_{l_0+1}=0] + \\ &\quad +Pr[\Delta_{l_0+1}>0|\wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi], \end{aligned} \tag{25}$$

where Δ_{l_0+1} is the number of edges that are incident to x_{l_0+1} and some vertex in $\{x_1, \ldots, x_{l_0-1}\}$.

Given that all vertices in $\{x_1, \ldots, x_{l_0}\}$ are disagreeing, let δ_i be the number of vertices in $V\backslash\{x_1, \ldots, x_{l_0}\}$ that are adjacent to x_i, for $1 \leq i \leq l_0$. If $\delta_i = t$, then all the possible subsets of $V\backslash\{x_1, \ldots, x_{l_0}\}$ with cardinality t are equiprobably adjacent to x_i. This implies that the probability for x_{l_0+1} to be adjacent to x_i is $\frac{E[\delta_i]}{n-l_0}$. By the linearity of expectation we have

$$E[\Delta_{l_0+1}| \wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] \leq \frac{1}{n-l_0}\sum_{s=1}^{l_0} E[\delta_s| \wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] \leq n^{-0.97}, \tag{26}$$

the last inequality follows from the fact that $l_0 \leq \Theta(\log^6 n)$ and all the expectations in the sum are upper bounded by $\log^6 n$, due to conditioning on Q_π. By (26) and Markov's inequality, we get that

$$Pr[\Delta_{l_0+1} > 0| \wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] \leq n^{-0.97}. \tag{27}$$

Also, we have that

$$\begin{aligned}
\varrho &= Pr[D_{l_0+1}| \wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi, \Delta_{l_0+1} = 0] \\
&\leq \sum_{j=0}^{k-3} \frac{1}{k-2-j}\binom{n}{j}(d/n)^j(1-d/n)^{n-j} + \sum_{j=k-2}^{n-2}\binom{n}{j}(d/n)^j(1-d/n)^{n-j} \\
&\leq \frac{1}{(2+\epsilon)d/2}\sum_{j=0}^{(2+\epsilon)d/2}\binom{n}{j}(d/n)^j(1-d/n)^{n-j} + \sum_{j=(2+\epsilon)d/2+1}^{n-2}\binom{n}{j}(d/n)^j(1-d/n)^{n-j} \\
&\leq \frac{1}{(2+\epsilon)d/2} + \exp\left(-cd\right),
\end{aligned} \tag{28}$$

where $c = \log c' - 1 + 1/c'$ and $c' = (1 + \epsilon/2)$. The last inequality follows from Chernoff bounds, i.e. Corollary 2.4 in [15]. Plugging (28) and (27) into (25), for large d we get that

$$Pr[D_{l_0+1}| \wedge_{j=1}^{l_0} D_j, I_\pi, Q_\pi] \leq \frac{1}{(1+\epsilon/2)d} + d^{-20}.$$

That is, (24) is true. The lemma follows. $\diamondsuit$

Proof of Proposition 4.3: Consider an enumeration of all the permutations of $t \geq l$ vertices in $G(n, d/n)$ with first the vertex x and last some vertex of Ψ. Let $\pi_0(t), \pi_1(t), \ldots$ be the permutations in the order they appear in the enumeration. Also, w.r.t. the graph $G(n, d/n)$, consider the product measure $\mathcal{P}_{k,x}$ as it is defined in the statement of Theorem 3.1. Let $\Gamma_i(t)$ be the random variable such that

$$\Gamma_i(t) = \begin{cases} 1 & \text{the path that corresponds to } \pi_i(t) \text{ is a path of disagreement} \\ 0 & \text{otherwise.} \end{cases}$$

Let, also, $\Gamma(t) = \sum_i \Gamma_i(t)$.

Let $\mathcal{E} = 1$ if the event *"there is no path of disagreement that starts from x and has length larger than $t_0 = \frac{10\log n}{\log(1.04)}$"* occurs and $\mathcal{E} = 0$ otherwise. It holds that

$$\begin{aligned}
\mathcal{P}_{k,x_1}\left[\sum_{t\geq l}\Gamma(t) > 0\right] &\leq \mathcal{P}_{k,x_1}\left[\sum_{t\geq l}\Gamma(t) > 0|\mathcal{E} = 1\right]\mathcal{P}_{k,x_1}[\mathcal{E} = 1] + \mathcal{P}_{k,x_1}[\mathcal{E} = 0] \\
&\leq \mathcal{P}_{k,x_1}\left[\sum_{l\leq t<t_0}\Gamma(l) > 0\right] + Pr[\mathcal{E} = 0].
\end{aligned} \tag{29}$$

For convenience, we let $\varrho = \mathcal{P}_{k,x_1}\left[\sum_{t\geq l}\Gamma(t) > 0\right]$, $\varrho_1 = \mathcal{P}_{k,x_1}\left[\sum_{l\leq t<t_0}\Gamma(l) > 0\right]$ and $\varrho_2 = Pr[\mathcal{E} = 0]$. The proposition follows by deriving an appropriate upper bound for $E[\varrho]$, where the expectation is taken w.r.t. graph instances. For this we bound appropriately $E[\varrho_1]$ and $E[\varrho_2]$ and use the following inequality (which follows from (29))

$$E[\varrho] \leq E[\varrho_1] + E[\varrho_2]. \tag{30}$$

It holds that

$$\begin{aligned} E[\varrho_1] &\leq \sum_{l\leq t<t_0}\sum_i E[\Gamma_i(t)] \\ &\leq \sum_{l\leq t<t_0}\frac{|\Psi|}{n}d^t\cdot\left(\left(\frac{1}{(1+\epsilon/2)d}+d^{-20}\right)^t + 2n^{-\log^4 n}\right), \end{aligned}$$

where in the last inequality we use Lemma 4.4 and the fact that between x_1 and Ψ there are at most $|\Psi|\cdot n^{t-1}$ paths of length exactly t. Since $t\leq \log^2 n$, it is direct that

$$E[\varrho_1] \leq \sum_{l\leq t<t_0}\frac{|\Psi|}{n}(1+\epsilon/4)^{-t} \leq \frac{4+\epsilon}{\epsilon}\frac{|\Psi|}{n}(1+\epsilon/4)^{-l}. \tag{31}$$

Observe that $\mathcal{P}_{k,x_1}[\mathcal{E}=0] \leq \mathcal{P}_{k,x_1}[H(t_0)>0]$, where $H(t_0)$ denotes the number of paths of disagreement of length t_0 that start from vertex x_1. Note that the paths that $H(t_0)$ counts do not necessarily end in Ψ. By Markov's inequality, we have that

$$\mathcal{P}_{k,x_1}[\mathcal{E}=0] \leq E_{\mathcal{P}}[H(t_0)].$$

Clearly, the above implies that $E[\varrho_2]\leq E[H(t_0)]$, where the expectations is taken w.r.t. both $\mathcal{P}_{k,x_1}$ and the graph instances. We use Lemma 4.4 to bound $E[H(t_0)]$ and we get that

$$\begin{aligned} E[\varrho_2] &\leq n^{t_0}\left(\frac{d}{n}\right)^{t_0}\left(\left(\frac{1}{(1+\epsilon/2)d}+d^{-20}\right)^{t_0}+2n^{-\log^4 n}\right) \qquad \text{[from Lemma 4.4]} \\ &\leq \left(\frac{1}{1+\epsilon/4}\right)^{\log^2 n} + n^{-\frac{1}{2}\log^4 n}. \end{aligned} \tag{32}$$

The proposition follows by plugging (31) and (32) into (30). $\Diamond$

4.3 Proof of Corollary 1.1

For proving the corollary we are going to use Lemma 2.1. In particular, it suffices to have the following: W.h.p over $G(n,d/n)$ all but a vanishing fraction of the probability terms $Pr[X(v_i)\neq X(u_i)]$ are within distance $o(1)$ from $\left(1-\frac{1}{k}\right)$. Also, the remaining probability terms, i.e. those which are not close to $\left(1-\frac{1}{k}\right)$ are bounded well away from zero.

The corollary follows immediately from Lemmas 4.1, 4.3. That is, consider the sequence of subgraph $\mathcal{G}(G(n,d/n))$ we have for the counting algorithm. From Lemma 4.3 and Lemma 4.1 we have that w.h.p. the situation is as follows: There is a set of indices I such that for every $i\in I$ it holds that

$$|Pr[X(v_i)\neq X(u_i)] - \left(1-\frac{1}{k}\right)| \leq n^{-\frac{\log\gamma}{21\log d}}. \tag{33}$$

For the rest indices, i.e. $i \notin I$ it holds that

$$|Pr[X(v_i) \neq X(u_i)] - \left(1 - \frac{1}{k}\right)| = \Theta(1). \tag{34}$$

From Lemma 2.1 we can write $\frac{1}{n}\log(Z(G(n,d/n),k))$ as follows:

$$\begin{aligned}\frac{1}{n}\log Z(G(n,d/n),k) &= k + \frac{1}{n}\sum_{i=1}^{r}\log Pr[X(v_i) \neq X(u_i)] \\ &= k + \frac{1}{n}\sum_{i\in I}\log Pr[X(v_i) \neq X(u_i)] + \frac{1}{n}\sum_{i\notin I}\log Pr[X(v_i) \neq X(u_i)],\end{aligned}$$

while from Lemma 4.1 we get that w.h.p. $|I| \geq n - O(n^{3/10}\log n)$. We derive upper and lower bounds for $\frac{1}{n}\log Z(G(n,d/n),k)$ by working as follows:

$$\begin{aligned}\frac{1}{n}\log Z(G(n,d/n),k) &\leq k + \frac{|I|}{n}\left(\left(1-\frac{1}{k}\right) + n^{-\frac{\log\gamma}{21\log d}}\right) + \frac{n-|I|}{n} \\ &\leq k + \frac{d}{2}\left(1-\frac{1}{k}\right) + 2n^{-\frac{\log\gamma}{21\log d}},\end{aligned} \tag{35}$$

where in the last inequality we used the lower bound for the cardinality of the set I. Working in exactly the same manner we get the lower bound for $\frac{1}{n}\log Z(G(n,d/n),k)$. The corollary follows.

4.4 Proof of Corollary 1.2

Consider the following sequence of subgraphs $\mathcal{G}(G_{n,d/n})$ (different than what we used previously): The term-graph G_0 is edgeless. There is an index i_1 such that for $0 < i \leq i_1$, G_i contains all the edges that belong to cycles of length at most $\frac{\log n}{10\log d}$ in $G_{n,d/n}$ and only these edges. We refer to the cycle of length less than $\frac{\log n}{10\log d}$ as "small cycles".

Let $S(n,d)$ be the set of instances of $G_{n,d/n}$ which have (A) $\Theta(n)$ edges, (B) $i_1 \leq \Theta(n^{0.3}\log n)$ and (C) each $B(v_i, \frac{\log n}{4\log(e^2d/2)})$ is either a tree or unicyclic.

We are going to show that for every $G \in S(n,d)$ and every term $G_i \in \mathcal{G}(G)$ such that $i \geq i_1$, we can verify in polynomial time that

$$||\mu(\cdot|\sigma_{v_i}) - \mu(\cdot|\eta_{v_i})||_{u_i} \leq n^{-\epsilon_1}, \tag{36}$$

where $\epsilon_1 > 0$. Then the corollary follows by using standard arguments, i.e. from Lemma 2.1 and from the fact that $\left|Pr[X_i(u_i) \neq X_i(v_i)] - \left(1-\frac{1}{k}\right)\right| \leq \max_{\sigma,\eta\in[k]^{\{v_i\}}} ||\mu_i(\cdot|\sigma) - \mu_i(\cdot|\eta)||_{u_i}$.

The value of ϵ_1 in (36) depends on the function $h(n,k,d)$ and i_1. For $i < i_1$ it direct to see that G_i is so simple that we can compute $Pr[X_{u_i} \neq X_{v_i}]$ exactly. Theorem 3.1 and Corollary 4.1 suggest that

$$||\mu(\cdot|\sigma_{v_i}) - \mu(\cdot|\eta_{v_i})||_{u_i} \leq \mathcal{P}_{k,v_i}[\,\exists \text{ path of disagreement connecting } \{v_i\} \text{ and } \{u_i\}]. \tag{37}$$

where $\mathcal{P}_{k,v_i}$ is the product measure defined in Section 3.2 and it is taken w.r.t graph $G_{n,d/n}\backslash\{v_i,u_i\}$. For $i > i_1$ it holds that $dist(v_i,u_i) \geq \frac{\log n}{10\log(d)}$ in $G_{n,d/n}\backslash\{v_i,u_i\}$. Consider, now, the event

$$E_{v_i,c} = \text{"}\exists \text{ a path of disagreement that connects } v_i \text{ with } L(v_i, c\log n) \text{ in } G_{n,d/n}\backslash\{v_i,u_i\}\text{"}.$$

For each pair v_i u_i define

$$a_i = \min\left\{\frac{dist(v_i,u_i)}{\log n}, (4\log(e^2d/2))^{-1}\right\}.$$

Noting that, for fixed $c_1 > c_2$ it holds that $\mathcal{P}_{k,v_i}[E_{v_i,c_1}] \leq \mathcal{P}_{k,v_i}[E_{v_i,c_2}]$, we get that

$$\mathcal{P}_{k,v_i}[\exists \text{ path of disagreement connecting } \{v_i\} \text{ and } \{u_i\} \text{ in } G_{n,d/n}\backslash\{v_i,u_i\}] \leq \mathcal{P}_{k,v_i}[E_{v_i,a_i}]. \tag{38}$$

By (36) (37) and (38), we can verify (36) by using the criterion $\mathcal{P}_{k,v_i}(E_{v_i,a_i}) \leq n^{-\epsilon_1}$. It remains to show that $\mathcal{P}_{k,v_i}(E_{v_i,a_i}) \leq n^{-\epsilon_1}$, for $i \geq i_1$, can be verified in polynomial time. Let T_{v_i,a_i} be the set of all simple paths that connect v_i to $L(v_i, a_i \log n)$, it holds that

$$\mathcal{P}_{k,v_i}[E_{v_i,a_i}] \leq \sum_{m \in T_{v_i,a_i}} \mathcal{P}_{k,v_i}[\text{“}m \text{ is a path of disagreement”}]. \tag{39}$$

The computation of each probability term on the r.h.s. of the above inequality can be carried out in polynomial time. It suffices to show that w.h.p. the number of these terms is polynomially large.

Using Lemma 2.1 from [10] we get that for every $i > i_1$ the subgraph $B(v_i, a_i \log n)$ of $G_{n,d/n}\backslash\{v_i,u_i\}$, is a tree with at most an extra edge, with probability at least $1-n^{-0.1}$. In this case, the number of simple paths between v_i and $L(v_i, a_i \log n)$ is at most $2|L(v_i, a_i \log n)|$. Also, with standard arguments (e.g. see Theorem 6 in [24]), it holds that with probability at least $1 - o(n^{-2})$, $|L(v_i, a_i \log n)| \leq n^{0.26} \log n$, for every $i > i_1$. That is, for every $i > i_1$, $|T_{v_i,a_i}|$ is polynomially large with probability at least $1 - 2n^{-0.1}$. Thus, the probability term on the l.h.s. of (39) can be computed efficiently, for any $i > i_1$, w.h.p.

Using the arguments in the paragraph above and Lemma 4.1 it is direct to show that $Pr[G(n, d/n) \in S(n, d)] \geq 1 - 3n^{-0.1}$. Also, it is direct that we can decide whether $G(n, d/n) \in S(n, d)$ or not, efficiently. The corollary follows

5 Bounds for spatial correlation decay - Proof of Theorem 3.1

For some finite graph $G = (V, E)$ and some sufficiently large integer k, let $\mu(\cdot)$ be the Gibbs distribution of the k-colourings of G. For $x \in V$, $\Lambda \subseteq V$ and $\sigma_x, \eta_x \in [k]^{\{x\}}$, we are interested in deriving upper bounds for following quantity

$$||\mu(\cdot|\sigma_x) - \mu(\cdot|\eta_x)||_\Lambda. \tag{40}$$

Towards bounding the above quantity we introduce two random variables X^σ, $X^\eta \in [k]^V$ distributed as in $\mu(\cdot|\sigma_x)$ and $\mu(\cdot|\eta_x)$, respectively. We couple X^σ and X^η and we use the following inequality from the Coupling Lemma (see [2]),

$$||\mu(\cdot|\sigma_x) - \mu(\cdot|\eta_x)||_\Lambda \leq Pr[X^\sigma(\Lambda) \neq X^\eta(\Lambda) \text{ in the coupling}].$$

We provide a upper bound for the probability of the event "$X^\sigma(\Lambda) \neq X^\eta(\Lambda)$" in the coupling, in terms of k and the degrees of the vertices in G by using "*disagreement percolation*", [6]. In Section 5.1 we describe the coupling between X^σ and X^τ.

5.1 The coupling for the comparison

Let Ω_σ and Ω_η denote the k-colourings of G that assign the vertex x colour σ_x and η_x, respectively. For the coupling of X^σ and X^η we need to develop, first, a bijection $T : \Omega_\sigma \to \Omega_\eta$ as follows:

Given $\xi \in \Omega_\sigma$, we let $G_\xi = (V_\xi, E_\xi)$, *induced subgraph* of G, be defined as follows: In the colouring ξ, let V_σ and V_η be the colour classes specified by the colours σ_x and η_x, respectively. Then $G_\xi = (V_\xi, E_\xi)$ is the *maximal*, *connected* graph such that $x \in V_\xi$ and $V_\xi \subseteq V_\sigma \cup V_\eta$. That is, G_ξ is the maximal, connected, induced subgraph of G which contains x and vertices only from the colour classes V_σ and V_η, in the colouring ξ. Then, given G_ξ, we derive $T\xi$ by working as follows: For every vertex

$u \notin G_\xi$ it holds that $\xi(u) = (T\xi)(u)$. For $u \in G_\xi$ if $\xi(u) = \sigma_x$, then $(T\xi)(u) = \eta_x$. Also, if $\xi(u) = \eta_x$, then $(T\xi)(u) = \sigma_x$.

In Figures 6 and 7, in Section 3.2, we illustrate how does the mapping T work. Of course, it is not direct that T is a bijection. For this we provide the following lemma.

Lemma 5.1. *It holds that* $T : \Omega_\sigma \to \Omega_\eta$ *is a bijection.*

Proof: For the colouring $\xi \in \Omega_\sigma$, consider $G_\xi = (V_\xi, E_\xi)$ as defined above. We need to focus on three properties that G_ξ has. First, it is easy to see that G_ξ should be bipartite (in the extreme case where $V_\xi = \{x\}$ we consider G_ξ bipartite too). Second, G_ξ is connected due to the way we consider it. Third, the fact that G_ξ is maximal implies the following: if $\partial V_\xi = \{v \in V \backslash V_\xi | \{v, u\} \in E \text{ for } u \in V_\xi\}$, then $\forall v \in \partial V_\xi$ it holds $\xi_v \notin \{\sigma_x, \eta_u\}$.

Clearly ξ specifies a proper 2-colouring for the vertices of G_ξ that uses only the colours σ_x and η_x. In particular, let $p_1, p_2 \subseteq V_\xi$ be the two parts of G_ξ and w.l.o.g. assume that x belongs to p_1. Then, ξ assigns to all the vertices in p_1 the colour σ_x and to all the vertices in p_2 the colour η_x. In that terms, the mapping T works as follows: For every vertex $v \in V \backslash V_\xi$ to hold $(T\xi)_v = \xi_v$. For the remaining vertices, i.e. those that belong to G_ξ, the mapping T swaps the colour assignments of the two parts of G_ξ. First we show that T maps every colouring of Ω_σ to Ω_σ.

Claim 5.1. *For every* $\xi \in \Omega_\sigma$ *it holds that* $(T\xi) \in \Omega_\eta$*.*

Proof: It is direct that $(T\xi)_x = \eta_x$. It remains to show that $T\xi$ is a proper colouring of G.

If $T\xi$ is a non proper colouring, then there should be , at least, two adjacent vertices (somewhere in G) having the same colour assignment. The swap of colour assignments that take place, when we apply T on ξ, involves only vertices in V_ξ. Thus if $(T\xi)$ is a non proper colouring, then the monochromatic pair of adjacent vertices has either both vertices in V_ξ or one vertex in V_ξ and the other in ∂V_ξ.

It is direct that swapping the colour assignments of the two parts of G_ξ, as these are specified by ξ, leads to a proper colouring of G_ξ. Thus, in $T\xi$ there is no monochromatic pair whose both vertices belong to G_ξ. Also, this swap of colourings cannot lead some vertex in V_ξ to have the same colour assignment with some vertex in ∂V_ξ. This is due to the maximality of G_ξ, i.e. the colouring ξ cannot not specify colour assignment that uses the colours σ_x and η_x for any vertex in ∂V_ξ. Thus, for every $\xi \in \Omega_\sigma$, it holds that $T\xi$ is a proper colouring of G. The claim follows. $\Diamond$

It remains to show that T is a bijection. The next claim shows that T is a surjective.

Claim 5.2. T *is surjective.*

Proof: Let ξ' be any member of Ω_η. We are going to show that there exists $\xi \in \Omega_\sigma$ such that $T\xi = \xi'$.

For the colouring ξ', let $G_{\xi'} = (V_{\xi'}, E_{\xi'})$ be the maximal, connected bipartite subgraph of G such that $x \in V_{\xi'}$ and $\forall v \in V_{\xi'}$ it holds $\xi'_v \in \{\sigma_x, \eta_x\}$, (i.e. $G_{\xi'}$ is derived in a similar way as G_ξ, above).

The colouring ξ' specifies a proper 2-colouring for $G_{\xi'}$ that uses only the colours σ_x and η_x. Let $p_1, p_2 \subseteq V_\xi$ be the two parts of $G_{\xi'}$ and w.l.o.g. assume that ξ' assigns to all the vertices in p_1 the colour η_x and to all the vertices in p_2 the colour σ_x.

Consider the colouring ξ which is derived by ξ' by swapping the colour assignments of the two parts of $G_{\xi'}$ while $\xi_v = \xi'_v$ for $v \in V \backslash V_{\xi'}$. With arguments similar to those in the proof of Claim 5.1 we can see that $\xi \in \Omega_\sigma$. The claim follows by noting, additionally, that $T\xi = \xi'$. $\Diamond$

In the following claim we show that T is one-to-one.

Claim 5.3. T *is one-to-one.*

Proof: Assume that there are two colourings $\xi^1, \xi^2 \in \Omega_\sigma$ such that $T\xi^1 = T\xi^2 = \xi^3$. We are going to show that it should hold $\xi^1 = \xi^2$. For this, assume the opposite, i.e. $\xi^1 \neq \xi^2$. We consider the graphs G_{ξ^1} G_{ξ^2} and G_{ξ^3}, as in the proofs of the two previous claims. By the proofs of these claims we know that the graphs G_{ξ^1}, G_{ξ^2} and G_{ξ^3} have the same subset of vertices of G.

Thus, we conclude that the colourings ξ^1 and ξ^2 should differ only on the colour assignment of the vertices in the graph G_{ξ^1}. We remind the reader that this graph is a connected bipartite graph with ξ^1 and ξ^2 specifying proper 2-colourings for G_{ξ^1} which both using the colours $\{\sigma_x, \eta_x\}$.

By assumption, the 2-colouring for G_{ξ^1} that ξ^1 specifies is different than that of $T\xi^1$. The same holds for colouring of ξ^2 and $T\xi^2$. Since $T\xi^1 = T\xi^2$ we deduce that there exist three different 2-colourings for G_{ξ^1}. There is a contradiction, here, since there can exist only two 2-colourings for G_{ξ^1}. The claim follows. $\Diamond$

Since the mapping $T : \Omega_\sigma \to \Omega_\eta$ is surjective (Claim 5.2) and one-to-one (Claim 5.3), it is a bijection. The lemma follows. $\Diamond$

Lemma 5.2. *There exists a coupling of X^σ with X^η such that*

$$X^\eta = TX^\sigma.$$

Proof: The existence of the bijection T implies that $|\Omega_\sigma| = |\Omega_\eta|$. Thus $\forall \xi \in \Omega(G, k, \sigma_x)$ it holds that

$$\mu\left(\xi|\sigma_x\right) = \mu\left((T\xi)|\eta_x\right) = \frac{1}{|\Omega_\sigma|}.$$

This implies that $Pr[X^\sigma = \xi] = Pr[X^\eta = T\xi], \forall \xi \in \Omega_\sigma$. The lemma follows by noting that

$$\left(\sum_{\xi \in \Omega_\sigma} Pr[X^\sigma = \xi]\right) = 1 \qquad \text{and} \qquad \left(\sum_{\xi \in \Omega_\sigma} Pr[X^\eta = (T\xi)]\right) = 1.$$

$\Diamond$

Let $\nu : [k]^V \times [k]^V \to [0, 1]$ denote the joint distribution of the colourings X^σ and X^η in the coupling where $X^\eta = TX^\sigma$. We close the section by providing a very useful property of ν, which we use in the disagreement percolation.

Lemma 5.3. *For every $u \in V \backslash \{x\}$, let N_u be the set that contains all the vertices which are adjacent to the vertex u in G. Also, let $\mathcal{B}_u \subseteq [k]^{N_u} \times [k]^{N_u}$ be defined such that*

$$\mathcal{B}_u = \{\xi \in [k]^{N_u} \times [k]^{N_u} | \nu(\xi) > 0\}.$$

If $k > \Delta$, then it holds that

$$\max_{\tau \in \mathcal{B}_u} \nu(X^\sigma(u) \neq X^\eta(u)|\tau) \leq \frac{1}{k - \Delta_u}$$

where Δ_u is the degree of vertex u in G.

Proof: Let $G_X = (V_X, E_X)$, denote the induced subgraph of G such that $v \in G_X$ if and only if $X^\sigma(v) \neq X^\eta(v)$, in the coupling. We remind the reader that under both X^σ and X^η, G_X is coloured using only the colours σ_x and η_x.

There are two necessary conditions for some vertex $v \in V \backslash \{x\}$ to be in V_X. The first one is that some vertex in N_u should, also, belong to V_X. This is due to the fact that G_X is connected. The second is the following one: Assume that $w_1 \in N_u$ and $w_1 \in V_X$. If there exists $w_2 \in N_u \backslash \{w_1\}$ and $X^\sigma(w_2) \in \{\sigma_x, \eta_x\}$, then it should hold $X^\sigma(w_1) = X^\sigma(w_2)$. This should hold under both X^σ and X^η, G_X is coloured using only the colours σ_x and η_x.

Considering the two previous conditions the worst case of $X^\sigma(N_u)$ is the following: At least one vertex in N_u belongs to V_X, call this vertex w. No vertex in N_u uses the colour $\{\sigma_x, \eta_x\} \backslash \{X^\sigma(w)\}$. $X^\sigma(N_u)$ is such that the number of different colour that are used is equal to $|N_u|$. In that case the probability of u to belong to V_X is $\frac{1}{k-\Delta_u}$. The lemma follows. $\Diamond$

Lemma 5.3 assumes that $k > \Delta$, otherwise it holds

$$\max_{\tau \in \mathcal{B}_u} \nu(X^\sigma(u) \neq X^\eta(u)|\tau) \leq 1.$$

5.2 Proof of Theorem 3.1

By Theorem 1 and Corollary 1.1 in [6], and Lemma 5.3 we get that

$$||\mu(\cdot|\sigma_x) - \mu(\cdot|\eta_x)||_\Lambda \leq \mathcal{P}_{k,x}[\exists \text{ path of disagreement between } \{x\} \text{ and a vertex in } \Lambda].$$

We have to remark here that the coupling on which the disagreement percolation is based, has the following property: Let t be the minimum integer such that there is no path of disagreement connecting x to $L(x,t)$. Then, our coupling specifies that no vertex in $L(x,t')$, for $t' \geq t$ can be disagreeing. This is a crucial property of our coupling, since otherwise we could not apply the disagreement percolation technique (see [13]).

6 Rest of the Proofs

6.1 Proof of Proposition 3.1

Let

$$err_i = |Pr[X_i(v_i) \neq X_i(u_i)] - Pr[Y_i(v_i) \neq Y_i(u_i)]| \qquad \text{for } 0 \leq i \leq r-1.$$

It holds that

$$\begin{aligned}
\log \mathcal{Z} &= \sum_{i=0}^{r-1} \log(P[Y_i(v_i) \neq Y_i(u_i)]) + \log Z(G_0, k) \\
&\leq \textstyle\sum_{i=0}^{r-1} \log\left(P[X_i(v_i) \neq X_i(u_i)] + err_i\right) + \log Z(G_0, k) \\
&\leq \textstyle\sum_{i=0}^{r-1} \log\left(P[X_i(v_i) \neq X_i(u_i)]\right) + \sum_{i=0}^{r-1} \log\left(1 + \frac{err_i}{P[X_i(v_i) \neq X_i(u_i)]}\right) + \log Z(G_0, k) \\
&\leq \log Z(G, k) + \sum_{i=0}^{r-1} \log\left(1 + \frac{err_i}{P[X_i(v_i) \neq X_i(u_i)]}\right) \\
&\leq \log Z(G, k) + \sum_{i=0}^{r-1} \frac{err_i}{P[X_i(v_i) \neq X_i(u_i)]}.
\end{aligned}$$

The final derivation follows by the fact that $\log(x)$ is an increasing function (the base is of the logarithm is $e > 1$) and by $1 + x \leq e^x$, for any x. Similarly we get the lower bound for $\log(\mathcal{Z})$. The theorem follows.

6.2 Proof of Proposition 3.2

Proposition 3.2 follows as a corollary of the two following lemmas.

Lemma 6.1. *It holds that*

$$|Pr[X_i(v_i) \neq X_i(u_i)] - Pr[Y_i(v_i) \neq Y_i(u_i)]| \leq \sum_{j=0}^{r_i-1} ||\mu_{i,j}(\cdot) - \mu_{i,j+1}(\cdot)||_{\Psi_i}.$$

Proof: Let $\mu_{i,j}$ be the Gibbs distribution of the k-colourings of $G_{i,j}$. It holds that

$$|Pr[X_i(v_i) \neq X_i(u_i)] - Pr[X_{i,0}(v_i) \neq X_{i,0}(u_i)]| \leq \max_{A \subseteq [k]^{\Psi_i}} |\mu_{i,0}(A) - \mu_{i,r_i}(A)| \leq ||\mu_{i,0}(\cdot) - \mu_{i,r_i}(\cdot)||_{\Psi_i}$$

By the triangle inequality we get that $||\mu_{i,0}(\cdot) - \mu_{i,r_i}(\cdot)||_{\Psi_i} \leq \sum_{j=0}^{r_i-1} ||\mu_{i,j}(\cdot) - \mu_{i,j+1}(\cdot)||_{\Psi_i}$ $\Diamond$

Lemma 6.2. *Let Λ be any subset of vertices of $G_{i,j}$ that does not contain $v_{i,j}$ and $u_{i,j}$. It holds that*

$$||\mu_{i,j}(\cdot) - \mu_{i,j+1}(\cdot)||_\Lambda \leq C_{i,j} \max_{\sigma,\tau \in [k]^{\{v_{i,j}\}}} \left\{ ||\mu_{i,j}(\cdot|\sigma) - \mu_{i,j}(\cdot|\tau)||_{\Lambda \cup \{u_{ij}\}} + ||\mu_{i,j}(\cdot|\sigma) - \mu_{i,j}(\cdot|\tau)||_{\{u_{ij}\}} \right\}$$

where $C_{ij} = C_{i,j}(G_{i,j}, k) = \max_{s,t \in [k]} \left\{ (Pr[X_{i,j}(u_{i,j}) = s | X_{i,j}(v_{i,j}) = t])^{-2} \right\}$.

Proof: Let $\Omega_{i,j}$ denote the set of k-colourings of G_{ij} and μ_{ij} the uniform distribution over $\Omega_{i,j}$. It is straightforward that

$$||\mu_{i,j}(\cdot) - \mu_{i,j+1}(\cdot)||_\Lambda \leq \max_{\sigma,\tau} ||\mu_{i,j}(\cdot|\sigma_{\Psi_{i,j}}) - \mu_{i,j+1}(\cdot|\tau_{\Psi_{i,j}})||_\Lambda,$$

where τ varies in $\Omega_{i,j+1}$ and σ varies in $\Omega_{i,j}$. By the fact that $\Omega_{i,j+1} \subseteq \Omega_{i,j}$ and by the conditional independence, it holds that $\mu_{i,j+1}(\cdot|\tau_{\Psi_{i,j}}) = \mu_{i,j}(\cdot|\tau_{\Psi_{i,j}})$. Hence, we have that

$$||\mu_{i,j}(\cdot) - \mu_{i,j+1}(\cdot)||_\Lambda \leq \max_{\sigma,\tau} ||\mu_{i,j}(\cdot|\sigma_{\Psi_{i,j}}) - \mu_{i,j}(\cdot|\tau_{\Psi_{i,j}})||_\Lambda. \tag{41}$$

By definition (see (4)), there exists a set $\mathcal{A} \subseteq [k]^\Lambda$ such that

$$||\mu_{i,j}(\cdot|\sigma_{\Psi_{i,j}}) - \mu_{i,j}(\cdot|\tau_{\Psi_{i,j}})||_\Lambda = |\mu_{i,j}(\mathcal{A}|\sigma_{\Psi_{i,j}}) - \mu_{i,j}(\mathcal{A}|\tau_{\Psi_{i,j}})|.$$

Let $Q_{ij} = \mu_{ij}(\tau_{u_{ij}}|\tau_{v_{ij}}) - \mu_{ij}(\sigma_{u_{ij}}|\sigma_{v_{ij}})$. Using elementary probability theory relations we get the following:

$$\begin{aligned}
|\mu_{i,j}(\mathcal{A}|\sigma_{\Psi_{i,j}}) - \mu_{i,j}(\mathcal{A}|\tau_{\Psi_{i,j}})| &\leq \left| \frac{\mu_{i,j}(A, \tau_{u_{ij}}|\tau_{v_{ij}})}{\mu_{i,j}(\tau_{u_{ij}}|\tau_{v_{ij}})} - \frac{\mu_{i,j}(A, \sigma_{u_{ij}}|\sigma_{v_{ij}})}{\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}})} \right| \\
&\leq \left| \frac{\mu_{i,j}(A, \tau_{u_{ij}}|\tau_{v_{ij}})}{\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}}) + Q_{ij}} - \frac{\mu_{i,j}(A, \sigma_{u_{ij}}|\sigma_{v_{ij}})}{\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}})} \right| \\
&\leq \left| \frac{\mu_{i,j}(A, \tau_{u_{ij}}|\tau_{v_{ij}})}{\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}})} - \frac{\mu_{i,j}(A, \sigma_{u_{ij}}|\sigma_{v_{ij}})}{\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}})} \right| + \\
&\quad + \frac{|Q_{i,j}|}{\mu_{i,j}(\tau_{u_{ij}}|\tau_{v_{ij}})\mu_{i,j}(\sigma_{u_{ij}}|\sigma_{v_{ij}})}.
\end{aligned}$$

It is direct to see that

$$|\mu_{i,j}(A, \tau_{u_{ij}}|\tau_{v_{ij}}) - \mu_{i,j}(A, \sigma_{u_{ij}}|\sigma_{v_{ij}})| \leq \max_{\tau,\sigma} ||\mu_{i,j}(\cdot|\tau_{v_{ij}}) - \mu_{i,j}(\cdot|\sigma_{v_{ij}})||_{\Lambda^*}$$

$$|\mu_{ij}(\tau_{u_{ij}}|\tau_{v_{ij}}) - \mu_{ij}(\sigma_{u_{ij}}|\sigma_{v_{ij}})| \leq \max_{\tau,\sigma} ||\mu_{i,j}(\cdot|\tau_{v_{ij}}) - \mu_{i,j}(\cdot|\sigma_{v_{ij}})||_{u_{ij}},$$

where $\Lambda^* = \Lambda \cup \{u_{ij}\}$. The lemma follows. $\Diamond$

6.3 Proof of Lemma 2.1

Consider the sequence of subgraphs $\mathcal{G}(G) = G_0, \ldots, G_r$, where $r = |E|$ and G_0 is empty. Consider, also, the following telescopic relation

$$|\Omega(G,k)| = |\Omega(G_0,k)| \cdot \prod_{i=0}^{r-1} \frac{|\Omega(G_{i+1},k)|}{|\Omega(G_i,k)|} = k^n \cdot \prod_{i=0}^{r-1} \frac{|\Omega(G_{i+1},k)|}{|\Omega(G_i,k)|}.$$

The lemma will follow by showing that

$$Pr[X_i(u_i) \neq X_i(v_i)] = \frac{|\Omega(G_{i+1},k)|}{|\Omega(G_i,k)|}.$$

The above relation clearly holds by noting the following: The set of k-colourings of G_{i+1} is the same as the subset of k-colourings of G_i that contains all the colourings that assign v_i and u_i different colours. The lemma follows.

6.4 Proof of Lemma 3.1.

$$\begin{aligned}
||\mu(\cdot|\sigma_x) - \mu(\cdot)||_\Lambda &= \frac{1}{2} \sum_{\sigma_\Lambda \in [k]^\Lambda} |\mu(\sigma_\Lambda|\sigma_x) - \mu(\sigma_\Lambda)| \\
&= \frac{k}{2}\mu(\sigma_x) \sum_{\sigma_\Lambda \in [k]^\Lambda} |\mu(\sigma_\Lambda|\sigma_x) - \mu(\sigma_\Lambda)| \\
&= \frac{k}{2} \sum_{\sigma_\Lambda \in [k]^\Lambda} \mu(\sigma_\Lambda)|\mu(\sigma_x|\sigma_\Lambda) - \mu(\sigma_x)| \\
&\leq \frac{k}{2} \sum_{\sigma_\Lambda \in [k]^\Lambda} \mu(\sigma_\Lambda) \sum_{\tau_x \in [k]} |\mu(\tau_x|\sigma_\Lambda) - \mu(\tau_x)| \\
&\leq k \sum_{\sigma_\Lambda \in [k]^\Lambda} \mu(\sigma_\Lambda)||\mu(\cdot|\sigma_\Lambda) - \mu(\cdot)||_x.
\end{aligned}$$

Noting that it holds

$$||\mu(\cdot|\sigma_x) - \mu(\cdot|\tau_x)||_\Lambda \leq ||\mu(\cdot|\sigma_x) - \mu(\cdot)||_\Lambda + ||\mu(\cdot) - \mu(\cdot|\tau_x)||_\Lambda,$$

the lemma follows.

6.5 Proof of Lemma 4.1

Let $\epsilon = 1/(10\log(d))$. Assume that after removing all the edges in R there are two cycles of length at most $\epsilon \log n$ which are connected, i.e. these two cycles share edges. Then, there must exist a subgraph of $G_{n,d/n}$ that contains at most $2\epsilon \log n$ vertices while the number of edges exceeds by 1, or more, the number of vertices.

Let D be the event that in $G_{n,d/n}$ there exists a set of r vertices which have $r+1$ edges between them. For $r \leq \epsilon \log n$ we have the following:

$$\begin{aligned}
Pr[D] &\leq \sum_{r=1}^{\epsilon \log n} \binom{n}{r} \binom{\binom{r}{2}}{r+1} (d/n)^{r+1} (1-d/n)^{\binom{r}{2}-(r+1)} \\
&\leq \sum_{r=1}^{\epsilon \log n} \left(\frac{ne}{r}\right)^r \left(\frac{r^2 e}{2(r+1)}\right)^{r+1} (d/n)^{r+1} \leq \frac{e \cdot d}{2n} \sum_{r=1}^{\epsilon \log n} \left(\frac{e^2 d}{2}\right)^r \\
&\leq \frac{C}{n} \left(\frac{e^2 d}{2}\right)^{\epsilon \log n}.
\end{aligned}$$

Having $\epsilon \cdot \log(e^2 d/2) < 1$, the quantity in the r.h.s. of the last inequality is $o(1)$, in particular it is of order $\Theta(n^{\epsilon \log(e^2 d/2)-1})$. Thus, for $\epsilon = 1/(10\log(d))$ there is no connected component that contains two cycles with probability at least $1 - n^{-0.85}$.

Let C_l denote the number of cycles of length at most l in $G(n, d/n)$. It is direct to show that $E[C_l] \leq 2d^l$. Furthermore, $E[C_{\epsilon \log n}] \leq 2n^{1/10}$. It is not hard to see that the expected number of edges whose one end is on a cycle of length less than $\epsilon \log n$ is $O(n^{1/10} \log^2 n)$. That is $E[|R|] = O(n^{1/10} \log^2 n)$.

Employing the Markov inequality, we have $Pr[|R| \geq n^{3/10}] = O(n^{-0.2}/\log^2 n)$ while $Pr[C_{\epsilon \log n} \geq n^{3/10}] \leq 2n^{-0.2}$. The lemma follows.

Acknowledgement. The author would like to thank the anonymous reviewer as well as Prof. Mike Paterson for the time they spent to read this manuscript and for their corrections and their suggestions to improve the presentation of the result. Also, the author would like to thank Amin Coja-Oghlan for the numerous fruitful discussions.

References

[1] D. Achlioptas, and A. Naor. *The two possible values of the chromatic number of a random graph*. Annals of Mathematics, **162** (3), (2005), 1333-1349.

[2] D. Aldous. *Random walks of finite groups and rapidly mixing Markov chains.* In: Séminaire de Probabilités XVII 1981/82, Springer-Verlag, Berlin. pp. 243-297.

[3] A. Bandyopadhyay and D. Gamarnik. *Counting without sampling: new algorithms for enumeration problems using statistical physics.* In Random Structures and Algorithms, **33** (4), (2008), 452-479.

[4] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali. *Simple deterministic approximation algorithms for counting matchings.* In Proc. of the ACM 39th annual Symposium of Theory of Computing (STOC '07) pp. 122-127, 2007.

[5] N. Bhatnagar, J. Vera, E. Vigoda and D. Weitz. *Reconstruction for coloring on Trees.* SIAM J. Discrete Math. **25** (2), (2011) pp 809-826.

[6] J. van den Berg and C. Maes. *Disagreement percolation in the study of Markov fields.* Annals of Probability **22**, (1994) pp. 794-763.

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

[8] M. Dyer, L. A. Goldberg and M. Jerrum. *Counting and sampling H-colourings*. Information and Computation, **189** (1), (2004), pp 1-16.

[9] C. Efthymiou. *A simple algorithm for random colouring $G(n, d/n)$ using $(2+\epsilon)d$ colours*. In proc. of the 23th annual Symposium on Discrete Algorithms (SODA 2012), pp 272-280.

[10] C. Efthymiou and P. G. Spirakis. *Random sampling of colourings of sparse random graphs with a constant number of colours*. In Theoretical Computer Science **407**, (2008), pp 134-154.

[11] D. Gamarnik and D. Katz. *Correlation decay and deterministic FPTAS for counting list-colorings of a graph*. Journal of Discrete Algorithms, **12**. pp. 29 - 47.

[12] H. O. Georgii. *Gibbs measures and phase transitions*, de Gruyter Studies in Mathematics 9, Walter de Gruyter & Co., Berlin, 1988.

[13] O. Häggström. *A note on disagreement percolation.* In Random Structures and Algorithms **18** (2001), 267-278.

[14] T. P. Hayes, J. C. Vera and E. Vigoda. *Randomly coloring planar graphs with fewer colours than the maximum degree.* In proc. of the 39th ACM annual Symposium of Theory of Computing (STOC 2007), pp 450-458.

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

[16] M. Jerrum *Sampling, Counting and Integrating:Algorithms and Complexity*. Chapter 3, Birkäuser, Basel, 2003.

[17] M. Jerrum and A. Sinclair. *Approximate Counting, Uniform generation and Rapidly Mixing Markov Chains.* Infomation and Computation **82**(1), (1986), pp 93-133.

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

[19] M. Jerrum, L. G. Valiant and V. V. Vazirani. *Random generation of combonatorial structures from a uniform distribution.* Theoretical Computer Science, **43**, (1986), 169-188.

[20] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjianc, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems.* In Proc. National Academy of Sciences 104 (2007) 10318-10323.

[21] M. Mézard and A. Montanari. *Infomation, Physics and Computation*. Oxford University Press 2009.

[22] A. Montanari, R. Restrepo and P. Tetali *Reconstruction and Clustering in Random Constraint Satisfaction Problems.* SIAM J. Discrete Math. **25** (2), (2011), 771-808.

[23] A. Montanari and D. Shah. *Counting good truth assignments of random k-SAT formulae.* In proc. of the 18th annual Symposium on Discrete Algorithms (SODA 2007), pp. 1255-1264.

[24] E. Mossel and A. Sly. *Rapid Mixing of Gibbs Sampling on Graphs that are Sparse on Average.* In Random Structures and Algorithms **35** (2), (2009), 250-270.

[25] E. Mossel and A. Sly. *Gibbs Rapidly Samples Colorings of* $G_{n,d/n}$. In Probability Theory and Related Fields, **148**, (2010), 37-69.

[26] L. Valiant. *The complexity of enumeration and reliability problems*. SIAM Journal of Computing **8**,(1979), 410-421.

[27] E. Vigoda. *Improved bounds for sampling colorings.* In J. Math. Phys. **41**, (2000), pp. 1555-1569.

[28] M. Wainwright and M. Jordan. *Graphical Models, exponential families and variational inference*. Technical report 649, Dept. of Statistics, University of Berkeley.

[29] D. Weitz. *Counting independent sets up to the tree threshold.* In Proc. of the ACM 38th annual Symposium of Theory of Computing (STOC 2006).