Counting without sampling. New algorithms for enumeration problems using statistical physics

Antar Bandyopadhyay* David Gamarnik †

October 21, 2005

Abstract

We propose a new type of approximate counting algorithms for the problems of enumerating the number of independent sets and proper colorings in low degree graphs with large girth. Our algorithms are not based on a commonly used Markov chain technique, but rather are inspired by developments in statistical physics in connection with correlation decay properties of Gibbs measures and its implications to uniqueness of Gibbs measures on infinite trees, reconstruction problems and local weak convergence methods.

On a negative side, our algorithms provide ϵ-approximations only to the logarithms of the size of a feasible set (also known as free energy in statistical physics). But on the positive side, our approach provides deterministic as opposed to probabilistic guarantee on approximations. Moreover, for some regular graphs we obtain explicit values for the counting problem. For example, we show that every 4-regular n-node graph with large girth has approximately $(1.494\ldots)^n$ independent sets, and in every r-regular graph with n nodes and large girth the number of $q \geq r+1$-proper colorings is approximately $[q(1-\frac{1}{q})^{\frac{r}{2}}]^n$, for large n. In statistical physics terminology, we compute explicitly the limit of the log-partition function. We extend our results to random regular graphs. Our explicit results would be hard to derive via the Markov chain method.

1 Introduction

Counting is a natural counterpart to a combinatorial optimization problem. The typical set up involves counting the number of feasible solutions to some combinatorially constrained problem. The most widely studied such problems involve counting the number of solutions to a bin packing problem [JS97], counting the number of independent sets (also known as hard-core model in statistical physics) [LV97],[DGJ04], matchings [JS97], proper colorings in graphs (Potts model in statistical physics) [DGJ04],[DFHV04], volume of a convex body [DaRK91],[KLS97], [LV03], permanent of a matrix (counting the number of full matchings of a bi-partite graph) [Val79],[JS89], [JSV04], [JS97], [BSVV] etc. Typically, the set of feasible solutions is exponentially large and exhaustive search is computationally prohibited. This complexity appears to be fundamentally unavoidable, Valiant [Val79]. Modulo a complexity theoretic conjecture, the problems in $\#P$ do not admit polynomial time algorithms, and thus research focused on approximation algorithms. Here the most powerful method comes from the theory of rapidly mixing Markov chains. The typical setup involves relating counting problem to a sampling problem via certain telescoping trick (see for example identity (1) below) and then computing some marginal probabilities

*Department of Mathematics and Mathematical Statistics Chalmers University of Technology, SE-412 96 Gothenburg, Sweden, e-mail:antar@math.chalmers.se

†IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, e-mail: gamarnik@watson.ibm.com

using sampling technique. The main technical challenge is establishing that the underlying Markov chain mixes in polynomial time (rapid mixing). The scope of Markov chains for which rapid mixing has been established includes such notable breakthrough results as Jerrum and Sinclair's [JS89], and Jerrum, Sinclair and Vigoda's [JSV04] proof of rapid mixing of a Markov chain related to permanents, and Dyer, Frieze and Kannan [DaRK91] proof of rapid mixing of a Markov chain related to computing the volume of a convex body. Subsequent improvements in running time for computing volumes have been established in Kannan, Lovasz and Simonovits [KLS97] and Lovasz and Vempala [LV03]. Somewhat closer to the topic of this paper, Luby and Vigoda [LV97] showed that a Markov chain related to counting independent sets is rapidly mixing, when the underlying graph has degree at most 4.

A natural extension of the counting problem is (exponentially) weighted counting, that is computing the partition function. Partition function is a fundamental object in statistical physics and thus the connection between the counting and statistical physics is well known. There are many results in statistical physics literature on computing partition functions in various statistical physics models, but unfortunately, most of these results are not rigorous and involve what is known as replica-symmetry and replica symmetry breaking cavity method also known as replica symmetry breaking Ansatz [MPV87]. The process of rigorization of these spectacular but unproven results by physicists was undertaken relatively recently in mathematics: Talgrand [Tal03] proved the validity of the Parisi formula for the partition function limit of a Sherrington-Kirpatrick's model. Also Talagrand [Tal01] proved the existence and showed a method for computing the partition function limit of a random K-SAT problem in an appropriately defined high temperature regime. However, the process of building a full mathematical picture of the cavity and replica-symmetry methods is still largely under way.

In this paper we propose new methods for counting the number of independent sets and colorings (computing the partition function) in low degree graphs with large girth. In particular we propose a simple polynomial time algorithm for computing approximately the number of independent sets in graphs with maximum degree ≤ 4 and large girth. Similarly, for every q we propose a simple computable expression for the number of proper q-colorings of any graph with maximum degree $r \leq q-1$ and large girth.

On a negative side our algorithms only approximate exponents of the partition function: for every $\epsilon > 0$ we compute ϵ-approximation of the log-partition function (free energy). Also our computation time, while polynomial in the size of the graph, is not polynomial in ϵ. Thus our algorithm is PAS (Polynomial Time Approximation Scheme) as opposed to FPRAS (Fully Polynomial Time Randomized Approximation Scheme) as is typically established using Markov chains method. But there are two crucial advantages to our method. First, our algorithms are deterministic and do not suffer from sampling error. Second, in special cases involving regular graphs we obtain the values of the partition function *explicitly*. For example we show that in every 4-regular graph with n nodes and large girth, the number of independent sets is approximately $(1.494\ldots)^n$ *irrespectively of the graph!* Precisely, we show that the logarithm of the number of independent sets divided by n approaches $\log(1.494\ldots)$ as girth increases. The class of regular graphs with large girth is very rich and the fact that the number of independent sets is the same in all of them is an interesting by-product of our analysis. The value $1.494\ldots$ is a numeric approximation of a solution to a certain fixed-point equation. We obtain similar limiting numeric values for the case of r-regular graphs when $r = 2, 3, 4, 5$. For the problem of counting the number of proper colorings, we show that for every constant $q \geq r+1$, the number of q colorings in every r-regular graphs with large girth is approximately $(q(1-\frac{1}{q})^{\frac{r}{2}})^n$, when n is large. We note, that our results allow both q and r to be arbitrarily small. All of the known results for counting which are based on Markov chain method require q/r to be at least a large positive constant [DFHV04].

The main technical approach underlying our results is the progress in understanding properties of Gibbs distributions on regular infinite trees for independent sets, coloring, Ising and some other related

models in the context of *correlation decay* and the connection of thereof to the uniqueness of Gibbs measure. We use this stream of work to propose a different method for computing marginal probability featuring in cavity equation (1) below. In one of the earliest results in this area, Kelly [Kel85] established the following phase transition property for independent set on infinite r-regular trees: the probability that a root of the tree belongs to an independent set selected according to the Gibbs measure is asymptotically independent from the finite depth boundary of a tree, provided that inverse temperature λ is sufficiently small. The "counting" case $\lambda = 1$ satisfies this condition for $r \leq 5$ but breaks down for larger r. A recent extension of this result to general Galton-Watson type random trees and Erdos-Renyie type random graphs was done by Bandyopadhyay [Ban]. Similar uniqueness property is also known for Ising model [Geo88] and recently was established for coloring in the case of $q \geq r + 1$ colors by Jonasson [Jon02], closing an open problem posed earlier by Brightwell and Winkler [BW02]. The correlation decay property (long-range independence) featured lately very prominently in a variety of contexts including Aldous' proof of the ζ_2-limit for the random assignment problem [Ald01], bivariate uniqueness and endogeny of recursive distributional equations in Aldous and Bandyopadhyay [AB05], Bandyopadhyay [Ban02], Bandyopadhyay [Ban], Warren [War05], the local weak convergence properties Aldous and Steele [AS03], Gamarnik, Nowicki and Swirscsz [GNSa],[GNSb], Gamarnik [Gam04], and the problems of reconstruction on a tree, Mossel[Mos04], Yet, the importance of the correlation decay property for the uniqueness of Gibbs distribution was well recognized long time ago in the fundamental works by Dobrushin [Dob70] dating back to 70's. While Dobrushin's work was conducted primarily for lattices, there is a recent extension of this work by Weitz [Wei05] to more general graphs.

In this paper we establish the correlation decay property for independent sets, similar to the one considered by Kelly [Kel85] but for an arbitrary (not necessarily regular) tree with maximum degree at most 4. This property coupled with the cavity trick (1) almost immediately leads to a simple algorithm for computing approximately the partition function for independent sets. The corresponding algorithm for colorings is obtained by a simple extension of the Jonasson's [Jon02] uniqueness theorem for colorings. Methodologically, our approach consists of implementations of the following 3 steps. First computing appropriate marginal probabilities on a tree. This step typically involves a very simple recursive type computation. Then showing that the boundary has a vanishing impact on this marginally probability (correlation decay). Finally, the correlation decay is used to project the results of computation of marginal probabilities to non-tree graphs with locally tree-like structure.

Our explicit results for regular graphs are obtained by explicit computations of marginal probabilities for regular trees. An additional technical difficulty is the fact that the cavity step "destroys" the regularity of the graph. A simple trick introduced by Mezard and Parisi [MP05], (see also Rivoire et.al [RBMM04]) fixes this problem via some "rewiring" step. The regime corresponding to the correlation-decay property in our sense, is called a *liquid phase*. Our results then can be viewed as a rigorous treatment of liquid phase solution for independent sets model. Thus our work strengthens further an interesting and intriguing connection between the statistical physics and the theory of algorithms.

The rest of the paper is organized as follows. In the following section we provide the necessary background and definitions. Main results and their extensions, including the extensions to random regular graphs are presented in Section 3. Proofs are derived in Sections 4,5,6. Some conclusions and open problems are presented in the Section 7.

2 Notations and basics

Throughout the paper we consider a simple graph G with the node set $V = \{v_1, \ldots, v_n\}$ and edge set $E = \{e_1, \ldots, e_m\}$. We also write $n = n(G) = |V|$ for the number of nodes in the graph. With some

abuse of notation we will be writing $v \in G$, if node v belongs to the node set V of the graph G. For every $v \in G$, $r(v) = r(v, G)$ denotes the degree of v in G. $N(v, G)$ denotes the set of neighbors of v in G. The maximum degree and the girth (size of the smallest cycle) of G are denoted by $r = r(G) = \max_{1 \le k \le n} r(v_k)$ and $g = g(G)$ respectively. Let $\mathcal{G}_0(n, g, r)$ be the set of all degree-r graphs G with n nodes and girth at least g. Let also $\mathcal{G}(n, g, r)$ be the set of all r-regular graphs G with n nodes and girth at least g. Typically, we will be considering graphs with constant r, but girth diverging to infinity as a function of n. For every positive integer t and every node v_i, we denote by $T(v_i, t)$ the depth-t neighborhood of v_i – the set of nodes reachable from v_i by paths of lengths at most t. Clearly $g > 2t$ implies that $T(v_i, t)$ is a tree for every node v_i. A set $I \subset V$ is independent (stable) if no two nodes of I share an edge. $\mathcal{I} = \mathcal{I}(G)$ denotes the set of all independent sets in G. A proper coloring $C \in \mathcal{C}(q)$ is an assignment $C : V \to \{1, \dots, q\}$ of nodes V to colors $1, 2, \dots, q$ such that no two nodes which share an edge are assigned to the same color. For every $q \in \mathbb{N}$, $\mathcal{C}(q, G) = \mathcal{C}(q)$ denotes the set of all proper colorings of the nodes of G by colors $1, 2, \dots, q$. Throughout the paper we will only consider the case $q \ge r+1$. Then, as is well-known (and straightforward to show), the set $\mathcal{C}(q)$ is non-empty. In statistical physics literature it is common to call independent sets hard-core model and call colorings q-state Potts model [Geo88]. There is a way of defining a general model which simultaneously includes the model for independent sets and colorings by means of graph homomorphisms. This formalism has been used in a variety of papers [DGJ04], [BW04a]. Here, for simplicity we do not resort to this formalism.

A classical object in statistical physics is Gibbs probability distribution on the sets $\mathcal{I}, \mathcal{C}(q)$. Fix $\lambda > 0, \lambda_j, 1 \le j \le q$ called activity parameters. The Gibbs distribution on the set $\mathcal{I}$ assigns a probability proportional to $\lambda^{|I|}$ to each independent set I. More precisely,

$$\mathbb{P}(\boldsymbol{I} = I) = \frac{\lambda^{|I|}}{Z(\lambda)},$$

where $\boldsymbol{I}$ is the random (with respect to Gibbs measure) independent set, and $Z(\lambda) = Z(\lambda, G) = \sum_{I \in \mathcal{I}} \lambda^{|I|}$, the normalizing constant, is called the partition function. λ is called inverse temperature and the quantity $\log Z(\lambda)$ is also called *free energy*. In order to emphasize the underlying graph, sometimes we will denote the Gibbs measure by $\mathbb{P}_G(\cdot)$. When $\lambda = 1$, $Z(\lambda, G) = Z(1, G) = |\mathcal{I}|$ and the Gibbs distribution is simply the uniform distribution on the set of all independent sets.

There exists a way to represent the partition function $Z(\lambda, G)$ in terms of marginals of the Gibbs measure in the following sense. Let $G_0 = G$ and $G_k = G \setminus \{v_1, \dots, v_k\}, k = 1, 2, \dots, n$.

Proposition 1 *The following relation holds*

$$\frac{Z(\lambda, G_k)}{Z(\lambda, G_{k-1})} = \mathbb{P}_{G_{k-1}}(v_k \notin \boldsymbol{I}). \tag{1}$$

As a result,

$$Z(\lambda, G) = \prod_{k=1}^{n} \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I}). \tag{2}$$

This proposition is well known and is used for Markov chain based approximation algorithms for counting. We provide the proof for completeness. For convenience we assume that a partition function of an empty graph is equal to the unity.

Proof : The proof is obtained by considering a telescoping product

$$Z(\lambda, G) = \prod_{k=1}^{n} \frac{Z(\lambda, G_{k-1})}{Z(\lambda, G_k)}$$

and observing

$$\mathbb{P}_{G_{k-1}}(v_k \notin \boldsymbol{I}) = \frac{\sum_{I \in \mathcal{I}(G_{k-1}): v_k \notin I} \lambda^{|I|}}{Z(\lambda, G_{k-1})} = \frac{Z(\lambda, G_k)}{Z(\lambda, G_{k-1})}.$$

■

For the case of coloring, the Gibbs distribution on the set $\mathcal{C}(q)$ of proper colorings is introduced similarly as

$$\mathbb{P}(\boldsymbol{C} = C) = \frac{\prod_{1 \le j \le q} \lambda_j^{|C_j|}}{Z(\lambda)},$$

where $\boldsymbol{C}$ is the (Gibbs) random coloring and $\lambda = (\lambda_1, \ldots, \lambda_q)$ is a fixed vector of activity parameters, $C_j = \{v \in V : C(v) = j\}$, and $Z(\lambda) = Z(\lambda, G) = \sum_{C \in \mathcal{C}(q)} \prod_{1 \le j \le q} \lambda_j^{|C_j|}$ is again the normalizing partition function. Again the special case $\lambda_j = 1, 1 \le j \le q$ corresponds to the uniform distribution on the set $\mathcal{C}(q)$ of proper q-colorings. In this paper we focus exclusively on this special case and use notation $Z(q, G)$ or $Z(G)$ instead. The corresponding analogue of Proposition 1 is somewhat more complicated. For a random coloring $\boldsymbol{C}$ selected according to the Gibbs distribution and for any subset of nodes A, denote by $\boldsymbol{C}(A)$ the set of colors assigned to A. In particular, $\boldsymbol{C}(N(v_k, G_{k-1}))$ is the set of colors used by coloring $\boldsymbol{C}$ for the neighbors of the node v_k in the graph G_{k-1}. We will also write $\boldsymbol{C}(v)$ for $\boldsymbol{C}(\{v\})$ for every node $v \in G$. Again for convenience we assume that the number of proper q-colorings of an empty graph is equal to unity.

Proposition 2 *The following relation holds*

$$\frac{Z(q, G_{k-1})}{Z(q, G_k)} = q - \mathbb{E}_{G_k}\big[|\boldsymbol{C}(N(v_k, G_{k-1}))|\big]. \tag{3}$$

As a result,

$$Z(q, G) = \prod_{k=1}^{n} \Big[q - \mathbb{E}_{G_k}\big[|\boldsymbol{C}(N(v_k, G_{k-1}))|\big]\Big]. \tag{4}$$

Proof : The second part is obtained again by considering a telescoping product $Z(q, G) = \prod_{1 \le k \le n} \frac{Z(q, G_{k-1})}{Z(q, G_k)}$. To prove the first part we observe that

$$Z(q, G_{k-1}) = \sum_{1 \le m \le r(v_k, G_{k-1})} (q - m) \Big|\{C \in \mathcal{C}(G_k) : C(N(v_k, G_{k-1})) = m\}\Big|$$

where we simply observe that if the coloring C uses m colors for the neighbors of v_k in G_{k-1} then there are $q - m$ colors left for v_k itself. Then we divide both parts by $Z(q, G_k)$ and observe that

$$\sum_{1 \le m \le r(v_k, G_{k-1})} m \frac{\Big|\{C \in \mathcal{C}(G_k) : C(N(v_k, G_{k-1})) = m\}\Big|}{Z(q, G_k)} = \mathbb{E}_{G_k}[|\boldsymbol{C}(N(v_k, G_{k-1}))|].$$

■

3 Problem formulation and results

The enumeration (counting) problem we are concerned with in this paper is of computing approximately the sizes of the sets $\mathcal{I}$ and $\mathcal{C}(q)$. Specifically, we are interested in approximating the exponents corresponding to the cardinalities of these sets:

Definition 1 *Value $\alpha > 0$ is defined to be ϵ-approximation of the log-partition function $\log Z(\lambda, G)$ if*

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

where $\epsilon > 0$ is the error tolerance.

Given a family of graphs $\mathcal{G}$, an algorithm $\mathcal{A}$ is said to be Polynomial Approximation Scheme (PAS) for computing the log-partition function if for every $G \in \mathcal{G}$ it produces an ϵ-approximation of $\log Z(G)$ in time which is polynomial in n.

The Markov chain based approach for solving the counting problems typically provides approximation for the partition function itself and not just a logarithm of the partition function (as our approach does). Also it typically runs in time which is also polynomial in ϵ^{-1}. Thus it is called Fully Polynomial Randomized Approximation Scheme (FPRAS). On the other hand it provides approximation only with some probabilistic guarantee. We stress that the algorithms proposed in this paper provide deterministic guarantee, and thus are PAS, albeit the dependence on ϵ can be exponential. A natural intersection of two classes is Fully Polynomial Approximation Scheme (FPAS). The difference between different types of approximations is non-trivial and is not fully understood. For example, it is yet not clear that FPAS is always possible whenever FPRAS is possible. In fact Dyer, Goldberg and Jerrum [DGJ04] provide an evidence to the contrary.

An (infinite) family of graphs $\mathcal{G}$ is defined to have *large girth* if there exists an increasing function $f : \mathbb{N} \to \mathbb{N}$ such that $\lim_{s\to\infty} f(s) = \infty$ and for every $G \in \mathcal{G}$ with n nodes

$$g(G) \geq f(n).$$

3.1 Counting independent sets and colorings

Our first result establishes existence of PAS for computing the logarithm of the number of independent sets in graphs.

Theorem 1 *For every family $\mathcal{G}$ of graphs G with maximum degree $r \leq 4$ and large girth, the problem of computing $\log Z(\lambda, G)$ when $\lambda = 1$ is PAS.*

We have noted in the introduction that a Markov chain based FPRAS has been established by Luby and Vigoda [LV97] for all graphs with maximum degree at most 4. We do not know whether these apparently similar restrictions are merely a coincidence or not.

Our corresponding result for counting proper colorings does not require any upper bound on the maximum degree. Also it is more explicit and its algorithmic implication is immediate. In Section 5, we do though describe an algorithm for completeness.

Theorem 2 *Given constants $q \geq r+1$, the number of q-coloring of graphs $G \in \mathcal{G}_0(n, g, r)$ satisfies*

$$\lim_{g\to\infty} \sup_{G\in\mathcal{G}_0(n,g,r)} \left| \frac{\log Z(q, G)}{n} - \frac{1}{n} \sum_{1\leq k\leq n} \log \left[q(1-\frac{1}{q})^{r(v_k, G_{k-1})} \right] \right| = 0.$$

In particular, for every family $\mathcal{G}$ of graphs G with maximum degree r and large girth, the problem of computing $\log Z(q, G)$ is PAS.

Note that the bound in theorem above does not put any lower bound restriction on the number of nodes n. This is because the quality of approximation is completely controlled by the girth size. Implicitly, however, there is a trivial restriction, since when $n < g$, the graph has in fact infinite girth, namely, it is a tree. In this case, it can be verified directly, that the expression Z is exact number of colorings.

Our next results provide explicit estimates for the cardinality of the number of independent sets $\mathcal{I}$ and colorings $\mathcal{C}(q)$ in the special case of regular graphs with high girth.

Theorem 3 *Suppose $\lambda < (r-1)^{r-1}/(r-2)^r$. Then the partition function $Z(\lambda, G)$ corresponding to independent sets satisfies*

$$\lim_{g\to\infty} \sup_{G\in\mathcal{G}(n,g,r)} \Big| \frac{\log Z(\lambda, G)}{n} - \log\big(x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}}\big)\Big| = 0.$$

When $r = 2, 3, 4, 5$ and $\lambda = 1$, the corresponding limits for $n^{-1}\log|\mathcal{I}(G)|$ are respectively, $\log 1.618\ldots$, $\log 1.545\ldots$, $\log 1.494\ldots$ *and* $\log 1.453\ldots$.

Remarks : One important corollary of this result is that the asymptotic value of the log-partition function (limit of free energy) is the same for every r-regular graph with large girth. In particular, this result validates the non-rigorous statistical physics approach for computing free energy, where only locally-tree like structure and regularity is used in computation of free energy. Such insensitivity result cannot be obtained by the Markov Chain sampling technique.

We now state our main results for coloring. As we already mentioned, we only consider the special case $\lambda_j = 1, 1 \le j \le q$, that is the problem of counting the number of colorings. The reason for this limitation will be apparent when we discuss the recent result by Jonasson [Jon02].

Theorem 4 *For every $q \ge r+1$, the number of q-colorings of graphs $G \in \mathcal{G}(n,g,r)$ satisfies*

$$\lim_{g\to\infty} \sup_{G\in\mathcal{G}(n,g,r)} \Big| \frac{\log Z(q, G)}{n} - \log\Big[q\big(1-\frac{1}{q}\big)^{\frac{r}{2}}\Big]\Big| = 0.$$

As an immediate corollary of Theorem 4 we obtain that for every constant $\alpha \ge 1$, the number of $q = \lfloor \alpha r \rfloor + 1$ colorings of graphs $G \in \mathcal{G}(n,g,r)$ is approximately $(qe^{-\frac{1}{2\alpha}})^n$ as $g, r \to \infty$. Recently Bezakova, et.al [BSVV] obtained the following lower bound on $|\mathcal{C}(q,G)|$ in arbitrary n-node graph with maximum degree r: $|\mathcal{C}(q,G)| \ge (q - r(1-e^{-1}))^n$. Thus, when r is large and $q = \alpha r$ for some constant α, their bound becomes approximately $(q(1-\alpha^{-1} + (\alpha e)^{-1})^n$. It is not hard to see that our lower bound is strictly superior. For example, when $\alpha = 1$, their bound gives approximately $(qe^{-1})^n$ colorings, whereas, per our result, the correct limiting value (in log scale) is $(q/\sqrt{e})^n$. Of course out tight estimate comes at a cost of the large girth requirement.

3.2 Applications to random regular graphs

Random graphs are obtained by drawing a graph from some family of graphs at random according to some (typically uniform) distribution. Specifically, an r-regular n-node random graph $G_r(n)$ is obtained by selecting an r-regular graph uniformly at random from the set of all r-regular graphs on n-nodes. An important feature of such a regular graph is that the number of small cycles is small. In particular, for every constant C the expected number of size-C cycles is $O(1)$ in terms of the number of nodes n, [JŁR00]. Thus, *essentially* such graphs have a large girth and we may expect that our results for regular graphs with large girth extend to this class of graphs. It is indeed the case as we state below. The derivation of these results is very similar to the one used for the class $\mathcal{G}(n,g,r)$.

Theorem 5 *For every r and every $\lambda < (r-1)^{r-1}/(r-2)^r$, the (random) partition function $Z(\lambda, G_r(n))$ of a random r-regular graph $G_r(n)$ corresponding to the Gibbs distribution on independent sets satisfies*

$$\frac{\log Z(\lambda, G_r(n))}{n} \to \log\left[x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}}\right],$$

with high probability (w.h.p.), as $n \to \infty$, where x is the unique positive solution of $x = 1/(1+\lambda x^{r-1})$. In particular, when $r = 2, \ldots, 5$ and $\lambda = 1$, $\log Z(\lambda, G_r(n))/n$ converges w.h.p. to $\log 1.618\ldots$, $\log 1.545\ldots$, $\log 1.494\ldots$ and $\log 1.453\ldots$, respectively, as $n \to \infty$.

Our corresponding result for colorings is as follows.

Theorem 6 *For every r and every $q \geq r+1$, the (random) partition function $Z(q, G_r(n))$ of a random r-regular graph $G_r(n)$ corresponding to the uniform distribution on proper q-colorings satisfies*

$$\frac{\log Z(q, G_r(n))}{n} \to \log\left[q\big(1-\frac{1}{q}\big)^{\frac{r}{2}}\right].$$

w.h.p. as $n \to \infty$.

Theorem 6 is in fact not new. Using the second moment method it was established in [AM04], that that logarithm of the number of q colorings of a graph $G_r(n)$ divided by n converges w.h.p. to $\log\left[q\big(1-\frac{1}{q}\big)^{\frac{r}{2}}\right]$, matching our expression. In fact the range for q for which this is the case includes $q < r$. However, the (second moment) argument relies strongly on randomness of the graph. We stress that our general result Theorem 4 holds for every regular graph with large girth.

4 Counting independent sets

The key method for obtaining the results in this paper is establishing a very strong form of correlation decay, appropriately defined. Correlation decay is one of the key concepts in statistical physics which has been used to established the uniqueness of Gibbs distribution on infinite graphs (on finite graphs Gibbs distribution is unique by definition). These questions of uniqueness and correlation decay have been considered primarily in on regular trees. Here we reconstruct some of these results and extend them to non-regular trees. A strong form of correlation decay which we will establish will then be used to project our results to arbitrary graphs with large girth (and additional restrictions dictated by a particular context).

4.1 Independent sets on trees and correlation decay

Let T be an arbitrary tree with depth at most t. That is the distance from the root (denoted v_0) to any other node $v \in T$ is at most t. Denote by $B(T)$ the boundary of the tree – the set of nodes with distance exactly t from the root. Any function $b : B(T) \to \{0,1\}$ is called a boundary condition b. When $B(T)$ is empty the boundary condition is not defined. We think of boundary condition as conditioning on which nodes on the boundary belong to an independent set (corresponding value is 1) and which do not (value is zero). In particular, for any boundary condition b, we denote by $\mathbb{P}(v_0 \in \boldsymbol{I}|b)$ the probability of the event "v_0 belongs to the random independent set $\boldsymbol{I}$", conditioned on the event $\{v \in B(T) : v \in \boldsymbol{I}\} = \{v \in B(T) : b(v) = 1\}$, with respect to the Gibbs measure. Denote by $\mathcal{B}(T)$ the set of all boundary conditions b on T, and denote by $\mathcal{T}(t,r)$ the set of all trees with maximum degree at most r and depth at most t.

Our first result establishes the key correlation decay property of Gibbs distributions of independent sets on trees with maximum degree at most 4.

Proposition 3 *The following bounds holds for every $t \geq 2$, $T \in \mathcal{T}(t,4), b, b_1, b_2 \in \mathcal{B}(T)$*

$$\frac{1}{2} \leq \mathbb{P}(v_0 \notin \boldsymbol{I}|b) \leq \frac{8}{9}. \tag{5}$$

and

$$\Big|\mathbb{P}(v_0 \notin \boldsymbol{I}|b_1) - \mathbb{P}(v_0 \notin \boldsymbol{I}|b_2)\Big| \leq (.9)^{t-2}. \tag{6}$$

where $\mathbb{P}(\cdot)$ is with respect to the Gibbs distribution with $\lambda = 1$.

Moreover, given λ satisfying $\lambda < (r-1)^{r-1}/(r-2)^r$, let x be the unique non-negative solution of the equation $x = 1/(1+\lambda x^{r-1})$. Suppose all the nodes of T except for leaves and the root have degree r, and suppose the root has degree $r-1$. Then for all $b \in \mathcal{B}(T)$

$$\big|\mathbb{P}(v_0 \notin \boldsymbol{I}|b) - x\big| \leq \alpha^t, \tag{7}$$

for some constant $\alpha = \alpha(\lambda) < 1$. If, on the other hand, all the nodes except for leaves, have degree r (including the root), then

$$\big|\mathbb{P}(v_0 \notin \boldsymbol{I}|b) - \frac{1}{2-x}\big| \leq \alpha^t, \tag{8}$$

for the same constant α.

Remark : The second part of the proposition is a known result established first in Kelly [Kel85]. and we simply refer to Kelly's work for the proof. See also [BW04b] (where w corresponds to $1/x - 1$), and Bandyopadhyay [Ban] where the latter work is concerned with the extension of Kelly's result to general Galton-Watson type random trees. The constant $\alpha(\lambda)$ approaches unity as λ approaches $(r-1)^{r-1}/(r-2)^r$ and can expressed explicitly, but this is not required for our paper.

Proof : We fix a tree $T \in \mathcal{T}(t,r)$ and activity λ. Denote by $v_1, \ldots, v_k, k \leq r$ the neighbors $N(v_0, T)$ of the root. This includes the possibility $k = 0$ (the tree consists of only node v_0). For every node $v \in T$, $T(v)$ denotes the subtree rooted at v not containing v_0, and $b(T(v))$ denotes the natural restriction of a boundary condition $b \in \mathcal{B}(T)$ to $T(v)$. For every node v, let $T(v|b)$ be the tree obtained by deleting the leaves $v' \in T(v)$ which have value $b(v') = 1$ as well as their parent nodes. Let $J = I \cap T(v|b)$. It is immediate that for every independent set $I \subset T$, its Gibbs probability with boundary condition b is

$$\mathbb{P}_T(\boldsymbol{I} = I | I \cap B(T) = b) = \mathbb{P}_{T(v|b)}(\boldsymbol{I} = J) = \frac{\lambda^{|J|}}{\sum_{J' \in \mathcal{I}(T(v|b))} \lambda^{|J'|}},$$

Using convention $\mathbb{P}_{1 \leq j \leq k} = 1$ when $k = 0$, we obtain

$$Z(\lambda, T(v_0|b)) = \sum_{I \in \mathcal{I}(T(v_0|b))} \lambda^{|I|} = \prod_{1 \leq j \leq k} \Big(\sum_{I \in \mathcal{I}(T(v_j|b))} \lambda^{|I|} \Big) + \lambda \prod_{1 \leq j \leq k} \Big(\sum_{I \in \mathcal{I}(T(v_j|b)), v_j \notin I} \lambda^{|I|} \Big)$$

We recognize that

$$\frac{\prod_{1 \leq j \leq k} \Big(\sum_{I \in \mathcal{I}(T(v_j|b))} \lambda^{|I|} \Big)}{Z(\lambda, T(v_0|b))} = \frac{\prod_{1 \leq j \leq k} Z(\lambda, T(v_j|b))}{Z(\lambda, T(v_0|b))} = \mathbb{P}_{T(v_0)}(v_0 \notin \boldsymbol{I}|b)$$

Using the previous expression for $Z(\lambda, T(v_0|b))$, we obtain

$$\mathbb{P}_{T(v_0)}(v_0 \notin \boldsymbol{I}|b) = \frac{1}{1 + \lambda \prod_{1 \leq j \leq k} \mathbb{P}_{T(v_j)}(v_j \notin \boldsymbol{I}|b)}. \tag{9}$$

Note, that similar recursion applies to any node v substituting the root v_0 by replacing T with $T(v)$. Specifically, take any node v which is a parent of a leaf in level t in a main tree T, if any exist. That is v is located on level $t-1$. It has $r(v)-1$ children which we denote by $v_1, \ldots, v_{r(v)-1}$ its children. For every child $v_j, j \leq r(v)-1$ (if there are any) the value $\mathbb{P}(v_j \notin \boldsymbol{I}|b)$ is either zero or one depending on whether $b(v_j) = 0$ or $= 1$. The recursive equation (9) implies that $\mathbb{P}_{T(v)}(v \notin \boldsymbol{I}|b) \in [(1+\lambda)^{-1}, 1]$.

Now, suppose that v is any node on level $t-2$ and suppose it has $r(v)-1$ children. Then applying the same recursion and the previously obtained bounds, we get

$$\frac{1}{1+\lambda} \leq \mathbb{P}(v \notin \boldsymbol{I}|b) \leq \frac{1}{1+\lambda(1+\lambda)^{-r(v)+1}} \leq \frac{1}{1+\lambda(1+\lambda)^{-r+1}}.$$

For every node v in level $t-2$ define $a(v) = 1/(1+\lambda)$ and $c(v) = 1/(1+\lambda(1+\lambda)^{-r+1})$ and now we obtain bounds on probability $\mathbb{P}(v \notin \boldsymbol{I}|b)$ nodes at lower levels. Given a node v in level $\tau \leq t-2$, suppose $\mathbb{P}(v \notin \boldsymbol{I}|b)$ belongs to an interval $[a(v), c(v)]$. Then for every node v with children nodes $v_1, \ldots, v_{r(v)-1}$ we obtain

$$a(v) = \frac{1}{1+\lambda\prod_{1\leq j\leq r(v)-1} c(v_j)} \leq \mathbb{P}(v \notin \boldsymbol{I}|b) \leq \frac{1}{1+\lambda\prod_{1\leq j\leq r(v)-1} a(v_j)} = c(v). \tag{10}$$

Also, inductively assuming $a(v_j) \geq 1/(1+\lambda), c(v_j) \leq 1/(1+\lambda(1+\lambda)^{-r+1}$, we obtain by the same argument as above that the same bounds hold for $a(v), c(v)$ for all the node v in levels up to $t-2$:

$$\frac{1}{1+\lambda} \leq a(v) \leq c(v) \leq \frac{1}{1+\lambda(1+\lambda)^{-r+1}}. \tag{11}$$

We note that these bounds only depend on the tree T but not the boundary condition b. We now show that , the length of the bounding interval $c(v) - a(v)$ is geometrically decreasing in as a function of the level of v in our special case of interest.

Lemma 1 *Suppose $r = 4, \lambda = 1$. Then for every node $v \in T$ in level τ, $c(v) - a(v) \leq (.9)^{t-2-\tau}$.*

Proof : The proof proceeds by reverse induction in τ starting with $\tau = t-2$. For $\tau = t-2$ the bound holds trivially from $0 \leq a(v), c(v) \leq 1$. Assume it holds for levels $\tau+1, \ldots, t-2$ and consider any node v in level τ with children $v_1, \ldots, v_k, 0 \leq k \leq r-1$. If $k = 0$ then $a(v) = c(v) = 1/(1+\lambda)$ and the bound holds trivially. Now suppose $k > 0$. Introduce function $f : [(1+\lambda)^{-1}, (1+\lambda(1+\lambda)^{-r+1})^{-1}]^k \to \mathbb{R}$ given by $f(z) = f(z_1, \ldots, z_k) = (1+\lambda\prod_{1\leq j\leq k} z_j)^{-1}$. We rewrite (10) as $f(c(v_1), \ldots, c(v_k)) = a(v) \leq c(v) = f(a(v_1), \ldots, a(v_k))$, where $a(v_j), c(v_j)$ satisfy the bounds in (11). Function f is differentiable on its domain. By mean value theorem, there exists $z \in [(1+\lambda)^{-1}, (1+\lambda(1+\lambda)^{-r+1})^{-1}]^k$ such that

$$\begin{aligned} c(v) - a(v) &= \nabla f(z)(a(v_1) - c(v_1), \ldots, a(v_k) - c(v_k)) \\ &\leq \|\nabla f(z)\|_1 \max_{1\leq j\leq k} |a(v_j) - c(v_j)| \\ &\leq \|\nabla f(z)\|_1 .9^{t-2-\tau+1}, \end{aligned}$$

where the last bound follows from the inductive assumption. It then suffices to prove that $\|\nabla f(z)\|_1 < .9$. We expand $\|\nabla f(z)\|_1$ as

$$\|\nabla f(z)\|_1 = \frac{\lambda\prod_{1\leq j\leq k} z_j \sum_{1\leq j\leq k} z_j^{-1}}{(1+\lambda\prod_{1\leq j\leq k} z_j)^2}.$$

We now resort to our specific assumption $r \leq 4, \lambda = 1$. The remainder of the proof is computer assisted. For given $k \leq 4$, consider a resolution .001 grid on the rectangle $[(1+\lambda)^{-1}, (1+\lambda(1+\lambda)^{-r+1})^{-1}]^k, 1 \leq$

$k \leq 4$. We note that the right end $(1+\lambda(1+\lambda)^{-r+1})^{-1}$ of the rectangle is largest when $r=4$, so we consider the set of vectors $z=(z_1,\ldots,z_k)$ of the form $z_j = .001m_j$, for some $m_j \in \mathbb{N}$ such that $1/2 = (1+\lambda)^{-1} \leq z_j \leq (1+2^{-3})^{-1}$ for all j. We have checked numerically using MATLAB that for every $k=2,3,4$ and every point z on this k-dimensional grid, the value of $\|\nabla f(z)\|_1$ is at most .8736. Specifically, the maximum values for $k=2,3,4$ (using rational computations) turn out to be $1089/2500 \approx .4356$, $109/165 \approx .6606$, $825/943 \approx .8749$, respectively. We now use first order Taylor approximation to argue that the maximums $\max \nabla f(z)$ over the domain of f are at most .9 for all $k=2,3,4$. For every z in the rectangle find any of its grid point approximation $\hat{z}=(\hat{z}_1,\ldots,\hat{z}_k)$, meaning $|z_j - \hat{z}_j| < .001$ (typically many such approximations exist and we choose any of them). Let $g = \|\nabla f\|_1$. We now show that for every two vectors z^1, z^2 which coincide in all the coordinates except for one, and such that $\|z^1 - z^2\| < .001$, we have

$$|g(z^1) - g(z^2)| < .013. \tag{12}$$

This results in $|f(z) - f(\hat{z})| < .013k \leq .013 \cdot 3 < .039$ and, combining with the bound on points on the grid we obtain that for every point z on the domain $\nabla f(z) < .8749 + .039 < .9$ and the proof of the lemma would be complete.

To estimate the difference $|g(z^1) - g(z^2)|$ we assume, w.l.g. that the two vectors differ in the first variable z_1. Applying second order Taylor expansion for the first variable z_1 we obtain that for some value θ between z_1^1 and z_1^2,

$$g(z^2) = g(z^1) + \frac{\partial g(z^1)}{\partial z_1}(z_1^2 - z_1^1) + \frac{1}{2}\frac{\partial_2 g(\theta)}{\partial^2 z_1}(z_1^2 - z_1^1)^2 \tag{13}$$

For convenience, denote generically $\prod_{2\leq j\leq k} z_j$ by A, $\prod_{2\leq j\leq k} z_j \sum_{2\leq j\leq k} z_j^{-1}$ by B, and $\prod_{2\leq j\leq k} z_j$ by C. Trivially, we have $A<1, B<k-1\leq 2, C<1$. We have $g(z) = \frac{Bz_1+C}{(1+Az_1)^2}$, and

$$\begin{aligned}\frac{\partial g(z)}{\partial z_1} &= \frac{B(1+Az_1)^2 - 2(Bz_1+C)(1+Az_1)A}{(1+Az_1)^4} \\ &= \frac{B(1+Az_1) - 2(Bz_1+C)A}{(1+Az_1)^3}\end{aligned}$$

Which in absolute value does not exceed $\max(B/(1+A)^2, 2(B+C)A/(1+A)^3)) \leq \max(B, 2(B+C)A) \leq 6$, using the bounds on A, B, C and $1+Az_1 > 1, 0 < z_1 < 1$. Then the absolute value of the second term in the sum in (13) is bounded by $6 \cdot .001 = .012$. We now bound the term corresponding to the second derivative, which find to be

$$\frac{\partial_2 g(z)}{\partial^2 z_1} = \frac{BA(1+Az_1)^3 - 2BA(1+Az_1)^3 - \Big(B(1+Az_1) - 2(Bz_1+C)A\Big)3(1+Az_1)^2 A}{(1+Az_1)^6}.$$

We very crudely upper bound the absolute value of $\frac{\partial_2 g(z)}{\partial^2 z_1}$ as

$$BA + 2BA + (B + 2(B+C)A)(3A) < 2 + 4 + (2+6)3 = 24,$$

again using the bounds $A<1, B<2, C<1, 0<z_1<1, Az_1+1>1$. Thus the third term in the sum (13) is upper bounded by $(1/2)12 \cdot .001^2 = 6 \cdot 10^{-4}$. Combining, we obtain from (13) and the obtained bounds on the first and second derivative, that $|g(z^1) - g(z^2)| < .012 + 6 \cdot 10^{-4} < .013$. We established (12). This completes the proof of the lemma. ∎

Application of the lemma to the root node v_0 yields, $c(v_0) - a(v_0) \leq (.9)^{t-2}$. Combining this with (10) applied to v_0 gives for every two boundary conditions b_1, b_2

$$\left|\mathbb{P}(v_0 \notin \boldsymbol{I}|b_1) - \mathbb{P}(v_0 \notin \boldsymbol{I}|b_2)\right| \leq c(v_0) - a(v_0) \leq (.9)^{t-2}.$$

This establishes (6) and completes the proof the first part of the proposition.

The second part of the proposition is the result already established by Kelly [Kel85] and we simply refer to his paper. ■

4.2 Algorithm and the proof of Theorem 1

Proposition 3 establishes the key correlation decay property for independent sets for trees with maximum degree at most 4. It shows that the marginal Gibbs probability at the root is asymptotically independent from the boundary. Equipped with this result and Proposition 1, we propose the following algorithm for estimating the number of independent sets of a given graph G.

Algorithm CountIND

INPUT: A graph G with a node set $v_1, \ldots, v_n$ and parameter $\epsilon > 0$.
BEGIN
1. Compute the girth $g(G)$. If $(.9)^{\frac{g(G)}{2}-2} \geq \epsilon$ compute $\mathcal{I}(G)$ by exhaustive enumeration. Otherwise
2. Set $G' = G$, $Z = 1$, $t = g(G)/2$.
3. Find any node $v \in G'$ and identify its depth$-t$ neighborhood $T(v)$ -- the set of all nodes at distance $\leq t$ from v.
4. Perform subroutine CountingTREE on $T(v)$ which results in some value $p(v)$. Set Z equal to $Zp^{-1}(v)$.
5. Set $G' = G' \setminus \{v\}$ and go to step 3.
END
OUTPUT: Z.

Subroutine CountingTREE

INPUT: A tree T with an identified root v and depth t.
BEGIN
1. Identify the nodes u in level t (if any exist) and set $p(u) = 1/2$.
FOR $l = t-1, t-2, \ldots, 0$
Identify a node u in level l (if any exist). If u has no children, set $p(u) = 1/2$. Otherwise set $p(u) = 1/(1+\prod p(u_i))$, where the product runs over children u_i of u in level $l+1$ and the values $p(u_i)$ were obtained in an earlier step.
END
OUTPUT: $p(v)$.

Proof : *Proof of Theorem 1.* We claim that the algorithm CountIND provides PAS. Fix a family of graphs $\mathcal{G}$ with maximum degree $r \leq 4$ and large girth, a graph $G \in \mathcal{G}$ and $\epsilon > 0$. The algorithm first checks whether $g(G) > 4 + 2\log(1/\epsilon)/\log(10/9)$. By definition there exists a finite number of graphs in $\mathcal{G}$ with girth $\leq 4 + 2\log(1/\epsilon)/\log(10/9)$ and their corresponding values of $\mathcal{I}$ can be found in constant time, where the constant depends on ϵ and the growth rate f of girth.

Otherwise the girth satisfies $(.9)^{\frac{g(G)}{2}-2} < \epsilon$ and in the remaining n steps of the algorithm the Gibbs marginal probability $\mathbb{P}(v_k \in \boldsymbol{I})$ is computed with respect to the depth $t = g(G)/2$ neighborhood $T(v_k)$

of the node v_k with respect to the graph G_{k-1}. By selection of t, $T(v_k)$ is a tree (the girth of each subgraph G_{k-1} is trivially at least $g(G)$). Let $B(T(v_k))$ be the boundary of $T(v_k)$ and consider the graph $\hat{G}_{k-1} = (G_{k-1} \setminus T(v_k)) \cup B(T(v_k))$, that is everything but the first $t-1$ levels of $T(v_k)$. Every independent set I which is a subset of $\hat{G}_{k-1}$ induces a boundary condition $b = b(I)$ on $T(v_k)$ via its intersection with $B(T(v_k))$. Let b_0 denote an empty boundary condition on $T(v_k)$ (also called free boundary). This corresponds to all independent sets I which do not intersect with $B(T(v_k))$. Then with respect to the tree $T(v_k)$ we have $\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}|b_0) = \mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I})$. We have for every independent subset $I \subset \hat{G}_{k-1}$ that $\mathbb{P}_{G_{k-1}}(v_k \notin \boldsymbol{I}|\boldsymbol{I} \cap \hat{G}_{k-1} = I) = \mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}|b(I))$ since $T(v_k)$ intersects with $\hat{G}_{k-1}$ only on $B(T(v_k))$. Proposition 3 implies that

$$\left|\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}|b_0) - \mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}|b(I))\right| < (.9)^{t-2} = (.9)^{\frac{g(G)}{2}-2} < \epsilon.$$

Then by summing over all possible realizations of I we obtain

$$|\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}) - \mathbb{P}_{G_{k-1}}(v_k \notin \boldsymbol{I})| < \epsilon.$$

The lower bound part of (5) gives $\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}) \geq 1/(1+\lambda) = .5$. Then

$$\begin{aligned}|\mathbb{P}^{-1}_{T(v_k)}(v_k \notin \boldsymbol{I}) - \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})| &= \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})\left|\frac{\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I}) - \mathbb{P}_{G_{k-1}}(v_k \notin \boldsymbol{I})}{\mathbb{P}_{T(v_k)}(v_k \notin \boldsymbol{I})}\right| \\ &< \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})\frac{\epsilon}{.5}.\end{aligned}$$

We conclude

$$\mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})(1-2\epsilon) \leq \mathbb{P}^{-1}_{T(v_k)}(v_k \notin \boldsymbol{I}) \leq \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})(1+2\epsilon)$$

The value $\mathbb{P}^{-1}_{T(v_k)}(v_k \notin \boldsymbol{I})$ is what algorithm CountTREE outputs as $p^{-1}(v)$. Therefore, applying Proposition 1, we have that Z, the product of these outputs satisfies

$$Z(1,G)(1-2\epsilon)^n = \prod_{k=1}^{n} \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})(1-2\epsilon)^n \leq Z \leq \prod_{k=1}^{n} \mathbb{P}^{-1}_{G_{k-1}}(v_k \notin \boldsymbol{I})(1+2\epsilon)^n = Z(1,G)(1-2\epsilon)^n.$$

Using $|\log(1-2\epsilon)| < 3\epsilon$ for sufficiently small ϵ, we obtain

$$\left|\frac{\log Z}{n} - \frac{\log Z(1,G)}{n}\right| < 3\epsilon.$$

Finally, we observe that since, by bounds (11) each element of the product Z belongs to the interval $[1+\lambda(1+\lambda)^{-r+1}, (\lambda+1)/\lambda] = [9/8, 2/1]$, then $\log Z/n \geq \log(9/8)$. Therefore

$$(1 - 3\epsilon \log^{-1}(9/8)) \leq \frac{\log Z}{\log Z(1,G)} \leq (1 + 3\epsilon \log^{-1}(9/8)).$$

Thus the algorithm CountIND is PAS for counting independent sets. ∎

4.3 Regular graphs and proof of Theorem 3

The second part of Proposition 3 provides an explicit limiting expression for the probability that a given node belongs to an independent set selected according to the Gibbs distribution. In this subsection we use it to obtain explicit asymptotics for the logarithm of the number of independent sets in regular

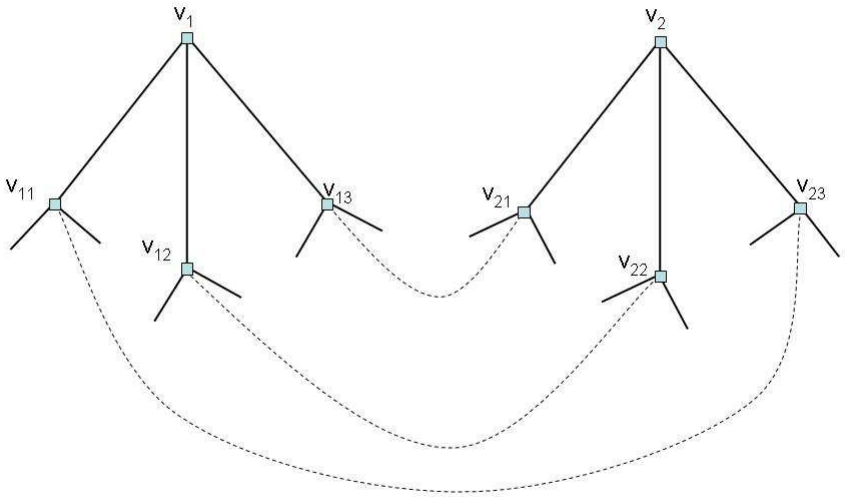

Figure 1: Rewiring on nodes v_1 and v_2

graphs. Theorem 1 provides a way in principle for computing number of independent sets in regular graph. The problem is, however, in the fact that the cavity step expressed in (1) destroys regularity: when node v_1 is removed, the remaining graph is no longer regular and it is not clear how to estimate product (2) explicitly. The help comes from a trick introduced by Mezard and Parisi [MP05], also used in [RBMM04] in the context of random regular graph. Given an n-node r-regular G fix any two nodes v_1, v_2 which are not neighbors, and do not have common neighbors (if there are any) and denote their non-overlapping neighbor sets by $v_{11}, \ldots, v_{1r}$ and $v_{21}, \ldots, v_{2r}$, respectively. Consider a modified graph G^o obtained by from G by deleting v_1, v_2 and connecting v_{1j} to $v_{2j}, j = 1, \ldots, r$ by an edge, see Figure 4.3 for an example with $r = 3$. The resulting graph is r-regular again. We call this operation "rewiring" or "rewire" operation. Rewiring was used in [MP05] and [RBMM04] was in a context of random regular graphs and was performed on two nodes selected randomly from the graph. The main question is whether we can relate the partition functions of the original and modified graphs and whether the resulting graph still has a sufficiently large girth, provided the original one does. The first issue has been addressed in [RBMM04] and is essentially a simple combination of type (1) arguments. The second issue was not addressed in [RBMM04] in a rigorous way. It was just postulated that the resulting graph again has a large girth if the two nodes are selected uniformly at random.

We begin by addressing the second issue first.

Lemma 2 *Given an n-node r-regular graph G, consider any integer $4 \leq g \leq g(G)$. The rewiring operation can be performed for at least $(n/2) - (2g+1)r^{2g}$ steps on pairs of nodes which are at least $2g+1$ distance apart. In every step the resulting graph is r-regular with girth at least g.*

Proof : In every step of the rewiring we delete two nodes in the graph. Thus when (if) we performed $t \leq (n/2) - (2g+1)r^{2g}$ successful rewiring steps, in the end we obtain a graph with at least

$n - 2((n/2) - (2g+1)r^{2g}) = 2(2g+1)r^{2g}$ nodes. Suppose in step $t \leq (n/2) - (2g+1)r^{2g}$ we have a graph G_t which is r-regular and has girth at least g. We claim that the diameter of this graph is at least $2g+1$. Indeed, if the diameter is smaller, then for a given node v any other nodes is reachable from v by a path with distance at most $2g$ and the total number of nodes is at most $\sum_{0\leq k\leq 2g} r^k < (2g+1)r^{2g}$ – contradiction. Now select any two nodes $v_1, v_2 \in G_t$ which are at the distance equal to the diameter of this graph, and thus are at least $2g+1$ edges apart. We already showed that the graph G_{t+1} obtained by rewiring G_t on v_1, v_2 is r-regular. It remains to show it has a girth at least g. Suppose, for the purposes of contradiction, G_t has girth $\leq g-1$ and $k \geq 1$ out of r newly created edges participate in creating a cycle with length $\leq g-1$. If $k=1$ and v_{1j}, v_{2j} is the pair creating the unique participating edge, then the original distance between v_{1j} and v_{2j} was at most $g-2$ by following a path on the cycle which does not use the new edge. But then the distance between v_1 and v_2 is at most $g < 2g+1$ – contradiction. Suppose there are $k > 1$ edges which create a cycle with length $\leq g-1$. Then there exists a path of length at most $(g-1)/k \leq (g-1)/2$ which uses only the original edges (the edges of the graph G_t) and connects a pair v, v' of nodes from the set $v_{11}, \ldots, v_{1r}, v_{21}, \ldots, v_{2r}$. If the pair is from the same set, for example $v = v_{1j}, v' = v_{1l}$, then, since these two nodes are connected to v_1, we obtain a cycle in G_t with length $(g-1)/2 + 2 < g$ – contradiction, since, by assumption $g > 3$. If these two nodes are from different sets, for example $v = v_{1j}, v' = v_{2l}$, then we obtain that the distance between v_1 and v_2 is at most $(g-1)/2 + 2 < 2g+1$ – again contradiction. We conclude that G_t has girth at least g as well. ∎

We now turn to the second problem of estimating the relative change of the partition function after rewiring. This relative change is called *energy shift* in [RBMM04]. First we provide an elementary analogue of (1).

Lemma 3 *Given an r-regular graph G, given $\lambda > 0$ and graph G^o obtained from G by rewiring on nodes $v_1, v_2 \in G$, the following relation holds*

$$\frac{Z(\lambda, G^o)}{Z(\lambda, G)} = \mathbb{P}_G(v_1, v_2 \notin \boldsymbol{I}) \mathbb{P}_{G\setminus\{v_1,v_2\}}(\wedge_{1\leq j\leq r}(v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I}))$$

where $v_{ij}, j = 1, \ldots, r$ is the set of neighbors of $v_i, i = 1, 2$ in G.

Proof : The proof is almost identical to the one of Proposition 1. The partition function $Z(\lambda, G^o)$ is obtained as a sum $\lambda^{|I|}$ over the set of independent subsets $I \subset V(G)$, which do not contain v_1, v_2 and which contain at most one of the two nodes v_{1j}, v_{2j} for each $j = 1, 2, \ldots, r$. ∎

We now obtain a very simple limiting expression for the probability in Lemma 3.

Lemma 4 *Given $r \in \mathbb{N}, \lambda < (r-1)^{r-1}/(r-2)^r$ and $\epsilon > 0$, there exists a sufficiently large constant $g = g(r, \epsilon, \lambda)$ such that for every graph G with girth $g(G) > g$, and for every pair of nodes $v_1, v_2 \in G$ at distance at least $2g+1$*

$$\left| \mathbb{P}_G((v_1, v_2 \notin \boldsymbol{I})) - \frac{1}{(2-x)^2} \right| < \epsilon, \tag{14}$$

and

$$\left| \mathbb{P}_{G\setminus\{v_1,v_2\}}\big(\wedge_{1\leq j\leq r} (v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I})\big) - (2x - x^2)^r \right| < \epsilon, \tag{15}$$

where $v_{ij}, j = 1, \ldots, r$ is the set of neighbors of v_i in G, $i = 1, 2$, and x is the unique solution of $x = 1/(1 + \lambda x^{r-1})$.

Proof : The proof consists of several steps, each ideologically very similar to the one for Theorem 1. Fix $\epsilon > 0$ and let $g = g(\epsilon, r, \lambda)$ be a large value to be specified later. Select $\alpha = \alpha(\lambda)$ is selected as in Proposition 3. We consider any r-regular graph with girth at least g and consider any two nodes v_1, v_2 in G at distance at least $2g+1$, if such two nodes exist. Consider depth $t = g/2$ neighborhoods $T(v_1), T(v_2)$. By the distance assumption, they do not intersect, and by the girth assumption, each neighborhood is a depth-t r-regular tree. First estimate the impact of deleting these nodes v_1, v_2 from G. That is we first take $G_1^o = G \setminus \{v_1, v_2\}$ and consider $Z(\lambda, G \setminus \{v_1, v_2\})/Z(\lambda, G)$. Then we will take G^o obtained by rewiring G on v_1, v_2 and estimate $Z(\lambda, G^o)/Z(\lambda, G \setminus \{v_1, v_2\})$.

Fix any independent set I on $\hat{G} = B(T(v_1)) \cup B(T(v_1)) \cup (G \setminus (T(v_1) \cup T(v_2)))$, where $B(T)$ is again the boundary of a tree T. Let $b_i = I \cap B(T(v_i)), i = 1, 2$. Let $\boldsymbol{I}$ be the random independent set in G selected according to the Gibbs distribution with parameter λ. We have by Gibbs property that

$$\begin{aligned}\mathbb{P}_G(v_1, v_2 \notin \boldsymbol{I} | \boldsymbol{I} \cap \hat{G} = I) &= \mathbb{P}_G(v_1 \notin \boldsymbol{I} | \boldsymbol{I} \cap \hat{G} = I) \mathbb{P}_G(v_2 \notin \boldsymbol{I} | \boldsymbol{I} \cap \hat{G} = I) \\ &= \mathbb{P}_{T(v_1)}(v_1 \notin \boldsymbol{I} | b_1) \mathbb{P}_{T(v_2)}(v_2 \notin \boldsymbol{I} | b_2)\end{aligned}$$

From the second part of Proposition 3

$$|\mathbb{P}_{T(v_i)}(v_i \notin \boldsymbol{I} | b_i) - \frac{1}{2-x}| < \alpha^t, \quad i = 1, 2,$$

which results in

$$|\mathbb{P}_G(v_1, v_2 \notin \boldsymbol{I} | \boldsymbol{I} \cap \hat{G} = I) - (\frac{1}{2-x})^2| \leq \alpha^t + \alpha^t \frac{1}{2-x}.$$

By summing over all the realizations of I we also obtain

$$|\mathbb{P}_G(v_1, v_2 \notin \boldsymbol{I}) - (\frac{1}{2-x})^2| \leq \alpha^t + \alpha^t \frac{1}{2-x}.$$

We take $t = g/2 = g(\epsilon, r, \lambda)$ sufficiently large, so that the absolute difference above is at most ϵ (note that the choice depends on α which in itself is controlled by λ). This concludes the proof of the first part.

Now consider $\mathbb{P}_{G_1^o}(\wedge_{1 \leq j \leq r}(v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I}))$. We take depth-$(t-1)$ neighborhoods of $v_{ij}, k = 1, 2, \ j = 1, \ldots, r$ and again observe that they are all non-intersecting trees because of the girth and distance between v_1 and v_2 assumption. By conditioning on the realizations I of a random independent set $\boldsymbol{I}$ in $\hat{G}_1 = (G_1^o \setminus \cup_{i,j} T(v_{ij})) \cup (\cup_{i,j} B(T(v_{ij})))$, letting $b_{ij} = I \cap B(T(v_{ij}))$ and using the same argument as above, we obtain

$$\begin{aligned}&\mathbb{P}_{G_1^o}\Big(\wedge_{1 \leq j \leq r} (v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I}) | \boldsymbol{I} \cap \hat{G}_1 = I \Big) \\ &\quad = \prod_{1 \leq j \leq r} \Big(\mathbb{P}_{T(v_{1j})}(v_{1j} \notin \boldsymbol{I} | b_{1j}) + \mathbb{P}_{T(v_{2j})}(v_{2j} \notin \boldsymbol{I} | b_{2j}) - \mathbb{P}_{T(v_{1j})}(v_{1j} \notin \boldsymbol{I} | b_{1j}) \mathbb{P}_{T(v_{2j})}(v_{2j} \notin \boldsymbol{I} | b_{2j}) \Big) \\ &\quad = \prod_{1 \leq j \leq r} \Big(1 - \mathbb{P}_{T(v_{1j})}(v_{1j} \in \boldsymbol{I} | b_{1j}) \mathbb{P}_{T(v_{1j})}(v_{1j} \in \boldsymbol{I} | b_{1j}) \Big)\end{aligned}$$

Again we use bound provided by Proposition 3

$$|\mathbb{P}_{T(v_{ij})}(v_{1j} \in \boldsymbol{I} | b_{ij}) - (1-x)| < \alpha^{t-1}, \quad i = 1, 2, \ j = 1, 2, \ldots, r,$$

(we recall that each tree $T(v_{ij})$ has depth $t-1$ and the root v_{ij} of this tree has degree $r-1$). We now take $t = g/2 = g(\epsilon, r, \lambda)/2$ sufficiently large so that

$$\left| \mathbb{P}_{G_1^o}\Big(\wedge_{1\le j\le r} (v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I}) | \boldsymbol{I} \cap \hat{G}_1 = I \Big) - (1-(1-x)^2)^r \right| < \epsilon.$$

By summing over all the realizations of I we obtain

$$\left| \mathbb{P}_{G_1^o}\Big(\wedge_{1\le j\le r} (v_{1j} \notin \boldsymbol{I} \vee v_{2j} \notin \boldsymbol{I}) \Big) - (2x-x^2)^r \right| < \epsilon.$$

■

Proof : *Proof of Theorem 3.* The proof is obtained by combining the results of Lemmas 2,3,4. From the last two lemmas, for every ϵ we can find $g = g(\epsilon, r, \lambda)$ sufficiently large so that for every graph G with girth at least $g+1$ and for every two nodes v_1, v_2 at distance at least $2g+1$, the graph G^o obtained from G by rewiring on v_1, v_2 satisfies, after simplifying $(2-x)^{-2}(2x-x^2)^r$ to $x^r(2-x)^{r-2}$, the following bounds.

$$(1-\epsilon)x^r(2-x)^{r-2} \le \frac{Z(\lambda, G^o)}{Z(\lambda, G)} \le (1+\epsilon)x^r(2-x)^{r-2}.$$

Here we note that in order to combine the individual absolute differences (14) and (15), we need to take $g = g(\epsilon, r, \lambda)$ which is sufficiently large with taking x into account. But x itself depends only on λ. Therefore such g indeed exists. By Lemma 2, if the original graph G has n nodes, then the rewiring can be performed for at least $N = n/2 - C = n/2 - C(g,r) = n/2 - C(\epsilon, r, \lambda)$ steps, and at most $n/2$ steps, where constant $C = C(g,r) = (2g+1)r^{2g}$. Let G^* denote the graph obtained from G after N rewiring steps. Then from the bound above

$$(1-\epsilon)^{\frac{n}{2}-C}(x^r(2-x)^{r-2})^{\frac{n}{2}-C} \le \frac{Z(\lambda, G^*)}{Z(\lambda, G)} \le (1+\epsilon)^{\frac{n}{2}}(x^r(2-x)^{r-2})^{\frac{n}{2}}$$

Since the number of nodes in G^* is at most $2C$, then trivially $Z(\lambda, G^*) \le (1+\lambda)^{2C}$, then we obtain for sufficiently large $n(\epsilon, r, x, C) = n(\epsilon, r, \lambda)$, that for all $n \ge n(\epsilon, r, \lambda)$

$$\left| \frac{\log Z(\lambda, G)}{n} - \log x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}} \right| < 2\epsilon.$$

This concludes the proof of the first part of the theorem.

The case $\lambda = 1$ corresponds to the counting problem. We check that $(r-1)^{r-1}/(r-2)^r > 1$ only for $r = 2, 3, 4, 5$ and thus for these values we can obtain the asymptotics of the log-partition function, and we do so now.

In the special case $r = 2$ and $\lambda = 1$ we find that $x = \frac{\sqrt{5}-1}{2} \approx 0.6180$, derived from the golden ratio equation $x = 1/(1+x)$. Thus the total number of independent sets $\mathcal{I}(G)$ in every 2-regular graphs with large girth is $\approx (\frac{2}{\sqrt{5}-1})^n \approx (1.618\ldots)^n$. As a sanity check there is a simple way to check the validity of this answer, for example in a special case when the graph is an n-cycle. We note that for every node v on a cycle, if it belongs to the independent set, its right-hand side neighbor v' does not, but if v does not, then v' either belongs or does not belong to the independent set. It is a simple exercise to see that the number of independent sets which can be created on a path of length k starting from v and going to the right is

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{k-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The growth rate of this expression is determined by the largest eigenvalue of the matrix, which is the golden ration value $2/(\sqrt{5}-1)$. Thus on the path of length n the number of independent sets is $\approx (2/(\sqrt{5}-1))^n$. The number of independent sets on a cycle differs from this only by a constant factor (to adjust for a fact that the last node and the first node v do belong to the independent set at the same time).

When $r = 3, \lambda = 1$, the solution x to the equation $x = 1/(1+x^2)$ is found numerically to be $x = 0.682\ldots$. Thus $\mathcal{I}(G)$ for every 3-regular is $\approx (1.545\ldots)^n$. When $r = 4, \lambda = 1$, we find similarly that $\mathcal{I}(G)$ for every 4-regular is $\approx (1.494\ldots)^n$ and when $r = 5$ it is $\approx (1.453\ldots)^n$. This concludes the proof of Theorem 3. ■

5 Counting Colorings

The general approach for solving the problem of counting the number of proper colorings is the same as for independent sets. We establish correlation decay property for arbitrary graphs with bounded degree and large girth. We construct an algorithm exploiting this correlation decay. Then we focus on regular graphs, where explicit results can be obtained. Unlike the results for independent sets, our results for coloring do not have explicit bounds on the degree of the graph.

5.1 Coloring of trees and correlation decay

We use the definitions and notations of Subsection 4.1: $T, B(T), \mathcal{B}(T)$ denote respectively an arbitrary depth-t tree with maximum degree at most r, the boundary of the tree and the set of boundary conditions. The latter, however, is defined as the set of functions $b : B(T) \to \{1, 2, \ldots, q\}$ mapping nodes to colors. The root of this tree is v_0. Similarly to the case of independent set, we use notation $\mathbb{P}(\boldsymbol{C}(v) = j|b)$ to indicate probability that the random coloring $\boldsymbol{C}$ assigns color j to the node $v \in T$, subject to the boundary condition b, where probability is with respect to the Gibbs measure, (in this case uniform distribution) on the set of all proper colorings.

We need an analogue of Proposition 3, and in this case we use the following result by Jonasson [Jon02]. This result was used to establish uniqueness of Gibbs measures for coloring on infinite trees, but the main underlying result is a very strong form of correlation decay. (We note that Jonasson uses $r+1$ in place of r for the degree of a tree).

Theorem 7 (Jonasson [Jon02].) *Suppose $q \geq r+1$. There exists a computable value $\beta = \beta(r) < 1$ such that for every r-regular tree T with depth t*

$$\sup_{b \in \mathcal{B}(T)} \left| \mathbb{P}(\boldsymbol{C}(v_0) = j|b) - \frac{1}{q} \right| \leq \beta^t,$$

for every $j = 1, 2, \ldots, q$.

This result says that the color received by the root v_0 is independent from the colors of the boundary in a uniform way as a function of the depth. Note that the decay constant β does not even depend on q provided that $q \geq r+1$. The analysis of the proof in [Jon02] reveals that the same result holds for non-regular trees as well.

Corollary 1 *The result of Theorem 7 holds when T is an arbitrary depth-t tree with maximum degree r.*

5.2 Algorithm and the proof of Theorem 2

We propose the following algorithm for estimating the number of q-colorings of a given graph G.

Algorithm CountCOLOR

INPUT: A graph G with maximum degree r such that $q \geq r+1$, a node set $v_1, \ldots, v_n$, and a parameter $\epsilon > 0$.
BEGIN
1. Compute the girth $g(G)$. If $\beta^{\frac{g(G)}{2}-2} \geq \epsilon$ compute $\mathcal{C}(G)$ by exhaustive enumeration. Otherwise
2. Set $G' = G$, $Z = 1$, $t = g(G)/2$.
3. Find any node $v \in G'$ and its degree $r' = r(v, G) \leq r$. Set Z equal to

$$Z[q(1 - \frac{1}{q})^{r'}]$$

4. Set $G' = G' \setminus \{v\}$ and go to step 2.
END
OUTPUT: Z.

Proof : *Proof of Theorem 2.* The proof is very similar to the one of Theorem 1. Applying Proposition 2 we need to estimate in each step of the algorithm the expected value of used colors $\mathbb{E}_{G_k}\big[|\boldsymbol{C}(N(v_k, G_{k-1}))|\big]$. By fixing any boundary condition on depth-t neighborhood of v_k in the graph G_{k-1} the probability of any particular coloring of the nodes in $N(v_k, G_{k-1})$ is product of individual coloring probabilities. Each individual coloring probability is asymptotically $1/q$ provided t is large by Corollary 1. Therefore given a fixed color $i \leq q$, the probability that this color was never used in coloring nodes $N(v_k, G_{k-1})$ is asymptotically $(1-1/q)^{r'}$, where r' is the degree of v_k in the graph G_{k-1}. Therefore $q - \mathbb{E}_{G_k}\big[|\boldsymbol{C}(N(v_k, G_{k-1}))|\big]$ is asymptotically $q(1-1/q)^{r'}$, provided that $t = g(G)/2$ is sufficiently large.

The rest of the argument follows the lines the proof of Theorem 1. ■

5.3 Regular graphs and proof of Theorem 4

Our main tool is again rewiring performed on regular graphs with large girth. Given an arbitrary graph G and nodes $v_1, v_2 \in G$ such that v_1 and v_2 are not neighbors, and they do not have a common neighbor, let G^o be obtained from G by rewiring on v_1, v_2. Proposition 2 already relates the partition function of G to the one of $G \setminus \{v_1, v_2\}$. We now relate it to the one of G^o. Let $G' = G \setminus \{v_1, v_2\}$. That is G' is G^o before the pairs v_{1j}, v_{2j} are connected. Consider a random uniform q-coloring $\boldsymbol{C}$ selected in G'. The lemma below does not rely on assumptions of regularity or the girth size of the underlying graph G.

Lemma 5 *The following relation holds*

$$\frac{Z(q,G)}{Z(q,G^o)} = \frac{\mathbb{E}_{G'}\Big[\big(q - |\boldsymbol{C}(N(v_1,G))|\big)\big(q - |\boldsymbol{C}(N(v_2,G))|\big)\Big]}{\mathbb{P}_{G'}(\boldsymbol{C}(v_{1j}) \neq \boldsymbol{C}(v_{2j}),\ 1 \leq j \leq r)},$$

where $v_{ij}, j = 1, \ldots, r$ is the set of neighbors of $v_i, i = 1, 2$ in G.

Proof : Using the same argument as in Proposition 2 we obtain that

$$\frac{Z(q,G)}{Z(q,G')} = \mathbb{E}_{G'}\Big[\big(q - |\boldsymbol{C}(N(v_1,G))|\big)\big(q - |\boldsymbol{C}(N(v_2,G))|\big)\Big].$$

On the other hand $\frac{Z(q,G_0)}{Z(q,G')}$ is the probability that a randomly selected coloring in G' assigns different colors to each pair $v_{1j}, v_{2j}, j = 1, 2, \ldots, r$. Combining, we obtain the result. ∎

The following lemma is an analogue of Lemma 4.

Lemma 6 *Given $r \in \mathbb{N}$, $q \geq r+1$ $\epsilon > 0$, there exists a sufficiently large constant $g = g(r, \epsilon)$ such that for every r-regular graph G with girth $g(G) > g$, for every pair of nodes $v_1, v_2 \in G$ at distance at least $2g+1$*

$$\left| \mathbb{E}_{G'} \Big[\big(q - |\boldsymbol{C}(N(v_1, G))|\big) \big(q - |\boldsymbol{C}(N(v_2, G))|\big) \Big] - q^2 (1 - \frac{1}{q})^{2r} \right| < \epsilon. \tag{16}$$

$$\left| \mathbb{P}_{G'} (\boldsymbol{C}(v_{1j}) \neq \boldsymbol{C}(v_{2j}),\ 1 \leq j \leq r) - (\frac{q-1}{q})^r \right| < \epsilon. \tag{17}$$

Proof : The proof is very similar to the one of Lemma 4. In the graph G' consider depth-$t = g/2$ neighborhoods of nodes v_{ij}. By girth assumptions these neighborhoods are non-intersecting r-regular trees T_{ij}, with the exception that the each root v_{ij} has degree $r-1$. Fix any collection of colors $c_{ij} \in \{1, 2, \ldots, q\},\ i = 1, 2,\ j = 1, 2, \ldots, r$. Applying Corollary 1 and using the fact that the tree T_{ij} are non-intersecting, we obtain

$$\left| \mathbb{P}_{G'} (\boldsymbol{C}(v_{ij}) = c_{ij}, \forall i, j) - \frac{1}{q^{2r}} \right| \leq \epsilon, \tag{18}$$

provided $g = g(\epsilon, r, q)$ is sufficiently large. Thus, under $\mathbb{P}_{G'}$ the random colors $\{\boldsymbol{C}(v_{ij})\}$ are *approximately* independent and each uniformly distributed on the set of colors $\{1, 2, \ldots, q\}$. Thus (16) and (17) follows by choosing ϵ as ϵ/q^2 in (18). ∎

Proof : *Proof of Theorem 4.* The proof follows the same steps as the proof of Theorem 3. The results of Corollary 1 and Lemmas 2, 5, 6 are combined to obtain the limiting expression after the cancelation of $(\frac{q-1}{q})^r$. ∎

6 Random regular graphs

We prove now Theorems 5,6.

Proof : *Proof of Theorem 5.* We use the following fact about random regular graphs (see [JLR00]): given any constant $g > 0$ the total number of cycles with length $< g$ is w.h.p. at most some constant $c_1 = c_2(g)$. Thus given $G = G_r(n)$ there exists a graph $\hat{G}$ obtained from G by removing at most $(1 + 2 + \ldots + g)c_1(g) = c_2(g)$ edges, such that $\hat{G}$ has girth at least g. Observe that all but some constantly many nodes $c_3(g)$ of $\hat{G}$ have degree r. We now revisit the proof of Lemma 2 and apply the rewire operation to $\hat{G}$ with the following modification. First we observe that the result of the lemma still holds when we replace $2g+1$ by any large constant. Only the size of the remaining constant size graph may change. So we take some constant $c_4(g)$ instead of $2g+1$, which is to be specified later. In every step if the pair of nodes v_1, v_2 at a distance equal to the diameter of the current graph is such that v_1 and v_2 have depth-g neighborhoods which are regular trees, then we rewire on them. Otherwise we perform a breadth-first search for nodes v_1' and v_2' which do. Note that for this purpose it suffices to find nodes which are outside of depth-$g+1$ neighborhoods of $c_3(g)$ nodes which have degree $< r$. This will occur after our breadth-first choice inspects at most $c_3(g)(1 + r + \cdots + r^{g+1})$ nodes. The newly found nodes v_1', v_2' are at distance which is at least diameter minus $c_3(g)(1 + r + \cdots + r^{g+1})$. We rewire on v_1', v_2'. Since their depth-g neighborhood are regular trees, then using the same argument as for regular trees, we obtain that the ration of partition functions is approximately given $x^{-\frac{r}{2}}(2-x)^{-\frac{r-2}{2}}$, where

the level of approximation is controlled by g. We now select $c_4(g) = c_3(g)(1 + r + \cdots + r^{g+1})$ and use lemma2 with $c_4(g)$ replacing $2g + 1$. The rest of the argument is the same as for the case of regular graphs.

Theorem 6 is established in exactly the same manner. ∎

7 Conclusions

We have presented in this paper a new method for solving approximately some counting problems, which is not based on the Markov Chain sampling technique. We applied our method to independent sets and colorings in low degree graphs with large girth. The primary technical tool is a derivation of a certain correlation decay property which features prominently in statistical physics literature in connections with a completely different topic: uniqueness of Gibbs distributions on infinite trees. We certainly hope that our approach is more general and can be applied to other combinatorial problems. This constitutes an interesting direction for further research. Another research direction is removing the requirement of large girth, and here the difficulty is establishing correlation decay in non-tree like graphs. Such correlation decay was already established by Dobrushin [Dob70] back in 70's for lattice like graphs, but there is a recent extension by Weitz [Wei05] to a more general graphs. Perhaps this correlation decay (long-range independence) can be exploited to obtain non-Markov chain type algorithms for counting problems. Finally, it would be interesting to see if our approach can be converted to an algorithm for sampling from the uniform distribution, for example of independent set or coloring in the same class of low degree graphs with large girth. This would be a nice supplement to the classical approach of rapidly mixing Markov chains.

Acknowledgement. We gratefully acknowledge several fruitful conversations with Marc Mézard, Richardo Zecchina and Dimitris Achlioptas.

References

[AB05] D. Aldous and A. Bandyopadhyay, *A survey of max-type recursive distributional equations*, Annals of Applied Probability **15** (2005), no. 2, 1047–1110.

[Ald01] D. Aldous, *The $\zeta(2)$ limit in the random assignment problem*, Random Structures and Algorithms (2001), no. 18, 381–418.

[AM04] D. Achlioptas and C. Moore, *The chromatic number of random regular graphs*, 8th. Workshop on Randomization and Computation (RANDOM) (2004).

[AS03] D. Aldous and J. M. Steele, *The objective method: Probabilistic combinatorial optimization and local weak convergence*, Discrete Combinatorial Probability, H. Kesten Ed., Springer-Verlag, 2003.

[Ban] A. Bandyopadhyay, *Hard-core model on random graphs*, In preparation.

[Ban02] ———, *Bivariate uniqueness in the logistic fixed point equation*, Technical Report 629, Department of Statistics, UC, Berkeley (2002).

[BSVV] I. Bezakova, D. Stefankovic, V. Vazirani, and E. Vigoda, *Improved simulated annealing algorithm for the permanent and combinatorial counting problems*, Submitted.

[BW02] G. Brightwell and P. Winkler, *Random colorings of a Cayley tree*, in Contemporary Combinatorics, B. Bollobas, ed., Bolyai Society Mathematical Studies, 2002, pp. 247–276.

[BW04a] G.R. Brightwell and P. Winkler, *Graph homomorphisms and long range action*, in Graphs, morphisms and statistical physics (Nesetril and Winkler eds.), DIMACS series in discrete mathematics and computer science, 2004, pp. 29–47.

[BW04b] ———, *A second threshold for the hard-core model on a Bethe lattice*, Random Structures and Algorithms **24** (2004), no. 303-314.

[DaRK91] M. E. Dyer and A. Frieze an R. Kannan, *A random polynomial time algorithm for approximating the volume of convex bodies*, Journal of the Association for Computing Machinery **38** (1991), 1–17.

[DFHV04] M. Dyer, A. Frieze, T. Hayes, and E. Vigoda, *Randomly coloring constant degree graphs*, in Proceedings of 45th IEEE Symposium on Foundations of Computer Science, 2004.

[DGJ04] M. Dyer, L. A. Goldberg, and M. Jerrum, *Counting and sampling H-colourings*, Information and Computation **189** (2004), 1–16.

[Dob70] R. L. Dobrushin, *Prescribing a system of random variables by the help of conditional distributions*, Theory of Probability and its Applications **15** (1970), 469–497.

[Gam04] D. Gamarnik, *Linear phase transition in random linear constraint satisfaction problems*, Probability Theory and Related Fields. **129** (2004), no. 3, 410–440.

[Geo88] H. O. Georgii, *Gibbs measures and phase transitions*, de Gruyter Studies in Mathematics 9, Walter de Gruyter & Co., Berlin, 1988.

[GNSa] D. Gamarnik, T. Nowicki, and G. Swirscsz, *Maximum weight independent sets and matchings in sparse random graphs. Exact results using the local weak convergence method*, To appear in Random Structures and Algorithms.

[GNSb] D. Gamarnik, T. Nowicki, and G. Swirszcz, *Dynamics of exponential linear map in functional space*, Submitted.

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, John Wiley and Sons, Inc., 2000.

[Jon02] J. Jonasson, *Uniqueness of uniform random colorings of regular trees*, Statistics and Probability Letters **57** (2002), 243–248.

[JS89] M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM journal on computing **18** (1989), 1149–1178.

[JS97] ———, *The Markov chain Monte Carlo method: an approach to approximate counting and integration*, Approximation algorithms for NP-hard problems (D. Hochbaum, ed.), PWS Publishing Company, Boston, MA, 1997.

[JSV04] M. Jerrum, A. Sinclair, and E. Vigoda, *A polynomial-time approximation algorithms for permanent of a matrix with non-negative entries*, Journal of the Association for Computing Machinery **51** (2004), no. 4, 671–697.

[Kel85] F. Kelly, *Stochastic models of computer communication systems*, J. R. Statist. Soc. B **47** (1985), no. 3, 379–395.

[KLS97] R. Kannan, L. Lovasz, and M. Simonovits, *Random walks and $o^*(n^5)$ volume algorithm for convex bodies*, Random Structures and Algorithms **11** (1997), no. 1, 1–50.

[LV97] M. Luby and E. Vigoda, *Approximately counting up to four*, Proc. 29d Ann. ACM Symposium on the Theory of Computing (STOC) (1997).

[LV03] L. Lovasz and S. Vempala, *Simulated annealing in convex bodies and an $o^*(n^4)$ volume algorithm*, Proceedings of the 44th annual IEEE Symposium on Foundations of Computer Science, 2003, pp. 650–659.

[Mos04] E. Mossel, *Survey: information flow on trees*, J. Nestril and P. Winkler, editors. Graphs, Morphisms and Statistical Physiscs. DIMACS series in discrete mathematics and theoretical computer science. American Mathematical Society., 2004, pp. 155–170.

[MP05] M. Mezard and G. Parisi, *The cavity method at zero temperature*, http://fr.arxiv.org/ps/cond-mat/0207121 (2005).

[MPV87] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin-glass theory and beyond,* vol 9 of *Lecture Notes in Physics*, World Scientific, Singapore, 1987.

[RBMM04] O. Rivoire, G. Biroli, O. C. Martin, and M. Mezard, *Glass models on Bethe lattices*, Eur. Phys. J. B **37** (2004), 55–78.

[Tal01] M. Talagrand, *The high temperature case of the K-sat problem*, Probability Theory and Related Fields **119** (2001), 187–212.

[Tal03] ———, *Parisi formula*, Ann. of Mathematics, to apper (2003).

[Val79] L. G. Valiant, *The complexity of computing the permanent*, Theoretical computer science **8** (1979), 189–201.

[War05] J. Warren, *Dynamics and endogeny for recursive processes on trees*, http://arxiv.org/abs/math.PR/0506038 (2005).

[Wei05] D. Weitz, *Combinatorial criteria for uniqueness of gibbs measures*, Random Structures and Algorithms, to appear. (2005).