An FPTAS for Counting Proper Four-Colorings on Cubic Graphs

Pinyan Lu * Kuan Yang † Chihao Zhang ‡ Minshen Zhu §

Abstract

Graph coloring is arguably the most exhaustively studied problem in the area of approximate counting. It is conjectured that there is a fully polynomial-time (randomized) approximation scheme (FPTAS/FPRAS) for counting the number of proper colorings as long as $q \geq \Delta + 1$, where q is the number of colors and Δ is the maximum degree of the graph. The bound of $q = \Delta + 1$ is the uniqueness threshold for Gibbs measure on Δ-regular infinite trees. However, the conjecture remained open even for any fixed $\Delta \geq 3$ (The cases of $\Delta = 1, 2$ are trivial). In this paper, we design an FPTAS for counting the number of proper four-colorings on graphs with maximum degree three and thus confirm the conjecture in the case of $\Delta = 3$. This is the first time to achieve this optimal bound of $q = \Delta + 1$. Previously, the best FPRAS requires $q > \frac{11}{6}\Delta$ and the best deterministic FPTAS requires $q > 2.581\Delta + 1$ for general graphs. In the case of $\Delta = 3$, the best previous result is an FPRAS for counting proper 5-colorings. We note that there is a barrier to go beyond $q = \Delta + 2$ for single-site Glauber dynamics based FPRAS and we overcome this by correlation decay approach. Moreover, we develop a number of new techniques for the correlation decay approach which can find applications in other approximate counting problems.

1 Introduction

The problem of counting proper q-colorings has been extensively studied in computer science and statistical physics. It is known to be #**P**-hard for $q \geq 3$ even on graphs with bounded maximum degree $\Delta \geq 3$ [2]. A number of literature has been devoted to the design of approximation algorithms [1–5, 8–11, 17, 21]. The main algorithmic tool used in these works is the method of Markov chain Monte Carlo (MCMC), which is based on the simulation of a Markov chain on all proper q-colorings of a graph G whose stationary distribution is the uniform distribution. Although the Markov chains themselves are usually quite simple, it is challenging to prove the rapid mixing property of the chains and the interplay between the number q of colors and the maximum degree Δ of the graph G turns out to be a key measure for such property to hold.

The *Glauber dynamics* is a natural Markov chain to sample colorings and it converges to the uniform distribution as long as $q \geq \Delta + 2$. Jerrum [11] and Salas and Sokal [18] independently showed that the Glauber dynamics mixes rapidly if $q > 2\Delta$. The bound of 2Δ was considered as a barrier for the analysis of the Glauber dynamics and was even conjectured as a threshold for the rapid mixing property to hold for a period of time. Later, the conjecture was refuted by Bubley et al. [2] by showing that the Glauber dynamics indeed rapidly mixes when $\Delta = 3$ and $q = 5$. It is worth to note that this result attains the ergodicity threshold for Glauber dynamics ($q \geq \Delta + 2$) and thus it is the best one can achieve via this method. For general Δ, the state-of-the-art requires that $q > \frac{11}{6}\Delta$ [21].

All the above algorithms based on MCMC provide randomized algorithms. Can we get deterministic approximation algorithms? A deterministic FPTAS was obtained in [7] when $q \geq 2.8432\Delta + \beta$ for some sufficiently large β on triangle-free graphs. The bound was improved to $q \geq 2.581\Delta + 1$ on general graphs [16]. These new deterministic FPTASes are based on the *correlation decay* techniques.

Correlation decay approach is a relatively new approach to design approximate counting algorithm comparing to the MCMC method. One advantage of correlation decay approach is that the resulting algorithms are deterministic. Moreover, there are quite a few problems, for which an FPTAS based on correlation decay approach was provided while no MCMC based FPRAS is known. Among which, the most successful example is the problem of computing the partition function of anti-ferromagnetic two-spin systems [13, 14, 19], including counting independent sets [22]. The correlation decay

*Institute for Theoretical Computer Science, School of Information Management and Engineering, Shanghai University of Finance and Economics. lu.pinyan@mail.shufe.edu.cn

†Department of Computer Science, University of Oxford. kuan.yang.6@gmail.com

‡Institute of Theoretical Computer Science and Communications, The Chinese University of Hong Kong. chihao.zhang@gmail.com

§Purdue University. minshen.zh@gmail.com

This work was done in part while some of the authors were visiting the Simons Institute for the Theory of Computing.

based FPTAS is beyond the best known MCMC based FPRAS and achieves the boundary of approximability [6, 20], which is the uniqueness condition of the system. It is an important and challenging open question to extend this result to anti-ferromagnetic multi-spin systems. Coloring problem (or anti-ferromagnetic Potts model at zero temperature in the statistical physics terminology) is the most important and canonical example for anti-ferromagnetic multi-spin systems. It was proved that the uniqueness bound for this system on infinite regular trees is exactly $q = \Delta + 1$ [12]. This fact supports the conjecture that $q = \Delta + 1$ is the optimal bound for approximate counting in general graphs.

1.1 Our Results Our main result is to introduce new techniques to the correlation decay based algorithm and provide an FPTAS all the way up to the optimal bound of $q = \Delta + 1$ in the case of $\Delta = 3$.

THEOREM 1.1. *There exists an* **FPTAS** *to compute the number of proper four-colorings on graphs with maximum degree three.*

As the first algorithm achieving the optimal bound, we view it as a substantial step towards the optimal counting algorithms for general graphs. The contribution is three folds

- It overcomes an intrinsic barrier of MCMC (Glauber dynamics) based algorithms. For the case of $q = \Delta + 1$, the Glauber dynamics Markov chain is not ergodic and thus its stationary distribution is not unique. Nevertheless, we obtained FPTAS based on correlation decay technique.

- We provide a number of new design and analysis technique for correlation decay based algorithms, which can be used for general graph colorings or even other approximate counting problems.

- Our analysis is simpler than previous analysis of MCMC algorithms in similar settings. Even when the maximum degree $\Delta = 3$, it is already a very challenging problem to analyze the MCMC algorithms. In order to improve from $q = 6$ to $q = 5$, [2] did a very detailed case by case analysis and even require computer to verify the proof. We obtain the optimal bound of $q = 4$.

1.2 Our Techniques The key step in all the proofs of correlation decay analysis is to prove that a recursive function is contractive. For most of current known correlation decay based FPTASes for coloring problem, the following recursion, introduced in [7], is used

$$\mathbf{Pr}_{G,L}\left[c(v) = i\right] = \frac{\prod_{k=1}^{d}\left(1 - \mathbf{Pr}_{G_v,L_{k,i}}\left[c(v_k) = i\right]\right)}{\sum_{j\in L(v)}\prod_{k=1}^{d}\left(1 - \mathbf{Pr}_{G_v,L_{k,j}}\left[c(v_k) = j\right]\right)}.$$

The notation $\mathbf{Pr}_{G,L}\left[c(v) = i\right]$ denotes the marginal probability of the vertex v to be colored i in an instance (G, L) where G is a graph and L is a color list that associates each vertex a set of *feasible* colors. $\mathbf{Pr}_{G_v,L_{k,j}}\left[c(v_k) = j\right]$ denotes a similar marginal probability in a modified instance: G_v is the graph obtained from G by removing v and $L_{k,j}$ is obtained from L by removing color j from the color list of the vertex v_w where $w < k$ and v_w is the w-th neighbor of v in some canonical order. In this recursion, $\mathbf{Pr}_{G,L}\left[c(v) = i\right]$ can be computed from dq different variables of $\mathbf{Pr}_{G_v,L_{k,j}}\left[c(v_k) = j\right]$ with $k = 1, 2, \cdots, d$ and $j = 1, 2, \cdots, q$. In all previous analyses, one view them as dq free and independent variables and then bound the contraction in the worst case. For each single variable, one use the same recursion to expand to a set of dq new free and independent variables. This yields a computation tree of degree dq. However, the expansion of the underlying graph is of degree d and we usually call this gap the information loss or inefficiency of the recursion. However, these dq variables are not completely free and independent. The key new idea of this work is to make use of the relations among these variables to reduce redundancy and improve the efficiency of the recursion. Here are two key observations:

- For different colors i and j, the recursions for $\mathbf{Pr}_{G,L}\left[c(v) = i\right]$ and $\mathbf{Pr}_{G,L}\left[c(v) = j\right]$ involve exactly the same set of dq different variables.

- For $k = 1$, $L_{k,j}$ is identical for different color j.

Using these two observations, we can further expand the q different variables $\mathbf{Pr}_{G_v,L_{1,j}}\left[c(v_1) = j\right]$ with $j = 1, 2, \cdots, q$ into a set of dq different variables simultaneously. The expansion here is d (from q variables to dq different variables) rather than dq. In previous analyses, each single variable of these q different variables will further expand to dq free and independent variables. The total number becomes dq^2.

This can be viewed as a partial two-layer recursion: for a subset of variables in the one layer recursion, we use the same recursive function to further expand them. We note that the similar information loss or inefficiency for recursion appears in many correlation decay based approximation counting algorithms, and it is the main cause of the sub-optimality of the analysis. The approach introduced here can also be applied to improve their analyses and the key is to observe some relations among the redundant variables and make use of it. In the current case, the improvement becomes substantial when the number of variables is small.

Another crucial idea in our proof is to get better bounds for variables $\mathbf{Pr}_{G_v,L_{k,j}}[c(v_k)=j]$ and then we only need to prove contraction in these bounded range. This idea was used in previous analyses for counting colorings and many other problems. However, in our setting of $q=4$ and $\Delta=3$, these values could be as large as 1 and as small as 0. These are trivial bounds for a probability in general. Here, we use two observation to refine the bounds.

First, we notice that bound of 1 can only be achieved at the root of the recursion tree and for all other variable the value is between 0 and $\frac{1}{2}$. The boundaries of 0 and $\frac{1}{2}$ are both achievable and thus cannot be improved in general. To overcome this, we use the following *alternative argument*: When the two bounds of 0 and $\frac{1}{2}$ are achieved, we can easily detect it and thus compute the accurate values without error; otherwise, we can get better bounds. In the later case, we view the variables achieving 0 and $\frac{1}{2}$ as parameters rather than variables of the recursion function as we are sure that there is no errors for them, and just prove that the degenerated recursion function is contractive with respect to remaining variables. This is plausible since we have better bounds for remaining variables.

Last but not the least, as in most of the correlation decay approach, we use a potential function to amortize the decay rate. It remains the most important and magic ingredient of the proof. There is no general method to design potential function. Based on some numerical computation, we propose a new potential function in the paper. Comparing to the previous potential functions for coloring problem, the main new feature of our new function is its non-monotonicity, which captures the property of the problem. We remark that a potential function with a similar shape can be used for general graph coloring problem for similar set of recursions.

2 Preliminaries and the (New) Recursion

List coloring and Gibbs measure Although we start with a standard graph coloring instance, where each vertex can choose from the same set of 4 different colors, we need to modify the color list during our algorithms to get a list-coloring instance. Therefore we work on list-coloring problem in general. A list-coloring instance is specified by a graph-list pair (G,L), where $G=(V,E)$ is an undirected graph and $L: V \to 2^{[q]}$ associates each vertex v with a color list $L(v) \subseteq [q]$. A proper coloring of (G,L) is an assignment $c: V \to [q]$ such that (1) $c(v) \in L(v)$ for every $v \in V$ and (2) no two ends of an edge share the same color, i.e., $c(u) \neq c(v)$ for every $e=(u,v) \in E$.

The *Gibbs measure* is the uniform distribution over all proper colorings of (G,L). For every vertex $v \in V$ and color $i \in [q]$, we use $\mathbf{Pr}_{G,L}[c(v)=i]$ to denote the marginal probability that the vertex v is colored i in the Gibbs measure.

In the following, we use Δ to denote the maximum degree of the graph. If there exists an efficient algorithm to estimate the marginal probability $\mathbf{Pr}_{G,L}[c(v)=i]$, then one can construct an **FPTAS** to count the number of proper colorings.

LEMMA 2.1. *Suppose there exists an algorithm to compute a $(1 \pm \varepsilon)$ approximation of $\mathbf{Pr}_{G,L}[c(v)=i]$ for every list-coloring instance (G,L) with $G=(V,E)$, $q=4$, $\Delta=3$, $|L(v)| \geq d_v+1$ for every $v \in V$, and every $i \in [q]$ in time $\mathrm{poly}(|V|, \frac{1}{\varepsilon})$. Then there exists an* **FPTAS** *to compute the number of proper 4-colorings on graphs with maximum degree three.*

The proof Lemma 2.1 is routine, see e.g. [7]. Therefore, the remaining task is to approximate $\mathbf{Pr}_{G,L}[c(v)=i]$ for instances satisfying the conditions stated in Lemma 2.1.

Recursion Let (G,L) be an instance of list-coloring and $v \in V$ be a vertex. Let $N(v)=\{v_1,\ldots,v_d\}$ denote the set of neighbors of v in G, where d is the degree of v and let G_v be the graph obtained from G by removing vertex v and all its incident edges. For every $k \in [d]$ and $i \in [q]$, let

(2.1)
$$L_{k,i}(u) = \begin{cases} L(u) \setminus \{i\}, & \text{if } u = v_\ell \text{ for some } \ell < k, \\ L(u), & \text{otherwise} \end{cases}$$

be color lists. Then the following recursion for computing $\mathbf{Pr}_{G,L}[c(v)=i]$ first appeared in [7].

LEMMA 2.2. *Assuming above notations we have*

$$\text{(2.2)} \quad \mathbf{Pr}_{G,L}[c(v)=i] = \frac{\prod_{k=1}^{d}\left(1-\mathbf{Pr}_{G_v,L_{k,i}}[c(v_k)=i]\right)}{\sum_{j \in L(v)} \prod_{k=1}^{d}\left(1-\mathbf{Pr}_{G_v,L_{k,j}}[c(v_k)=j]\right)}$$

Then we can apply the same recursion to further expand $\mathbf{Pr}_{G_v,L_{k,j}}[c(v_k)=j]$ so on and so forth. It gives a computation tree to compute the value of the root $\mathbf{Pr}_{G,L}[c(v)=i]$. The condition that $q=4$, $\Delta=3$, and $|L(v)| \geq d_v+1$ for every $v \in V$, holds for all the list-coloring instances appearing in this computation tree. In the definition of new color lists (2.1), the list size is decreased by one only for the neighbors of v, but the degrees of its neighbors are also decreased by one in the new modified graph G_v since we have removed vertex v and all its incident edges. Therefore the condition $|L(v)| \geq d_v + 1$ remains satisfied for every $v \in V$ in the new instance. For every probability $\mathbf{Pr}_{G',L'}[c(u) = j]$ in the computation tree except the root, the degree $d_u \leq \Delta - 1 = 2$ since we come to this instance by removing a neighbor of u and thus the degree is decreased by at least one. All these observations are used in previous analyses. A more subtle and crucial new observation is that for every probability $\mathbf{Pr}_{G',L'}[c(u) = j]$ in the computation tree except the root, one have $|L(u)| \geq d_u + 2$ (which is stronger than $|L(u)| \geq d_u + 1$) since the degree of u is decreased by one while color list for u remains in the definition of (2.1).

We do not analyze this computation tree directly but turn to a more efficient one by taking the relation between variables into account. In the definition (2.1) of $L_{k,i}$, if $k = 1$ the new color lists remain the same for all the remaining vertices and thus is independent of the color i. Therefore, the $|L(v_1)|$ variables $\mathbf{Pr}_{G_v,L_{1,j}}[c(v_1) = j]$ are simply the marginal probabilities of vertex v_1 for different colors in the same instance. Therefore, when we further expand these variables, they involve same set of variables. We make use of this property and further expand these variables as follows. Let d_1 be the degree of v_1 in the graph G_v and $u_1, u_2, \ldots, u_{d_1}$ be the neighbors of v_k in the graph G_v. We use G_{v,v_1} to denote the graph obtained from G_v by removing the vertex v_1 and all its incident edges. For every $k \in [d_1]$ and $i \in [q]$, we use $L'_{k,i}$ to denote the color list such that

$$L'_{k,i}(u) = \begin{cases} L(u) \setminus \{i\}, & \text{if } u = u_\ell \text{ for some } \ell < k, \\ L(u), & \text{otherwise.} \end{cases}$$

Applying recursion (2.2), we obtain for every $j \in L(v_1)$, it holds that

(2.3)

$$\mathbf{Pr}_{G_v,L_{1,j}}[c(v_1) = j] = \frac{\prod_{k=1}^{d_1}\left(1 - \mathbf{Pr}_{G_{v,v_1},L'_{k,j}}[c(u_k) = j]\right)}{\sum_{l \in L(v_1)} \prod_{k=1}^{d_1}\left(1 - \mathbf{Pr}_{G_{v,v_1},L'_{k,l}}[c(u_k) = l]\right)}.$$

Then we substitute these into recursion (2.2) and get a new recursion for $\mathbf{Pr}_{G,L}[c(v) = i]$. We view this new recursion as one step in the computation tree and analyze its correlation decay property. From the algorithmic point of view, this does not make much difference but it do impact the analysis a lot. A similar situation appeared in [15], where one use the same algorithm to compute the number of independent sets in bipartite graphs as in general graphs. However, in that analysis, one combined two step of the recursion, and viewed it as one single step in the computation tree, and then analyze the contractive rate directly. Here, we analyze the partial two-step recursion, where one only further expand the variables for its first neighbor.

3 Algorithm

In this section, we describe our algorithm to estimate marginals.

The main idea of our algorithm to estimate $\mathbf{Pr}_{G,L}[c(v) = i]$ is to recursively apply recursions (2.2) and (2.3) up to some depth D. For the convenience of analysis, we distinguish between cases, depending on the degree of v and its neighbors.

- Our algorithm terminates in one of the following three boundary cases. (1) the color i is not in the color list $L(v)$, i.e., $i \notin L(v)$, in which case we return 0; (2) the recursion depth is zero, in which case we return $\frac{1}{|L(v)|}$ and (3) the degree of v in G is zero, i.e. v is an isolated vertex, in which case we return $\frac{1}{|L(v)|}$.
- If the degree of v in G is one, the algorithm branches into three cases according to the size of $L(v)$. In the case of $|L(v)| = 2$, we directly apply recursion (2.2). In the case of $|L(v)| = 4$, note that the sum of the marginal probabilities of colors $j \in L(v)$ on v_1 in G_v is 1, the denominator of the recursion (2.2) becomes a constant 3. For the same reason, in the case of $|L(v)| = 3$, we can denote the denominator of the recursion (2.2) by $2 + y$, where y is the marginal probability of color $j \in [4] \setminus L(v)$ (the absent color) on v_1 in G_v.
- If the degree of v in G is two or three, we faithfully apply recursion (2.2) and (2.3) to estimate the marginals. In order to simplify the analysis, we use the following convention in the case of $\mathsf{deg}_G(v) = 2$: Let the neighbors of v be v_1, v_2, then we always assume $\mathsf{deg}_G(v_1) \geq \mathsf{deg}_G(v_2)$ and if $\mathsf{deg}_G(v_1) = \mathsf{deg}_G(v_2) = 1$, then $i \notin L(v_1)$ implies $i \notin L(v_2)$.

The whole algorithm is described in Algorithm 1. We use procedure $P(G, L, v, i, D)$ to estimate $\mathbf{Pr}_{G,L}[c(v) = i]$ up to depth D.

The procedures $P1(G, L, v, i, D)$, $P2(G, L, v, i, D)$ and $P3(G, L, v, i, D)$ deal with the case of $\mathsf{deg}_G(v) = 1$, $\mathsf{deg}_G(v) = 2$ and $\mathsf{deg}_G(v) = 3$ respectively.

Case $\mathsf{deg}_G(v) = 1$: The algorithm for this case is described in Algorithm 2.

Case $\mathsf{deg}_G(v) = 2$: The algorithm for this case is described in Algorithm 3.

Algorithm 1: Estimate $\mathbf{Pr}_{G,L}[c(v)=i]$

```
Input  : Graph G; color lists L; vertex v; color
         i; recursion depth D;
Output: P ∈ [0,1]: Estimate of Pr_{G,L}[c(v) = i]
         up to depth D.
Function P(G, L, v, i, D)
    if i ∉ L(v) then
        return 0;
    end
    if D ≤ 0 then
        return 1/|L(v)|;
    end
    if deg_G(v) = 0 then
        return 1/|L(v)|;
    end
    if deg_G(v) = 1 then
        return P1(G, L, v, i, D);
    end
    if deg_G(v) = 2 then
        return P2(G, L, v, i, D);
    end
    if deg_G(v) = 3 then
        return P3(G, L, v, i, D);
    end
end
```

Algorithm 2: Estimate $\mathbf{Pr}_{G,L}[c(v)=i]$ when $\mathsf{deg}_G(v)=1$

```
Function P1(G, L, v, i, D)
    /* the vertex v has only one neighbor
       v_1.                                   */
    /* If the L(v) = {i, j}.                  */
    if |L(v)| = 2 then
        x ← P(G_v, L_{1,i}, v_1, i, D − 1);
        y ← P(G_v, L_{1,j}, v_1, j, D − 1);
        return (1−x)/(2−x−y);
    end
    if |L(v)| = 4 then
        x ← P(G_v, L_{1,i}, v_1, i, D − 1);
        return (1−x)/3;
    end
    /* in the following case, |L(v)| = 3.
       */
    if i ∈ L(v_1) then
        Let j be the color in the singleton set
         [4] \ L(v);
        x ← P(G_v, L_{1,i}, v_1, i, D − 1);
        y ← P(G_v, L_{1,j}, v_1, j, D − 1);
        return (1−x)/(2+y);
    end
end
```

Case $\mathsf{deg}_G(v)=3$: The algorithm for this case is described in Algorithm 4.

Proposition 3.1. *Let $q=4$. Given a list-coloring instance $(G=(V,E),L)$ with maximum degree 3, a vertex $v\in V$ satisfying $\mathsf{deg}_G(v)\le 2$ and $|L(v)|\ge \mathsf{deg}_G(v)+2$, a nonnegative integer D, we have $\sum_{i=1}^4 P(G,L,v,i,D)=1$.*

Proof. We will prove by induction on D. When $D=0$ we have $\sum_{i=1}^4 P(G,L,v,i,0)=\sum_{i\in L(v)}\frac{1}{|L(v)|}=1$. Suppose the proposition holds for $D-1$. To obtain the proof for D, we will discuss on degree of v.

(1) $\mathsf{deg}_G(v)=0$. Clearly $\sum_{i=1}^4 P(G,L,v,i,D)=\sum_{i\in L(v)}\frac{1}{|L(v)|}=1$.

(2) $\mathsf{deg}_G(v)=1$. Let $x_i=P(G_v,L_{1,i},v_1,i,D-1)$. By definition we have $L_{1,i}=L$ for all $i\in[4]$. Therefore $\sum_{i=1}^4 x_i=\sum_{i=1}^4 P(G_v,L,v_1,i,D-1)=1$ by induction hypothesis.

If $|L(v)|=4$, $\sum_{i=1}^4 P(G,L,v,i,D)=\sum_{i=1}^4\frac{1-x_i}{3}=\frac{4-1}{3}=1$.

If $|L(v)|=3$, assume $j\notin L(v)$. Then $\sum_{i=1}^4 P(G,L,v,i,D)=\sum_{i\in L(v)}\frac{1-x_i}{2+x_j}=\frac{3-(1-x_j)}{2+x_j}=1$.

(3) $\mathsf{deg}_G(v)=2$. In this case we have $|L(v)|=4$. So $\sum_{i=1}^4 P(G,L,v,i,D)=\sum_{i=1}^4\frac{(1-f_i)(1-y_i)}{\sum_{j\in L(v)}(1-f_j)(1-y_j)}=\frac{\sum_{i=1}^4(1-f_i)(1-y_i)}{\sum_{j=1}^4(1-f_j)(1-y_j)}=1$.

Using the same proof, we can also have $\sum_{j=1}^4 f_j=1$, where f_j is defined in Algorithm 3.

We conclude this section with the following lemma, whose proof is postponed to Section 7.

Lemma 3.1. *Let $q=4$. There exists an algorithm such that for every list-coloring instance (G,L) with $G=(V,E)$ and maximum degree at most three, every vertex $v\in V$, every coloring $i\in L(v)$ and every $0<\varepsilon<1$, it computes a number $\hat{p}$ in time $\mathrm{poly}\left(|V|,\frac{1}{\varepsilon}\right)$ satisfying $(1-\varepsilon)\hat{p}\le \mathbf{Pr}_{G,L}[c(v)=i]\le(1+\varepsilon)\hat{p}$.*

4 Bounds

In this section, we introduce upper and lower bounds for values computed in the algorithm. These bounds will play a crucial role in our proof.

Definition 5. *We call a a triple $(G=(V,E),L,v\in V)$ (a list-coloring instance together with a vertex in the*

Algorithm 3: Estimate $\mathbf{Pr}_{G,L}[c(v)=k]$ when $\mathsf{deg}_G(v)=2$

```
Function P2(G, L, v, i, D)
    /* the vertex v has two neighbors
       {v_1, v_2} with deg_G(v_1) ≥ deg_G(v_2);
       the vertex v_1 has neighbors
       {u_1, ..., u_{d_1}} in the graph G_v.  We
       also assume that if
       deg_G(v_1) = deg_G(v_2) = 1, then
       i ∉ L(v_1) implies i ∉ L(v_2).          */
    for j ∈ L(v) do
        if j ∉ L(v_1) then
            f_j ← 0;
        end
        else
            for k ∈ [d_1] do
                for w ∈ L(v_1) do
                    x_{k,w} ←
                      P(G_{v,v_1}, L'_{k,w}, u_k, w, D − 1);
                end
            end
            f_j ← ∏_{k=1}^{d_1}(1−x_{k,j}) / ∑_{w∈L(v_1)} ∏_{k=1}^{d_1}(1−x_{k,w});
        end
        y_j ← P(G_v, L_{2,j}, v_2, j, D − 1);
    end
    return (1−f_i)(1−y_i) / ∑_{j∈L(v)}(1−f_j)(1−y_j);
end
```

Algorithm 4: Estimate $\mathbf{Pr}_{G,L}[c(v)=i]$ when $\mathsf{deg}_G(v)=3$

```
Function P3(G, L, v, i, D)
    /* the vertex v has three neighbors
       {v_1, v_2, v_3}.                        */
    for j ∈ L(v) do
        x_j ← P(G_v, L_{1,j}, j, D);
        y_j ← P(G_v, L_{2,j}, j, D);
        z_j ← P(G_v, L_{3,j}, j, D);
    end
    return (1−x_i)(1−y_i)(1−z_i) / ∑_{j∈L(v)}(1−x_j)(1−y_j)(1−z_j);
end
```

graph) ***reachable*** *if the following condition is satisfied:* $\mathsf{deg}_G(u)\le 3$ *and* $|L(u)|\ge \mathsf{deg}_G(u)+1$ *for every* $u\in V$, $\mathsf{deg}_G(v)\le 2$ *and* $|L(v)|\ge \mathsf{deg}_G(v)+2$.

It follows from the discussion in section 2 that for all the probability $\mathbf{Pr}_{G,L}[c(v)=i]$ appeared in the computation tree except the root, (G,L,v) is reachable.

PROPOSITION 5.1. *Let (G,L,v) be reachable, $i\in[4]$ be a color , and D be a nonnegative integer. Then it holds that $0\le P(G,L,v,i,D)\le \frac{1}{2}$.*

Proof. We prove by induction on D. For base case, $P(G,L,v,i,D)$ will return $\frac{1}{|L(v)|}$ if $D=0$, so the proposition holds since $|L(v)|\ge 2$.

Suppose the proposition holds for $D-1$. We discuss on degree of v.

(1) $\mathsf{deg}_G(v)=0$. In this case $P(G,L,v,i,D)$ will return $\frac{1}{|L(v)|}$ where $|L(v)|\ge \mathsf{deg}_G(v)+2=2$, hence we have $0\le P(G,L,v,i,D)\le\frac{1}{2}$.

(2) $\mathsf{deg}_G(v)=1$. Let $x=P(G_v,L_{1,i},v_1,i,D-1)$ and $y=P(G_v,L_{1,j},j,D-1)$, as defined in Algorithm 2. Then $0\le x,y\le\frac{1}{2}$ by induction hypothesis.

According to algorithm, $P(G,L,v,i,D)$ will return $\frac{1-x}{3}$ or $\frac{1-x}{2+y}$, and in both cases this return value is bounded by $\frac{1}{2}$ given $x,y\ge 0$.

(3) $\mathsf{deg}_G(v)=2$. Let f_j, $x_{k,w}$ and y_j be the variables defined in Algorithm 3. By induction hypothesis we have $0\le x_{k,w},y_j\le\frac{1}{2}$. As for f_j, we need to further discuss on d_1.

If $d_1\in\{0,1\}$ then $f_j\le\frac{1}{2}$ immediately follows, as we have already seen in previous two cases. If $d_1=2$, we also have $f_j=\frac{\prod_{k=1}^{2}(1-x_{k,j})}{\sum_{w\in L(v_1)}\prod_{k=1}^{2}(1-x_{k,w})}\le(1-x_{1,j})\left((1-x_{1,j})+\frac{1}{2}\sum_{w\in L(v_1)\setminus\{j\}}(1-x_{1,w})\right)^{-1}=\frac{1-x_{1,j}}{(1-x_{1,j})+\frac{1}{2}(2+x_{1,j})}\le\frac{1}{2}$. Here we used the fact that $\sum_{w\in L(v_1)}x_{1,w}=1$, since $|L(v_1)|=4$ when $d_1=2$. Similarly we have $P(G,L,v,i,D)=\frac{(1-f_i)(1-y_i)}{\sum_{j\in L(v)}(1-f_j)(1-y_j)}\le\frac{1}{2}$.

PROPOSITION 5.2. *Let (G,L,v) be reachable, $i\in L(v)$ be a color , and D be a nonnegative integer. Then it holds that*

(1) if $\mathsf{deg}_G(v)=2$, then $P(G,L,v,i,D)\ge\frac{1}{13}$;

(2) if $\mathsf{deg}_G(v)\le 1$, then $P(G,L,v,i,D)\ge\frac{1}{6}$.

Proof. If $\mathsf{deg}_G(v)=0$ or $D=0$, we have $P(G,L,v,i,D)=\frac{1}{|L(v)|}\ge\frac{1}{4}$. In the following, we assume $D\ge 1$ and $\mathsf{deg}_G(v)\ge 1$. We discuss on degree of v.

(1) $\mathsf{deg}_G(v)=2$. It must be the case that $|L(v)|=4$. Therefore we have $P(G,L,v,i,D)=\frac{(1-f_i)(1-y_i)}{(1-f_i)(1-y_i)+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)(1-y_j)}\ge$

$\frac{(1-\frac{1}{2})^2}{(1-\frac{1}{2})^2+\sum_{j\in L(v)\setminus\{i\}}(1-0)} = \frac{\frac{1}{4}}{\frac{1}{4}+3} = \frac{1}{13}$. The upper bound $\frac{1}{2}$ for f_j and y_j is guaranteed by Proposition 5.1.

(2) $\mathsf{deg}_G(v) = 1$. If $|L(v)| = 4$, $P(G,L,v,i,D) = \frac{1-x}{3} \geq \frac{1-\frac{1}{2}}{3} = \frac{1}{6}$. If $|L(v)| = 3$, $P(G,L,v,i,D) = \frac{1-x}{2+y} \geq \frac{1-\frac{1}{2}}{2+\frac{1}{2}} = \frac{1}{5} > \frac{1}{6}$.

Note that overall we have lower bounds $\frac{1}{13}$ for $P(G,L,v,i,D)$, regardless of the degree of v. Furthermore, for f_j's defined in Algorithm 3 we can draw a similar conclusion: if $j \in L(v_1)$ then $f_j \geq \frac{1}{13}$.

PROPOSITION 5.3. *Let (G,L,v) be reachable and D be a nonnegative integer. Then for every color $i \in [4]$ such that $i \in L(u)$ for some neighbor u of v, we have*

(1) if $\mathsf{deg}_G(v) = 2$, then $P(G,L,v,i,D) \leq \frac{12}{25}$;

(2) if $\mathsf{deg}_G(v) = 1$, then $P(G,L,v,i,D) \leq \frac{6}{13}$.

Proof. If $i \notin L(v)$ then $P(G,L,v,i,D)$ returns 0 and we are done. If $D = 0$ and $i \in L(v)$, then $P(G,L,v,i,0) = \frac{1}{|L(v)|} \leq \frac{1}{3}$ since v have at least one neighbor. In the following, we assume $i \in L(v)$ and $D \geq 1$.

(1) $\mathsf{deg}_G(v) = 2$. If $i \in L(v_1)$, by Proposition 5.2 we know that $f_i \geq \frac{1}{13}$.

$P(G,L,v,i,D) = \frac{(1-f_i)(1-y_i)}{(1-f_i)(1-y_i)+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)(1-y_j)} \leq \frac{(1-f_i)(1-0)}{(1-f_i)(1-0)+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)\left(1-\frac{1}{2}\right)} = \frac{2(1-f_i)}{2(1-f_i)+(3-\sum_{j\in L(v)\setminus\{i\}}f_j)} = \frac{2(1-f_i)}{4-f_i} \leq \frac{24}{51} < \frac{12}{25}$.

Here we used the fact that $\sum_{j\in L(v)} f_j = 1$. On the other hand, if $i \in L(v_2)$ then $y_i \geq \frac{1}{13}$. So

$P(G,L,v,i,D) = \frac{(1-f_i)(1-y_i)}{(1-f_i)(1-y_i)+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)(1-y_j)} \leq \frac{(1-f_i)\left(1-\frac{1}{13}\right)}{(1-f_i)\left(1-\frac{1}{13}\right)+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)\left(1-\frac{1}{2}\right)} = \left(\frac{12}{13}(1-f_i)\right) \cdot \left(\frac{12}{13}(1-f_i) + \frac{1}{2}\sum_{j\in L(v)\setminus\{i\}}(1-f_j)\right)^{-1} = \frac{24(1-f_i)}{50-11f_i} \leq \frac{12}{25}$.

(2) $\mathsf{deg}_G(v) = 1$. Clearly $i \in L(u)$ where u is the only neighbor of v. So $x = P(G_v,L,u,i,D-1) \geq \frac{1}{13}$. If $|L(v)| = 4$ then $P(G,L,v,i,D) = \frac{1-x}{3} < \frac{1}{3} < \frac{6}{13}$. If $|L(v)| = 3$ then $P(G,L,v,i,D) = \frac{1-x}{2+y} \leq \frac{1-\frac{1}{13}}{2} = \frac{6}{13}$.

PROPOSITION 5.4. *Let (G,L,v) be reachable, $i \in [4]$ be a color. Assume $\mathsf{deg}_G(v) = 2$, then one of the following holds:*

(1) the vertex v and its two neighbors form a triangle in G;

(2) $\mathbf{P}(G,L,v,i,D) \leq \frac{13}{27}$ for all integer $D \geq 2$.

Proof. Without loss of generality we assume $i = 1$. Denote by v_1 and v_2 the two neighbors of v in G. We only need to consider the case when $1 \notin L(v_1)$ and $1 \notin L(v_2)$, i.e. $f_1 = y_1 = 0$, since otherwise by Proposition 5.3 we immediately have $P(G,L,v,1,D) \leq \frac{12}{25} < \frac{13}{27}$. In this case v_1 and v_2 each only have up to one neighbor in G_v, which we will denote by u_1 and u_2 respectively. We now continue to discuss in two cases.

(1) $\mathsf{deg}_{G_v}(v_2) = 0$. According to the algorithm $y_j = \frac{1}{|L(v)|}$ for $j \in L(v_2)$ and $y_j = 0$ for $j \notin L(v_2)$. If $|L(v_2)| = 3$ then $y_1 = 0$ and $y_2 = y_3 = y_4 = \frac{1}{3}$. We have $P(G,L,v,i,D) = \frac{1}{1+\sum_{j\in L(v)\setminus\{1\}}(1-f_j)(1-y_j)} = \frac{1}{1+\frac{2}{3}\left(3-\sum_{j\in L(v)\setminus\{1\}}f_j\right)} = \frac{1}{1+\frac{4}{3}} = \frac{3}{7} < \frac{13}{27}$. If $|L(v_2)| = 2$ we can assume $2 \notin L(v_2)$, thus $y_1 = y_2 = 0$ and $y_3 = y_4 = \frac{1}{2}$. We have $P(G,L,v,i,D) = \frac{1}{1+\sum_{j\in L(v)\setminus\{i\}}(1-f_j)(1-y_j)} \leq \left(1+(1-f_2)(1-0)+(1-f_3)\left(1-\frac{1}{2}\right)+(1-f_4)\left(1-\frac{1}{2}\right)\right)^{-1} = \left(1+\frac{1}{2}(1-f_2)+\frac{1}{2}\sum_{j\in L(v)\setminus\{1\}}(1-f_j)\right)^{-1} = \frac{1}{2+\frac{1}{2}(1-f_2)} \leq \frac{4}{9} < \frac{13}{27}$.

(2) $\mathsf{deg}_{G_v}(v_2) = 1$. Since v, v_1 and v_2 do not form a triangle, $L_{2,k}(u_2) = L(u_2)$ for all color k, thus it follows from Proposition 5.3 that for every $j \in L_{2,j}(u_2) = L(u_2)$, we have $y_j = P(G_v,L_{2,j},v_2,j,D-1) \leq \frac{6}{13}$. So if $L(v_2) \subseteq L(u_2)$ we have $P(G,L,v,1,D) = \frac{1}{1+\sum_{j\in L(v)\setminus\{1\}}(1-f_j)(1-y_j)} \leq \frac{1}{1+\left(1-\frac{6}{13}\right)\sum_{j\in L(v)\setminus\{1\}}(1-f_j)} = \frac{1}{1+\frac{7}{13}\cdot 2} = \frac{13}{27}$. On the other hand, consider $L(v_2) \not\subseteq L(u_2)$. Notice that $1 \notin L(v_2)$ so there should be some other color, say color 2, satisfying $2 \in L(v_2) \setminus L(u_2)$. This also forces $\mathsf{deg}_{G_{v,v_2}}(u_2) \leq 1$. Let $t_{kj} \triangleq P(G_{v,v_2},L_{2,k},u_2,j,D-2)$, where $G_{v,v_2} \triangleq (G_v)_{v_2}$, i.e., the graph obtained from G by removing v and v_2 and all edges incident to them. Since $2 \notin L(u_2)$ we have $t_{k2} = 0$ for all k. We need to further distinguish between two cases.

(1) $1 \in L(u_2)$. Recall that $L_{2,k}(u_2) = L(u_2)$ so $\forall k \in L(v)$, $1 \in L_{2,k}(u_2)$. Combining $\mathsf{deg}_{G_{v,v_2}}(u_2) \leq 1$, by Proposition 5.2 we have $\forall k \in L(v)$, $t_{k1} \geq \frac{1}{6}$. Specifically we have $t_{21} \geq \frac{1}{6}$. Now 1 is the color in singleton set $[4] \setminus L(v_2)$, so according to Algorithm 2, $y_2 = \frac{1-t_{22}}{2+t_{21}} \leq \frac{6}{13}$. As a consequence, we again have $y_j \leq \frac{6}{13}$ for all $j \in L(v_2)$ and the theorem follows.

(2) $1 \notin L(u_2)$. In this case u_2 is isolated in G_{v,v_2} with color list $\{3,4\}$. So it is clear that $t_{k1} = t_{k2} = 0$ and $t_{k3} = t_{k4} = \frac{1}{2}$ for every k.

Further we have $y_2 = \frac{1}{2}$ and $y_3 = y_4 = \frac{1}{4}$. Now $P(G, L, v, 1, D) = \frac{1}{1+\sum_{j \in L(v)\setminus\{1\}}(1-f_j)(1-y_j)} = \frac{1}{\left(1+\left(1-\frac{1}{2}\right)(1-f_2)+\left(1-\frac{1}{4}\right)(1-f_3)+\left(1-\frac{1}{4}\right)(1-f_4)\right)} = \frac{1}{1+\frac{3}{4}\sum_{j \in L(v)\setminus\{1\}}(1-f_j)-\frac{1}{4}(1-f_2)} = \frac{4}{9+f_2} \le \frac{4}{9} < \frac{13}{27}$.

Combining above propositions with the $\mathsf{deg}_G(v) \le 1$ case, we have the following theorem for bounds on marginal probabilities computed:

Theorem 5.1. *Let (G, L, v) be reachable, $i \in [4]$ be a color. Then one of the following propositions holds:*

(1) $P(G, L, v, i, D) = \frac{1}{2}$ for all integer $D \ge 2$;

(2) $P(G, L, v, i, D) \le \frac{13}{27}$ for all integer $D \ge 2$. Specifically $P(G, L, v, i, D) \le \frac{6}{13}$ when $\mathsf{deg}_G(v) \le 1$.

Furthermore, when $P(G, L, v, i, D) = \frac{1}{2}$ for some integer $D \ge 2$ the local structure of G around v falls into one of the following three cases (see Figure 1,2 and 3):

(1) $\mathsf{deg}_G(v) = 0$. Then $j, w \notin L(v)$ for two distinct colors j, w other than i (Figure 1).

(2) $\mathsf{deg}_G(v) = 1$. Denote by u neighbor of v. Then $i \notin L(u)$ and $j \notin L(u) \cup L(v)$ for some color $j \neq i$ (Figure 2).

(3) $\mathsf{deg}_G(v) = 2$. Denote by u_1 and u_2 two neighbors of v. Then v, u_1, u_2 form a triangle, and $i \notin L(u_1) \cup L(u_2)$ (Figure 3).

Proof. Assume w.l.o.g. $i = 1$, and we will assume $1 \in L(v)$, otherwise the statement is trivial. From Proposition 5.4 we know if $P(G, L, v, i, D) = \frac{1}{2}$ then v and its two neighbors v_1, v_2 must form a triangle, and $1 \notin L(v_1) \cup L(v_2)$, as depicted in Figure 3. If this is not the case then $P(G, L, v, i, D) \le \frac{13}{27}$. Now we focus on those $\mathsf{deg}_G(v) \le 1$ cases.

(1) $\mathsf{deg}_G(v) = 0$. We know $|L(v)| \ge 2$. If $|L(v)| \ge 3$ then apparently $P(G, L, v, 1, D) \le \frac{1}{3} < \frac{13}{27}$. If $|L(v)| = 2$ then $P(G, L, v, i, D) = \frac{1}{2}$ and this is just the case depicted in Figure 1

(2) $\mathsf{deg}_G(v) = 1$. By Algorithm 2, if $|L(v)| = 4$ then $P(G, L, v, 1, D)$ will return $\frac{1-x}{3} < \frac{13}{27}$. So we focus on $|L(v)| = 3$. Assume 2 is the color in singleton set $[4]\setminus L(v)$. According to the algorithm, $P(G, L, v, 1, D)$ now returns $\frac{1-x}{2+y}$ where

$$x = P(G_v, L_{1,1}, u, 1, D-1)$$

$$y = P(G_v, L_{1,2}, u, 2, D-1)$$

Notice that $\frac{1-x}{2+y}$ could reach $\frac{1}{2}$ if and only if $x = y = 0$. Otherwise at least one of x and y is bounded by $\frac{1}{6}$ from below, thus $\frac{1-x}{2+y}$ is bounded by $\max\left\{\frac{1-0}{2+1/6}, \frac{1-1/6}{2+0}\right\} = \frac{6}{13}$.

Moreover, $x = y = 0$ indicates that $1 \notin L_{1,1}(u)$ and $2 \notin L_{1,2}(u)$. Recall $L_{1,1} = L_{1,2} = L$, it immediately follows that $\mathsf{deg}_{G_v}(u) = 0$ and $1 \notin L(u)$ and $2 \notin L(u)$. Together with $2 \notin L(v)$ we have $2 \notin L(u) \cup L(v)$ which completes the proof (This case is depicted in Figure 2).

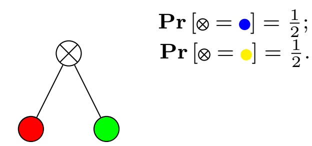

Figure 1: Boundary case one

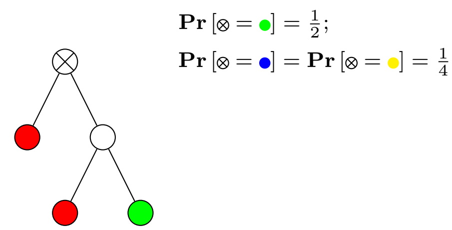

Figure 2: Boundary case two

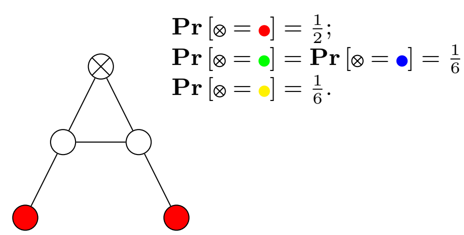

Figure 3: Boundary case three

Consider a depth D that is large enough(larger than the size of G), then clearly $P(G, L, v, i, D)$ should return $\mathbf{Pr}_{G,L}[c(v) = i]$. Therefore we can actually draw the same conclusions for true value $\mathbf{Pr}_{G,L}[c(v) = i]$. To make things clearer, we present the following theorem.

Theorem 5.2. *Let (G, L, v) be reachable, $i \in [4]$ be a color and $D \ge 2$ be an integer. Then*

(1) $P(G, L, v, i, D) = 0 \iff \mathbf{Pr}_{G,L}[c(v) = i] = 0$;

(2) $P(G, L, v, i, D) = \frac{1}{2} \iff \mathbf{Pr}_{G,L}[c(v) = i] = \frac{1}{2}$;

(3) $P(G, L, v, i, D) \in \left[\frac{1}{13}, \frac{13}{27}\right]$ $\iff \mathbf{Pr}_{G,L}[c(v) = i] \in \left[\frac{1}{13}, \frac{13}{27}\right]$.

6 Correlation Decay

In this section, we discuss the correlation decay property of our recursion. First we present the main theorem.

THEOREM 6.1. *Suppose $D \geq 3$ and $q = 4$. Let $\lambda = \frac{9996}{10000}$ be a constant, then for any list-coloring instance $(G = (V, E), L)$ satisfying $|L(v)| \geq \mathsf{deg}_G(v)+1$ for every $v \in V$, we have*

$$|P(G, L, v, i, D) - \mathbf{Pr}_{G,L}[c(v) = i]| \leq C \cdot \lambda^{D-3},$$

where $C > 0$ is some constant.

We can view the one step recursion $P(G, L, v, i, D)$ as a function F_i where each input of F_i is obtained by calling a depth-$(D-1)$ recursion on some list-coloring instance (G_k, L_k). Therefore F_i has 2 main variations, depending on whether $P1$ or $P2$ is called.

It is natural to conceive of a sufficient condition that probably looks like: the error of our estimation decays by a constant factor in every iteration. However, this is not generally true even for systems exhibiting correlation decay. This issue has already been addressed in [13,14], and in these works a potential-based analysis is adopted. We will once more utilize this method in our proof.

We choose $\varphi(x) = 2\ln x - 2\ln\left(\frac{1}{2} - x\right)$ whose derivative (potential function) is $\Phi(x) = \frac{1}{x(\frac{1}{2}-x)}$ and take $M \triangleq \frac{3}{2} - \sqrt{2} = \sup_{0\leq x\leq \frac{1}{2}} \frac{1}{(1-x)\Phi(x)}$.

Pick a list-coloring instance $(G = (V, E), L)$ with maximum degree 3, a vertex v in G with neighbor(s) v_1 and v_2 if exist satisfying $|L(v)| \geq \mathsf{deg}_G(v)+2$ and a color i. To prove Theorem 6.1, the idea is to apply induction on D, which can be formalized by the following lemma.

LEMMA 6.1. *Let $\lambda = \frac{9996}{10000}$ be a constant, then one of the following statements holds:*

(1) $F_i(\mathbf{x}) = F_i(\mathbf{x}^*) = 0$*;*

(2) $F_i(\mathbf{x}) = F_i(\mathbf{x}^*) = \frac{1}{2}$*;*

(3) $|\varphi(F_i(\mathbf{x})) - \varphi(F_i(\mathbf{x}^*))|$
$\leq \lambda \cdot \max_{j:x_j\in\left(0,\frac{1}{2}\right)} \left|\varphi(x_j) - \varphi(x_j^*)\right|$,

where $\mathbf{x}$ are the return values of subroutines called by $P(G, L, v, i, D)$ and $\mathbf{x}^*$ are true values of those called instances.

We shall point out here if the first two cases do not occur then $\varphi(F_i(\mathbf{x}))$ and $\varphi(F_i(\mathbf{x}^*))$ are always well-defined. This is a simple corollary of Lemma 5.2. Instead of proving this lemma, we will introduce Lemma 6.2 which can directly imply Lemma 6.1.

To ease the notation we first define the following. Let $\varphi(\mathbf{x}) = (\varphi(x_1), \varphi(x_2), \cdots, \varphi(x_d))$ for any d-dimensional vector $\mathbf{x}$, $d \in \mathbb{N}$, and similarly define $\varphi^{-1}(\mathbf{x})$.

LEMMA 6.2. *Suppose d is the arity of F_i. Define the contraction rate*

$$\alpha(\mathbf{x}) = \sum_{j=1}^{d} \frac{\Phi(F_i(\mathbf{x}))}{\Phi(x_j)} \left|\frac{\partial F_i(\mathbf{x})}{\partial x_j}\right|.$$

Then for all $\mathbf{x} \in Dom(F_i) \subseteq [0, \frac{1}{2}]^d$, we have $\alpha(\mathbf{x}) \leq \lambda$ where $\lambda = \frac{9996}{10000}$.

Before delving into the proof, we first show that how to prove Lemma 6.1 by Lemma 6.2.

Proof. [Proof of Lemma 6.1] Let $\mathcal{I}$ be the index set of variables of F_i. Let $\mathbf{x}_0 = \left\{x_i \mid i \in \mathcal{I}, x_i \in \left\{0, \frac{1}{2}\right\}\right\}$ and $\mathbf{x}_1 = \left\{x_i \mid i \in \mathcal{I}, x_i \in (0, \frac{1}{2})\right\}$. Let $\mathcal{I}_0$ and $\mathcal{I}_1$ be the corresponding index set of $\mathbf{x}_0$ and $\mathbf{x}_1$. Define $\mathbf{x}_0^*$ and $\mathbf{x}_1^*$ similarly.

Let $\mathbf{u}_1 = \varphi(\mathbf{x}_1)$, $\mathbf{u}_1^* = \varphi(\mathbf{x}_1^*)$, and since φ is strictly increasing we have $\mathbf{x}_1 = \varphi^{-1}(\mathbf{u}_1)$ and $\mathbf{x}_1^* = \varphi^{-1}(\mathbf{u}_1^*)$. Notice $\mathbf{u}_1^*$ is well-defined because we know $x_i \in (0, \frac{1}{2})$ if and only if $x_i^* \in (0, \frac{1}{2})$ by Lemma 5.2. In other words, $\mathbf{x}_0$ and $\mathbf{x}_1$ shares the same index set with $\mathbf{x}_0^*$ and $\mathbf{x}_1^*$, respectively.

Introduce

$$g(t) \triangleq \varphi(F_i(\mathbf{x}_0, \varphi^{-1}(t\mathbf{u}_1 + (1-t)\mathbf{u}_1^*))$$

so that $\varphi(F_i(\mathbf{x})) - \varphi(F_i(\mathbf{x}^*)) = \varphi(F_i(\mathbf{x}_0, \mathbf{x}_1)) - \varphi(F_i(\mathbf{x}_0^*, \mathbf{x}_1^*)) = g(1) - g(0)$. By *Mean Value Theorem* there exists $\tilde{t} \in (0, 1)$ such that

$$\frac{g(1) - g(0)}{1 - 0} = g'(\tilde{t}).$$

For convenience we denote $\tilde{\mathbf{u}}_1 = \tilde{t}\mathbf{u}_1 + (1-\tilde{t})\mathbf{u}_1^*$ and $\tilde{\mathbf{x}}_1 = \varphi^{-1}(\tilde{\mathbf{u}}_1)$. Clearly each component of $\tilde{\mathbf{x}}_1$ lies between 0 and $\frac{1}{2}$ since φ is a monotone function. Simple derivative calculation yields

$$\begin{aligned}
&|\varphi(F_i(\mathbf{x})) - \varphi(F_i(\mathbf{x}^*))| \\
=& \left|\sum_{j\in\mathcal{I}_1} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(\tilde{x}_j)} \frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j} \cdot (u_j - u_j^*)\right| \\
\leq& \sum_{j\in\mathcal{I}_1} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(\tilde{x}_j)} \left|\frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j}\right| \cdot |u_j - u_j^*| \\
\leq& \left(\sum_{j\in\mathcal{I}_1} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(\tilde{x}_j)} \left|\frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j}\right|\right) \cdot \max_{j\in\mathcal{I}_1} \left|u_j - u_j^*\right|.
\end{aligned}$$

Finally notice that if $x_j \in \left\{0, \frac{1}{2}\right\}$ then $\frac{1}{\Phi(x_j)} = 0$ and

$$\begin{aligned}
&\sum_{j\in\mathcal{I}_1} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(\tilde{x}_j)} \left|\frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j}\right| \\
&= \sum_{j\in\mathcal{I}_0} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(x_j)} \left|\frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j}\right| \\
&\quad + \sum_{j\in\mathcal{I}_1} \frac{\Phi(F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1))}{\Phi(\tilde{x}_j)} \left|\frac{\partial F_i(\mathbf{x}_0, \tilde{\mathbf{x}}_1)}{\partial x_j}\right| \\
&\le \sup_{\mathbf{x}\in[0,\frac{1}{2}]^d} \alpha(\mathbf{x}) \le \lambda.
\end{aligned}$$

This completes the proof.

We make some remarks. Here F_i is just a general concept representing the function of our algorithm. We use different recursions to compute the marginal probability as the degrees of v and its neighbors changes. As a consequence, the specific form, including arity of F_i has several variations, and depends on actual situations. Moreover, in our analysis we will frequently refine the domain of F_i because in some cases both true value and computed value never exceed a certain bound. Nevertheless, we can always obtain the expression of this contraction rate $\alpha(\mathbf{x})$, and it turns out that we can bound this rate for all variations of F_i.

The rest of this section is dedicated to prove Lemma 6.2. Our proof is based on the discussion on the degree of v. Thanks to the symmetry between colors, we will only need to prove for $i = 1$. The proofs for other colors are identical.

6.1 $\mathsf{deg}_G(v) = 1$ Denote by v_1 the only neighbor of v. In this case F_1 has three variations.

$$F_1 = \begin{cases} \frac{1-x}{3} & |L(v)| = 4 \\ \frac{1-x}{2+y} & 1 \in L(v), j \notin L(v) \\ 0 & 1 \notin L(v) \end{cases}$$

where $x = P(G_v, L_{1,1}, v_1, 1, D-1)$ and $y = P(G_v, L_{1,j}, v_1, j, D-1)$. We shall prove Lemma 6.1 for the first two variations since the last one is trivial.

1. $|L(v)| = 4$. The contraction rate writes as

$$\alpha(\mathbf{x}) = \frac{\Phi(F_1(x))}{\Phi(x)} \left|\frac{\partial F_1(x)}{\partial x}\right|.$$

Moreover we have the following upper bound $\frac{\Phi(F_1(x))}{\Phi(x)} \left|\frac{\partial F_1(x)}{\partial x}\right| = \frac{1}{3} \cdot \frac{x\left(\frac{1}{2}-x\right)}{\frac{1-x}{3}\left(\frac{1}{2}-\frac{1-x}{3}\right)} = \frac{3x(1-2x)}{(1-x)(1+2x)} \le \frac{3x}{1+2x} \le \frac{3}{4} < 1.$

2. $1 \in L(v), j \notin L(v)$. In this case F_1 is a binary function. The contraction rate writes as

$$\begin{aligned}
\alpha(x,y) = &\frac{\Phi(F_1(x,y))}{\Phi(x)} \left|\frac{\partial F_1(x,y)}{\partial x}\right| \\
&+ \frac{\Phi(F_1(x,y))}{\Phi(y)} \left|\frac{\partial F_1(x,y)}{\partial y}\right|.
\end{aligned}$$

We further discuss on three cases.

(a) $1 \notin L(v_1)$ and $j \notin L(v_1)$. In this case x and y are accurately computed, hence no error occurs in our computation.

(b) $1 \notin L(v_1)$ and $j \in L(v_1)$. Denote by d_1 degree of v_1 in graph G_v. Then $y = \frac{\prod_{k=1}^{d_1}(1-z_{jk})}{\sum_{l\in L(v_1)}\prod_{k=1}^{d_1}(1-z_{lk})} \ge \frac{\left(1-\frac{1}{2}\right)^2}{\left(1-\frac{1}{2}\right)^2+(1-0)\times 3} = \frac{1}{13}$. This lower bound also holds for y^*.

If $1 \notin L(v_1)$, then $x = x^* = 0$. Let $F_0 = F_1(0,\cdot)$ be the function obtained by fixing $x = 0$ in F_1. The contraction rate of F_0 is

$$\begin{aligned}
\alpha(y) &= \frac{\Phi(F_0(y))}{\Phi(y)} \left|\frac{\partial F_0(y)}{\partial y}\right| \\
&= \frac{y\left(\frac{1}{2}-y\right)}{\frac{1}{2+y}\left(\frac{1}{2}-\frac{1}{2+y}\right)} \cdot \frac{1}{(2+y)^2} \\
&= (1-2y) \le \frac{11}{13}.
\end{aligned}$$

(c) $1 \in L(v_1)$. Similarly we have $x, x^* \ge \frac{1}{13}$. Then

$$\begin{aligned}
\alpha(x,y) &= \frac{\Phi(F_1(x,y))}{\Phi(x)} \left|\frac{\partial F_1(x,y)}{\partial x}\right| \\
&\quad + \frac{\Phi(F_1(x,y))}{\Phi(y)} \left|\frac{\partial F_1(x,y)}{\partial y}\right| \\
&= \frac{1}{\frac{1-x}{2+y}\left(\frac{1}{2}-\frac{1-x}{2+y}\right)} \\
&\quad \cdot \left(\frac{x\left(\frac{1}{2}-x\right)}{2+y} + \frac{(1-x)y\left(\frac{1}{2}-y\right)}{(2+y)^2}\right) \\
&= \frac{x\left(\frac{1}{2}-x\right)(2+y) + y\left(\frac{1}{2}-y\right)(1-x)}{(1-x)\left(x+\frac{y}{2}\right)}.
\end{aligned}$$

We show that $\frac{x\left(\frac{1}{2}-x\right)(2+y)+y\left(\frac{1}{2}-y\right)(1-x)}{(1-x)\left(x+\frac{y}{2}\right)} \le \lambda$ for $\lambda = \frac{9996}{10000}$, which is equivalent to

$$\begin{aligned}
(6.4) \qquad & x\left(\frac{1}{2}-x\right)(2+y) + y\left(\frac{1}{2}-y\right)(1-x) \\
& \le \lambda \cdot (1-x)\left(x+\frac{y}{2}\right).
\end{aligned}$$

Inequality (6.4) can be simplified to

(6.5)
$$(1-x)y^2+\left(x^2-\frac{\lambda}{2}x-\frac{1-\lambda}{2}\right)y+x^2-(1-\lambda)x \geq 0.$$

Using the fact that $\frac{1}{13} \leq x \leq \frac{1}{2}$, we know the LHS of (6.5) is minimized at $y = \frac{2x^2-\lambda x-1+\lambda}{4x-4}$. Plugging this into (6.5) and it can be simplified to $1-2\lambda+\lambda^2+(16-14\lambda-2\lambda^2)x+(-52+36\lambda+\lambda^2)x^2+(32-20\lambda)x^3+4x^4 \leq 0$, which holds for $\frac{1}{13} \leq x \leq \frac{1}{2}$.

To summarize the analysis in Section 6.1, we have $\alpha(\mathbf{x}) \leq \lambda = \frac{9996}{10000}$.

6.2 $\mathsf{deg}_G(v) = 2$ Denote by v_1, v_2 two neighbors of v. Let $d_i = \mathsf{deg}_{G_v}(v_i)$ and we have $d_1 \geq d_2$. In this case

$$F_i = F_i(\mathbf{x},\mathbf{y}) = \frac{(1-f_i)(1-y_i)}{\sum_{j\in L(v)}(1-f_j)(1-y_j)}$$

where

$$f_i = \begin{cases} \frac{\prod_{k=1}^{d_1}(1-x_{k,i})}{\sum_{j\in L(v_1)}\prod_{k=1}^{d_1}(1-x_{k,j})} & i \in L(v_1) \\ 0 & i \notin L(v_1). \end{cases}$$

6.2.1 $d_1 = 2$ We first note that for $i,j \in L(v_1)$, f_i/f_j is bounded by constants.

PROPOSITION 6.1. *If $d_1 = 1$ or 2 and for every $1 \leq k \leq d_1$, $j \in L(v_1)$, we have $0 \leq x_{k,j} \leq \frac{1}{2}$, then for every $i,j \in L(v_1)$, it holds that $\frac{1}{4} \leq f_i/f_j \leq 4$ and $f_i \geq \frac{1}{13}$.*

Proof. For every $i,j \in L(v_1)$, we have $\frac{f_i}{f_j} = \frac{\prod_{k=1}^{d_1}(1-x_{k,i})}{\prod_{k=1}^{d_1}(1-x_{k,j})}$. Then the bound for the ratio follows from $d_1 = 1,2$ and $0 \leq x_{k,j} \leq \frac{1}{2}$ for every $1 \leq k \leq d_1$, $j \in L(v_1)$.

To see the lower bound for f_i, we note that $|L(v)| \leq 4$ and thus $1 = \sum_{j\in L(v)} f_j \leq f_i + 4\sum_{j\in L(v)\setminus\{i\}} f_i \leq 13 f_i$.

To prove Lemma 6.2 it suffices to bound the contraction rate

$$\alpha(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{2}\sum_{j\in L(v_1)}\frac{\Phi(F_1)}{\Phi(x_{ji})}\left|\frac{\partial F_1(\mathbf{x})}{\partial x_{ji}}\right| + \sum_{j=1}^{4}\frac{\Phi(F_1)}{\Phi(y_j)}\left|\frac{\partial F_1(\mathbf{y})}{\partial y_j}\right|.$$

Simple calculation yields

$$\begin{aligned}
&\sum_{i=1}^{2}\sum_{j\in L(v_1)}\frac{\Phi(F_1)}{\Phi(x_{ji})}\left|\frac{\partial F_1}{\partial x_{ji}}\right| \\
&= \sum_{i=1}^{2}\left(\frac{\Phi(F_1)}{\Phi(x_{1i})}\cdot\frac{F_1 f_1}{1-x_{1i}}\cdot\sum_{k=2}^{4}\frac{F_k}{1-f_k} + \right. \\
&\left.\sum_{j\in L(v_1)\setminus\{1\}}\frac{\Phi(F_1)}{\Phi(x_{ji})}\cdot\frac{F_1 f_j}{1-x_{ji}}\left|\frac{1}{1-f_1}-\sum_{\substack{k=1\\k\neq j}}^{4}\frac{F_k}{1-f_k}\right|\right) \\
&\leq \sum_{i=1}^{2}\left(M\cdot\Phi(F_1)F_1\left(f_1\sum_{k=2}^{4}\frac{F_k}{1-f_k} + \right.\right. \\
&\left.\left.\sum_{j\in L(v_1)\setminus\{1\}} f_j\left|\frac{1}{1-f_1}-\sum_{\substack{k=1\\k\neq j}}^{4}\frac{F_k}{1-f_k}\right|\right)\right) \triangleq 2\cdot P_1(\mathbf{f},\mathbf{y}),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{4}\frac{\Phi(F_1)}{\Phi(y_j)}\left|\frac{\partial F_1}{\partial y_j}\right| \\
&= \frac{\Phi(F_1)}{\Phi(y_1)}\cdot\frac{F_1(1-F_1)}{1-y_1} + \sum_{j=2}^{4}\frac{\Phi(F_1)}{\Phi(y_j)}\cdot\frac{F_1F_j}{1-y_j} \triangleq P_2(\mathbf{f},\mathbf{y}).
\end{aligned}$$

Now we only need to bound

$$\alpha(\mathbf{x},\mathbf{y}) = 2P_1(\mathbf{f},\mathbf{y}) + P_2(\mathbf{f},\mathbf{y}).$$

Notice that after substituting M for $\frac{1}{(1-x)\Phi(x)}$ we can ignore $\mathbf{x}$ and treat P_1 and P_2 as functions of $\mathbf{f}$ and $\mathbf{y}$, with some constraints on $\mathbf{f}$ as we will see soon.

Discussion on the absolute value. Let $D_j \triangleq \frac{1}{1-f_1} - \sum_{\substack{k=1\\k\neq j}}^{4}\frac{F_k}{1-f_k}$ for $j = 2,3,4$. We show that at least two of these D_j's are nonnegative. Assume for the contraction that $D_2, D_3 < 0$, then we obtain

$$\begin{aligned}
\frac{1}{1-f_1}-\frac{F_1}{1-f_1}-\frac{F_3}{1-f_3}-\frac{F_4}{1-f_4} &< 0 \\
\frac{1}{1-f_1}-\frac{F_1}{1-f_1}-\frac{F_2}{1-f_2}-\frac{F_4}{1-f_4} &< 0.
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
(6.6)\qquad &(1-f_2)(1-y_2)+(f_1-f_3)(1-y_3) \\
&\quad +(f_1-f_4)(1-y_4) < 0 \\
(6.7)\qquad &(1-f_3)(1-y_3)+(f_1-f_2)(1-y_2) \\
&\quad +(f_1-f_4)(1-y_4) < 0
\end{aligned}$$

(6.6)+(6.7) gives

$$\begin{aligned}
&(1+f_1-2f_3)(1-y_3)+(1+f_1-2f_2)(1-y_2) \\
&\quad +2(f_1-f_4)(1-y_4) < 0
\end{aligned}$$

Since $1+f_1-2f_3, 1+f_1-2f_2 > 0$ and $0 < y_2, y_3, y_4 < \frac{1}{2}$, it holds that $3f_1 + 1 - f_2 - f_3 - 2f_4 < 0$.

Since $d_1 = 2$ we have $|L(v)| = 4$ so Proposition 6.1 holds for all pairs of f_i, f_j, $1 \le i < j \le 4$. Applying $f_1 + f_2 + f_3 + f_4 = 1$, we obtain $4f_1 < f_4$, which is a contradiction.

Therefore, we have either all D_j for $j = 2, 3, 4$ are nonnegative or at most one of it is negative. Assume D_2 is negative, i.e.,

$$(1-f_2)(1-y_2)+(f_1-f_3)(1-y_3)+(f_1-f_4)(1-y_4) < 0.$$

Since $(1 - f_2)(1 - y_2) \ge 0$, we have either $f_1 < f_3$ or $f_1 < f_4$ or both. W.l.o.g. assume $f_1 < f_3$, now we distinguish between two cases:

- ($f_1 < f_4$) In this case, we can let $y_2 = \frac{1}{2}$ and $y_3 = y_4 = 0$, this gives

$$1 - f_2 + 2(f_1 - f_3) + 2(f_1 - f_4) < 0.$$

  Using the identity $f_1 + f_2 + f_3 + f_4 = 1$, we obtain

$$6f_1 + f_2 - 1 < 0.$$

- ($f_1 \ge f_4$) In this case, we can let $y_2 = \frac{1}{2}$, $y_3 = 0$ and $f_4 = f_1$, this gives

$$1 - f_2 + 2(f_1 - f_3) < 0.$$

  Using $f_3 = 1-f_1-f_2-f_4 \le \frac{7}{8}-f_1-f_2$, we obtain

$$4f_1 + f_2 - \frac{3}{4} < 0.$$

Now we can continue our analysis of $\alpha(\mathbf{x}, \mathbf{y})$.

Case 1: All D_j are nonnegative for $j = 2, 3, 4$. Introduce the following function of w and f

$$G_\xi(w, f) = \frac{1-f}{\Phi(1-\frac{w}{1-f})} + 4M\xi \cdot \frac{w}{1-f}$$

where $\xi \in [0, 1]$ is some constant parameter. The following two lemmas are used to symmetrize $\alpha(\mathbf{x}, \mathbf{y})$.

Lemma 6.3. *$G_0(w, f)$ is concave when $f \in [0, \frac{1}{2}]$ and $\frac{1-f}{2} \le w \le 1 - f$, hence for all w_i, f_i satisfying $f_i \in [0, \frac{1}{2}]$ and $\frac{1-f_i}{2} \le w_i \le 1 - f_i, i = 1, 2, \cdots, n$, we have*

$$\frac{G_0(w_1, f_1) + G_0(w_2, f_2) + \cdots + G_0(w_n, f_n)}{n} \le G_0\left(\frac{w_1 + w_2 + \cdots + w_n}{n}, \frac{f_1 + f_2 + \cdots + f_n}{n}\right).$$

Proof. The Hessian of $G_0(w, f)$ is $\begin{bmatrix} -\frac{2}{1-f} & -\frac{2w}{(1-f)^2} \\ -\frac{2w}{(1-f)^2} & -\frac{2w^2}{(1-f)^3} \end{bmatrix}$, which is negative semi-definite when $f \in [0, \frac{1}{2}]$.

Lemma 6.4. *For all $w_1, w_2, w_3 \in [0, \frac{1}{2}]$ and $f_1, f_2, f_3 \in [\frac{1}{13}, \frac{1}{2}]$ such that $\frac{1-f_i}{2} \le w_i \le 1 - f_i, i = 1, 2, 3$, we have*

$$\frac{1}{2}\left(G_\xi(w_1, f_1) + G_\xi(w_2, f_2)\right) \le \kappa \cdot G_\xi\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right);$$

$$\frac{1}{3}\left(G_\xi(w_1, f_1) + G_\xi(w_2, f_2) + G_\xi(w_3, f_3)\right) \le \kappa \cdot G_\xi\left(\frac{w_1 + w_2 + w_3}{3}, \frac{f_1 + f_2 + f_3}{3}\right),$$

holds for any $\xi \in [0, 1]$, where $\kappa = \frac{1038}{1000}$.

Proof. First we shall point out that if the lemma holds for $\xi = 1$, then it should hold for any other $0 \le \xi < 1$.

Suppose the lemma holds for $\xi = 1$. That is

$$\frac{1}{2}\left(G_1(w_1, f_1) + G_1(w_2, f_2)\right) \le \kappa \cdot G_1\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right).$$

Rewrite $G_\xi(w, f) = (1-\xi)G_0(w, f)+\xi G_1(w, f)$. Recall that G_0 is concave, thus

$$\begin{aligned}
&\frac{1}{2}\left(G_\xi(w_1, f_1) + G_\xi(w_2, f_2)\right) \\
&\le (1-\xi)G_0\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right) + \xi\kappa \cdot G_1\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right) \\
&\le (1-\xi)\kappa \cdot G_0\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right) + \xi\kappa \cdot G_1\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right) \\
&= \kappa \cdot G_\xi\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right).
\end{aligned}$$

The same argument works for the 3-variable case. So it remains to prove the $\xi = 1$ case.

It can be rigorously proved by *Mathematica* (the codes are in Section 8) that for all $w_1, w_2, w_3 \in [0, \frac{1}{2}]$ and $f_1, f_2, f_3 \in [\frac{1}{13}, \frac{1}{2}]$ such that $\frac{1-f_i}{2} \le w_i \le 1 - f_i, i =$

$1, 2, 3$, we have

$$\frac{1}{2}(G_1(w_1, f_1) + G_1(w_2, f_2)) \leq \kappa_1 \cdot G_1\left(\frac{w_1 + w_2}{2}, \frac{f_1 + f_2}{2}\right);$$

$$\frac{1}{3}(G_1(w_1, f_1) + 2G_1(w_2, f_2)) \leq \kappa_2 \cdot G_1\left(\frac{w_1 + 2w_2}{3}, \frac{f_1 + 2f_2}{3}\right).$$

Here $\kappa_1 = \frac{10195}{10000}, \kappa_2 = \frac{10181}{10000}$ and $\kappa_1\kappa_2 \leq \kappa$. As a consequence,

$$\begin{aligned}
&\frac{1}{3}(G_1(w_1, f_1) + G_1(w_2, f_2) + G_1(w_3, f_3)) \\
&= \frac{1}{2}\left(\frac{1}{3}(2G_1(w_1, f_1) + G_1(w_2, f_2)) + \frac{1}{3}(G_1(w_2, f_2) + 2G_1(w_3, f_3))\right) \\
&\leq \kappa_2 \cdot \frac{1}{2}\left(G_1\left(\frac{2w_1 + w_2}{3}, \frac{2f_1 + f_2}{3}\right) + G_1\left(\frac{w_2 + 2w_3}{3} + \frac{f_2 + 2f_3}{3}\right)\right) \\
&\leq \kappa_1\kappa_2 \cdot G_1\left(\frac{1}{2}\left(\frac{2w_1 + w_2}{3} + \frac{w_2 + 2w_3}{3}\right), \frac{1}{2}\left(\frac{2f_1 + f_2}{3} + \frac{f_2 + 2f_3}{3}\right)\right) \\
&\leq \kappa \cdot G_1\left(\frac{w_1 + w_2 + w_3}{3}, \frac{f_1 + f_2 + f_3}{3}\right).
\end{aligned}$$

Recall that $f_j = 0$ for $j \notin L(v_1)$, and

$$\begin{aligned}
&\sum_{j \in L(v_1)\setminus\{1\}} f_j \left(\frac{1}{1 - f_1} - \sum_{\substack{k=1\\k\neq j}}^{4} \frac{F_k}{1 - f_k}\right) \\
&= 1 - \sum_{j=2}^{4} f_j \sum_{\substack{k=1\\k\neq j}}^{4} \frac{F_k}{1 - f_k} \\
&= 1 - \sum_{j=1}^{4} f_j \sum_{\substack{k=1\\k\neq j}}^{4} \frac{F_k}{1 - f_k} + f_1 \sum_{k=2}^{4} \frac{F_k}{1 - f_k} \\
&= 1 - \sum_{k=1}^{4} \frac{F_k}{1 - f_k} \sum_{\substack{j=1\\j\neq k}}^{4} f_j + f_1 \sum_{k=2}^{4} \frac{F_k}{1 - f_k} \\
&= f_1 \sum_{j=2}^{4} \frac{F_j}{1 - f_j}.
\end{aligned}$$

So we have

$$\alpha = \Phi(F_1)F_1\left(\frac{1 - F_1}{(1 - y_1)\Phi(y_1)} + \sum_{j=2}^{4} \frac{F_j}{(1 - y_j)\Phi(y_j)} + 4Mf_1 \sum_{j=2}^{4} \frac{F_j}{1 - f_j}\right).$$

Define symmetric forms of F_k as follows.

$$\hat{F}_k(f_1, f_2, y_1, y_2) = \frac{(1 - f_k)(1 - y_k)}{(1 - f_1)(1 - y_1) + 3(1 - f_2)(1 - y_2)}, \quad k = 1, 2.$$

Then we can define the symmetric form of α

$$\hat{\alpha}(f_1, f_2, y_1, y_2) = \Phi(\hat{F}_1)\hat{F}_1 \cdot \left(\frac{1 - \hat{F}_1}{(1 - y_1)\Phi(y_1)} + \frac{3\hat{F}_2}{(1 - y_2)\Phi(y_2)} + 12Mf_1 \cdot \frac{\hat{F}_2}{1 - f_2}\right).$$

Lemma 6.5. *For all $\mathbf{f}, \mathbf{y} \in [0, \frac{1}{2}]^4$ such that $\frac{1}{13} \leq f_1, f_2, f_3, f_4 \leq \frac{1}{2}$ and $f_1 + f_2 + f_3 + f_4 = 1$, there exists $\hat{f}_2, \hat{y}_2 \in [0, \frac{1}{2}]$ such that $f_1 + 3\hat{f}_2 = 1$ and*

$$\alpha(\mathbf{f}, \mathbf{y}) \leq \kappa \cdot \hat{\alpha}(f_1, \hat{f}_2, y_1, \hat{y}_2)$$

where $\kappa = \frac{1038}{1000}$.

Proof. Let $w_k = (1 - f_k)(1 - y_k), k = 2, 3, 4$, $A(\mathbf{f}, \mathbf{y}) = \sum_{j=1}^{4}(1 - f_j)(1 - y_j)$ be the denominator of F_k, and $\hat{A}(f_1, f_2, y_1, y_2) = (1 - f_1)(1 - y_1) + 3(1 - f_2)(1 - y_2)$ be the denominator of $\hat{F}_k$. Then

$$\begin{aligned}
\alpha &= \Phi(F_1)F_1\left(\frac{1 - F_1}{(1 - y_1)\Phi(y_1)} + \frac{1}{A}\sum_{j=2}^{4} \frac{1 - f_j}{\Phi(1 - \frac{w_j}{1 - f_j})} + 4Mf_1 \cdot \frac{w_j}{1 - f_j}\right) \\
&= \Phi(F_1)F_1\left(\frac{1 - F_1}{(1 - y_1)\Phi(y_1)} + \frac{1}{A}\sum_{j=2}^{4} G_{f_1}(w_j, f_j)\right).
\end{aligned}$$

Take $\hat{w}_2$ and $\hat{f}_2$ such that $3\hat{w}_2 = w_1 + w_2 + w_3$, $3\hat{f}_2 = f_1 + f_2 + f_3$, and take $\hat{y}_2 = 1 - \frac{\hat{w}_2}{1 - \hat{f}_2}$. Therefore $f_1 + 3\hat{f}_2 = f_1 + f_2 + f_3 + f_4 = 1$ and

$$A(\mathbf{f}, \mathbf{y}) = \hat{A}(f_1, \hat{f}_2, y_1, \hat{y}_2)$$

$$F_1(\mathbf{f}, \mathbf{y}) = \hat{F}_1(f_1, \hat{f}_2, y_1, \hat{y}_2).$$

Furthermore, w_j and y_j satisfy the condition of Lemma 6.4 hence

$$\begin{aligned}\alpha &\le \Phi(F_1)F_1\left(\frac{1-F_1}{(1-y_1)\Phi(y_1)}+3\kappa\cdot\frac{G_{f_1}(\hat{w}_2,\hat{f}_2)}{A}\right)\\ &=\Phi(\hat{F}_1)\hat{F}_1\left(\frac{1-\hat{F}_1}{(1-y_1)\Phi(y_1)}\right.\\ &\quad\left.+3\kappa\cdot\left(\frac{\hat{F}_2}{(1-\hat{y}_2)\Phi(\hat{y}_2)}+4Mf_1\cdot\frac{\hat{F}_2}{1-\hat{f}_2}\right)\right)\\ &\le\kappa\cdot\Phi(\hat{F}_1)\hat{F}_1\left(\frac{1-\hat{F}_1}{(1-y_1)\Phi(y_1)}\right.\\ &\quad\left.+\frac{3\hat{F}_2}{(1-\hat{y}_2)\Phi(\hat{y}_2)}+12Mf_1\cdot\frac{\hat{F}_2}{1-\hat{f}_2}\right)\\ &=\kappa\cdot\hat{\alpha}(f_1,\hat{f}_2,y_1,\hat{y}_2).\end{aligned}$$

Lemma 6.6. *For all $f_1, f_2, y_1, y_2 \in [0, \frac{1}{2}]$ such that $\frac{1}{13} \le f_1 \le \frac{1}{2}$ and $f_1 + 3f_2 = 1$, we have*

$$\hat{\alpha}(f_1,f_2,y_1,y_2)\le\frac{963}{1000}.$$

Proof. The lemma can be rigorously proved by *Mathematica*. The codes are in Section 8.

Theorem 5.2 and Proposition 6.1 provide the condition for Theorem 6.5, and combining Theorem 6.6 gives

$$\alpha(\mathbf{f},\mathbf{y})\le\kappa\cdot\hat{\alpha}(f_1,\hat{f}_2,y_1,\hat{y}_2)\le\frac{1038}{1000}\cdot\frac{963}{1000}<\frac{9996}{10000}.$$

Case 2: D_j is negative for some j. Without loss of generality, we can assume $j = 2$, i.e., $\frac{1}{1-f_1} - \sum_{\substack{k=1\\k\neq 2}}^{4}\frac{F_k}{1-f_k} < 0$. Therefore,

$$\begin{aligned}&\sum_{j=2}^{4}f_j\left|\frac{1}{1-f_1}-\sum_{\substack{k=1\\k\neq j}}^{4}\frac{F_k}{1-f_k}\right|\\ &=f_1\sum_{j=2}^{4}\frac{F_j}{1-f_j}-2f_2\left(\frac{1}{1-f_1}-\sum_{\substack{k=1\\k\neq 2}}^{4}\frac{F_k}{1-f_k}\right).\end{aligned}$$

So we have

$$\begin{aligned}\alpha&=\Phi(F_1)F_1\left(\frac{1-F_1}{(1-y_1)\Phi(y_1)}+\sum_{j=2}^{4}\frac{F_j}{(1-y_j)\Phi(y_j)}\right.\\ &\left.+4Mf_1\sum_{j=2}^{4}\frac{F_j}{1-f_j}-4Mf_2\left(\frac{1}{1-f_1}-\sum_{\substack{k=1\\k\neq 2}}^{4}\frac{F_k}{1-f_k}\right)\right)\end{aligned}$$

which is a function of $\mathbf{f}, \mathbf{y} \in [0, 1]^4$ where $f_1 + f_2 + f_3 + f_4 = 1$.

Similarly, by exploiting the symmetry of f_3 and f_4, we define the symmetric form of F_1.

$$\hat{F}_1(f_1,f_2,f_3,y_1,y_2,y_3)=\frac{(1-f_1)(1-y_1)}{\hat{A}}$$

where

$$\hat{A}=(1-f_1)(1-y_1)+(1-f_2)(1-y_2)+2(1-f_3)(1-y_3).$$

Then we can define the symmetric form of α

$$\hat{\alpha}=\frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}}\left(\hat{A}(1-\hat{F}_1)P_1+P_2+P_3\right)$$

where

$$\begin{aligned}P_1&=\frac{1}{(1-y_1)\Phi(y_1)}-\frac{4Mf_2}{1-f_1},\\ P_2&=\frac{1-f_2}{\Phi(y_2)}+4Mf_1(1-y_2),\\ P_3&=\frac{2(1-f_3)}{\Phi(y_3)}+8M(f_1+f_2)(1-y_3).\end{aligned}$$

So $\hat{\alpha}$ is a function of $\mathbf{f}, \mathbf{y} \in [0, \frac{1}{2}]^3$.

Lemma 6.7. *For all $\mathbf{f}, \mathbf{y} \in [0, \frac{1}{2}]^4$ such that $\frac{1}{13} \le f_3, f_4 \le \frac{1}{2}$ and $f_1 + f_2 + f_3 + f_4 = 1$, there exists $\hat{f}_3, \hat{y}_3 \in [0, \frac{1}{2}]$ such that $f_1 + f_2 + 2\hat{f}_3 = 1$ and*

$$\alpha(\mathbf{f},\mathbf{y})\le\kappa\cdot\hat{\alpha}(f_1,f_2,\hat{f}_3,y_1,y_2,\hat{y}_3)$$

where $\kappa = \frac{1038}{1000}$.

Proof. Let $w_j = (1-f_j)(1-y_j)$ for $j = 3, 4$, and denote $A = A(w_1, w_2, w_3, w_4) = \sum_{j=1}^{4} w_j$ be the denominator of F_k. Let $t = f_1 + f_2$, then

$$\begin{aligned}\alpha&=\frac{\Phi(F_1)F_1}{A}\left(A(1-F_1)P_1+P_2+\right.\\ &\quad\left.\sum_{j=3}^{4}\frac{1-f_j}{\Phi(y_j)}+4M(f_1+f_2)(1-y_j)\right)\\ &=\frac{\Phi(F_1)F_1}{A}\left(A(1-F_1)P_1+P_2\right.\\ &\quad\left.+\sum_{j=3}^{4}\frac{1-f_j}{\Phi(1-\frac{w_j}{1-f_j})}+4M(f_1+f_2)\cdot\frac{w_j}{1-f_j}\right)\\ &=\frac{\Phi(F_1)F_1}{A}\left(A(1-F_1)P_1+P_2+\sum_{j=3}^{4}G_t(w_j,f_j)\right).\end{aligned}$$

Take $\hat{w}_3$ and $\hat{f}_3$ such that $2\hat{w}_3 = w_3 + w_4, 2\hat{f}_3 = f_3 + f_4$, and take $\hat{y}_3 = 1 - \frac{\hat{w}_3}{1-\hat{f}_3}$. Then we have $f_1 + f_2 + 2\hat{f}_3 = f_1+f_2+f_3+f_4 = 1$. Let $\hat{A}(w_1, w_2, w_3) = w_1+w_2+2w_3$, then clearly $A(w_1, w_2, w_3, w_4) = \hat{A}(w_1, w_2, \hat{w}_3)$. Since $f_1 + f_2 \in [0, 1]$ by Lemma 6.4 we have

$$\begin{aligned}
\alpha(\mathbf{f}, \mathbf{y}) &\leq \frac{\Phi(F_1)F_1}{\hat{A}} \left(\hat{A}(1-\hat{F}_1)P_1 + P_2 + 2G_t(\hat{w}_3, \hat{f}_3)\right) \\
&= \frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}} \Bigg(\hat{A}(1-\hat{F}_1)P_1 + P_2 \\
&\qquad + 2\kappa \cdot \left(\frac{1-\hat{f}_3}{\Phi(\hat{y}_3)} + 4M(f_1+f_2)(1-\hat{y}_3)\right)\Bigg) \\
&\leq \kappa \cdot \frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}} \Bigg(\hat{A}(1-\hat{F}_1)P_1 + P_2 \\
&\qquad + \frac{2(1-\hat{f}_3)}{\Phi(\hat{y}_3)} + 8M(f_1+f_2)(1-\hat{y}_3)\Bigg) \\
&= \kappa \cdot \hat{\alpha}(f_1, f_2, \hat{f}_3, y_1, y_2, \hat{y}_3).
\end{aligned}$$

Lemma 6.8. *For all $f_1, f_2, f_3, y_1, y_2, y_3 \in [0, \frac{1}{2}]$ satisfying*

$$\begin{aligned}
f_1 + f_2 + 2f_3 &= 1, \\
6f_1 + f_2 - 1 &< 0, \\
4f_1 + f_2 - \frac{3}{4} &< 0,
\end{aligned}$$

and

$$\frac{1}{13} \leq f_1 \leq \frac{1}{2}, \quad 0 \leq f_2, f_3 \leq \frac{1}{2},$$

we have

$$\hat{\alpha}(f_1, f_2, f_3, y_1, y_2, y_3) \leq \frac{9163}{10000}.$$

Proof. Recall that

$$\hat{\alpha} = \frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}} \left(\hat{A}(1-\hat{F}_1)P_1 + P_2 + P_3\right)$$

where

$$\begin{aligned}
P_1 &= \frac{1}{(1-y_1)\Phi(y_1)} - \frac{4Mf_2}{1-f_1}, \\
P_2 &= \frac{1-f_2}{\Phi(y_2)} + 4Mf_1(1-y_2), \\
P_3 &= \frac{2(1-f_3)}{\Phi(y_3)} + 8M(f_1+f_2)(1-y_3) \\
&= \frac{1+f_1+f_2}{\Phi(y_3)} + 8M(f_1+f_2)(1-y_3).
\end{aligned}$$

Denote

$$A_1 = \hat{A}(1-\hat{F}_1) = (1-f_2)(1-y_2) + 2(1-f_3)(1-y_3).$$

So

$$\hat{\alpha} = \frac{2\left(A_1P_1 + P_2 + P_3\right)}{A_1 - (1-y_1)(1-f_1)}.$$

We substitute P_1 for $P_1' = \frac{1}{(1-y_1)\Phi(y_1)} - \frac{4Mf_2}{1-1/13} \geq P_1$ and obtain an upper bound

$$\hat{\alpha} \leq \frac{2(A_1P_1' + P_2 + P_3)}{A_1 - (1-y_1)(1-f_1)}.$$

Notice now both numerator and denominator are linear functions of f_1. Therefore it reaches the maximum value when f_1 is at its boundary. The next step is to let f_1 take its boundary values and simplify the formula.

1. $f_1 = \frac{1}{6}(1-f_2)$.

$$\alpha_1 = \frac{2(A_1P_1' + P_2' + P_3')}{A_1 - (1-y_1)(1-\frac{1}{6}(1-f_2))}$$

where

$$\begin{aligned}
P_2' &= \frac{1-f_2}{\Phi(y_2)} + \frac{2}{3}M(1-f_2)(1-y_2), \\
P_3' &= \frac{7+5f_2}{6\Phi(y_3)} + \frac{4}{3}M(5f_2+1)(1-y_3).
\end{aligned}$$

It can be rigorously proved by *Mathematica* that $\alpha_1 \leq \frac{9138}{10000}$. The codes are in Section 8.

2. $f_1 = \frac{1}{4}\left(\frac{3}{4} - f_2\right)$.

$$\alpha_2 = \frac{2(A_1P_1' + P_2' + P_3')}{A_1 - (1-y_1)(1-\frac{1}{4}(\frac{3}{4}-f_2))}$$

where

$$\begin{aligned}
P_2' &= \frac{1-f_2}{\Phi(y_2)} + M\left(\frac{3}{4} - f_2\right)(1-y_2), \\
P_3' &= \frac{19+12f_2}{16\Phi(y_3)} + 6M\left(f_2 + \frac{1}{4}\right)(1-y_3).
\end{aligned}$$

It can be rigorously proved by *Mathematica* that $\alpha_1 \leq \frac{9163}{10000}$. The codes are in Section 8.

3. $f_1 = \frac{1}{13}$.

$$\alpha_3 = \frac{2(A_1P_1' + P_2' + P_3')}{A_1 - (1-y_1)(1-\frac{1}{13})}$$

where

$$\begin{aligned}
P_2' &= \frac{1-f_2}{\Phi(y_2)} + \frac{4}{13}M(1-y_2), \\
P_3' &= \frac{14+f_2}{13\Phi(y_3)} + 8M\left(f_2 + \frac{1}{13}\right)(1-y_3).
\end{aligned}$$

It can be rigorously proved by *Mathematica* that $\alpha_3 \leq \frac{9102}{10000}$. The codes are in Section 8.

To conclude we have
$$\hat{\alpha} \le \max\left\{\frac{9138}{10000}, \frac{9163}{10000}, \frac{9102}{10000}\right\} = \frac{9163}{10000}.$$

The discussion of absolute values provides the condition for Lemma 6.7, and combining Lemma 6.8 gives
$$\alpha(\mathbf{f}, \mathbf{y}) \le \kappa \cdot \hat{\alpha}(f_1, f_2, \hat{f}_3, y_1, y_2, \hat{y}_3) \le \frac{1038}{1000} \cdot \frac{9163}{10000} < \frac{9512}{10000}.$$

To summarize the analysis in section 6.2.1, we have
$$\alpha(\mathbf{f}, \mathbf{y}) \le \max\left\{\frac{9512}{10000}, \frac{9996}{10000}\right\} = \frac{9996}{10000}.$$

6.2.2 $d_1 = 1$ When $d_1 = 1$ we need to bound $\alpha(\mathbf{x}, \mathbf{y}) = P_1(\mathbf{f}, \mathbf{y}) + P_2(\mathbf{f}, \mathbf{y})$. Furthermore, if $1 \in L(v_1)$ then we still have $\frac{1}{13} \le f_1 \le \frac{1}{2}, 0 \le f_2, f_3, f_4 \le \frac{1}{2}$. In this case, the proof in Section 6.2.1 can all go through once we obtain the symmetric form of α by the following lemma. This is a modified version of Lemma 6.4 that can fit the situation of $d_1 = 1$.

LEMMA 6.9. *For all $w_1, w_2, w_3 \in [0, \frac{1}{2}]$ and $f_1, f_2, f_3 \in [0, \frac{1}{2}]$ such that $\frac{1-f_i}{2} \le w_i \le 1 - f_i$ where $i = 1, 2, 3$, we have $\frac{1}{2}\left(G_\xi(w_1, f_1) + G_\xi(w_2, f_2)\right) \le \kappa \cdot G_\xi\left(\frac{w_1+w_2}{2}, \frac{f_1+f_2}{2}\right)$ and $\frac{1}{3}\left(G_\xi(w_1, f_1) + G_\xi(w_2, f_2) + G_\xi(w_3, f_3)\right) \le \kappa \cdot G_\xi\left(\frac{w_1+w_2+w_3}{3}, \frac{f_1+f_2+f_3}{3}\right)$ hold for any $\xi \in [0, \frac{1}{4}]$, where $\kappa = \frac{1019}{1000}$.*

Proof. The proof is almost the same as Lemma 6.4, except that here we only need to prove for $\xi = \frac{1}{4}$. This is also achieved by proving that for all $w_1, w_2, w_3 \in [0, \frac{1}{2}]$ and $f_1, f_2, f_3 \in [0, \frac{1}{2}]$ such that $\frac{1-f_i}{2} \le w_i \le 1 - f_i, i = 1, 2, 3$, we have $\frac{1}{2}(G_{\frac{1}{4}}(w_1, f_1) + G_{\frac{1}{4}}(w_2, f_2)) \le \kappa_1 \cdot G_{\frac{1}{4}}\left(\frac{w_1+w_2}{2}, \frac{f_1+f_2}{2}\right)$ and $\frac{1}{3}(G_{\frac{1}{4}}(w_1, f_1) + 2G_{\frac{1}{4}}(w_2, f_2)) \le \kappa_2 \cdot G_{\frac{1}{4}}\left(\frac{w_1+2w_2}{3}, \frac{f_1+2f_2}{3}\right)$, where $\kappa_1 = \frac{1009}{1000}, \kappa_2 = \frac{1009}{1000}$ and $\kappa_1\kappa_2 \le \kappa$. The *Mathematica* code to verify the lemma is in Section 8.

Now it remains to handle the case when $1 \notin L(v_1)$. So in the rest of this section we will assume that $f_1 = 0$.

According to the convention in Algorithm 3, we have either $1 \notin L(v_2)$ or $d_2 = 0$. We will defer the discussion of this $d_2 = 0$ case to the end of this section. If $1 \notin L(v_2)$ we have $f_1 = y_1 = 0$ and this is true for both actual value and computed value. So we fix f_1, y_1 to be zero in our recursion and discuss the contraction rate of this partially fixed function
$$F_1 = \frac{1}{1 + \sum_{j=2}^4 (1 - f_j)(1 - y_j)}.$$

The contraction rate should not involve the derivatives of f_1 and y_1, namely
$$\alpha(\mathbf{f}, \mathbf{y}) = P_1(\mathbf{f}, \mathbf{y}) + P_2(\mathbf{f}, \mathbf{y})$$
where
$$P_1(\mathbf{f}, \mathbf{y}) = \Phi(F_1)F_1 \cdot M \cdot \sum_{j \in L(v_1)\setminus\{1\}} f_j \left|1 - \sum_{\substack{k=1\\k\ne j}}^4 \frac{F_k}{1 - f_k}\right|$$
$$P_2(\mathbf{f}, \mathbf{y}) = \Phi(F_1)F_1 \cdot \sum_{j=2}^4 \frac{F_j}{(1 - y_j)\Phi(y_j)}$$
and $F_k = \frac{(1-f_k)(1-y_k)}{1+\sum_{j=2}^4(1-f_j)(1-y_j)}$ is also partially fixed accordingly.

Discussion on the absolute values. Let $D_j \triangleq 1 - \sum_{\substack{k=1\\k\ne j}}^4 \frac{F_k}{1-f_k}$ for $j = 2, 3, 4$. Recall that
$$\sum_{j=2}^4 D_j = \sum_{j \in L(v_1)\setminus\{1\}} f_j \left(1 - \sum_{\substack{k=1\\k\ne j}}^4 \frac{F_k}{1 - f_k}\right) = f_1 \sum_{j=2}^4 \frac{F_j}{1 - f_j} = 0,$$
so it cannot be the case that all D_j's have the same sign. We will always, without loss of generality, assume D_2 has the opposite sign against others. Then $|D_2| + |D_3| + |D_4|$ is either $2D_2$ or $-2D_2$.

Case 1: D_2 is negative. In this case
$$\alpha(\mathbf{f}, \mathbf{y}) = \Phi(F_1)F_1\left(M \cdot (-2D_2) + \sum_{j=2}^4 \frac{F_j}{(1 - y_j)\Phi(y_j)}\right).$$

Denote $A \triangleq 1 + \sum_{j=2}^4 (1 - f_j)(1 - y_j)$ the denominator of F_1.

We first consider the case when $y_j = \frac{1}{2}$ for some $j \in \{2, 3, 4\}$. By Theorem 5.1 we know that all y_j's should be accurately computed given the recursion depth D is at least 3. So we can further discard all derivatives of y_j and obtain
$$\alpha(\mathbf{f}, \mathbf{y}) = 2Mf_2 \cdot \Phi(F_1)F_1\left(\sum_{\substack{k=1\\k\ne 2}}^4 \frac{F_k}{1 - f_k} - 1\right) = \frac{4Mf_2(3 - y_3 - y_4 - A)}{A - 2}.$$

Notice that α is monotonically increasing on y_2, so we take $y_2 = \frac{1}{2}$. After substituting $1 - f_2$ for $f_3 + f_4$ we get $\alpha(\mathbf{f}, \mathbf{y}) \le 4Mf_2 \cdot \frac{f_3(\frac{1}{2}-y_3)+f_4(\frac{1}{2}-y_4)}{(\frac{1}{2}-y_3)(1-f_3)+(\frac{1}{2}-y_4)(1-f_4)} \le 2M$ where the last inequality is due to $f_2, f_3, f_4 \le \frac{1}{2}$ and the monotonicity on f_3 and f_4.

On the other aspect, if $y_j \neq \frac{1}{2}$ for all $j \in \{2,3,4\}$, then by Theorem 5.1 we have $y_j \leq \frac{6}{13}$ for all $j \in \{2,3,4\}$ since d_2 is at most 1. Let $w_j = (1-f_j)(1-y_j)$, by Lemma 6.9

$$\alpha(\mathbf{f},\mathbf{y}) = \Phi(F_1)F_1\left(2Mf_2\left(\sum_{\substack{k=1\\k\neq 2}}^{4}\frac{F_k}{1-f_k}-1\right)+\sum_{j=2}^{4}\frac{F_j}{(1-y_j)\Phi(y_j)}\right)$$
$$= \frac{\Phi(F_1)F_1}{A}\left(\frac{1-f_2}{\Phi(y_2)}+2Mf_2(1-A)+\sum_{j=3}^{4}\frac{1-f_j}{\Phi(y_j)}+2Mf_2(1-y_j)\right)$$
$$= \frac{\Phi(F_1)F_1}{A}\left(\frac{1-f_2}{\Phi(y_2)}+2Mf_2(1-A)+\sum_{j=3}^{4}\frac{1-f_j}{\Phi(1-\frac{w_j}{1-f_j})}+2Mf_2\frac{w_j}{1-f_j}\right)$$
$$= \frac{\Phi(F_1)F_1}{A}\cdot\left(\frac{1-f_2}{\Phi(y_2)}+2Mf_2(1-A)+\sum_{j=3}^{4}G_{\frac{f_2}{2}}(w_j,f_j)\right)$$
$$\leq \kappa\cdot\frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}}\cdot\left(\frac{1-f_2}{\Phi(y_2)}+2Mf_2(1-\hat{A})+2G_{\frac{f_2}{2}}(\hat{w}_3,\hat{f}_3)\right),$$

where $\hat{w}_3 = \frac{w_3+w_4}{2}$, $\hat{f}_3 = \frac{f_3+f_4}{2}$, $\hat{F}_1 = \frac{1}{1+w_2+2\hat{w}_3}$ and $\hat{A} = 1+w_2+2\hat{w}_3$. If we take $\hat{y}_3 = 1-\frac{\hat{w}_3}{1-\hat{f}_3}$ then we can get the symmetric form of α:

$$\hat{\alpha}(\mathbf{f},\mathbf{y}) = \frac{\Phi(\hat{F}_1)\hat{F}_1}{\hat{A}}\left(\frac{1-f_2}{\Phi(y_2)}+2Mf_2(1-\hat{A})+\frac{2(1-\hat{f}_3)}{\Phi(\hat{y}_3)}+4Mf_2(1-\hat{y}_3)\right).$$

LEMMA 6.10. *For all $f_2, f_3, y_2, y_3 \in [0,\frac{1}{2}]$ satisfying*

$$f_2+2f_3 = 1,$$
$$\frac{1}{13}\leq f_2\leq\frac{1}{2},$$
$$0\leq y_2,y_3\leq\frac{6}{13},$$

we have

$$\hat{\alpha}(f_2,f_3,y_2,y_3)\leq\frac{9231}{10000}.$$

Proof. The lemma can be verified by *Mathematica*. The codes are in Section 8.

In conclusion we have

$$\alpha(\mathbf{f},\mathbf{y}) \leq \max\left\{2M, \kappa\cdot\hat{\alpha}(f_2,\hat{f}_3,y_2,\hat{y}_3)\right\}$$
$$\leq \max\left\{2M, \frac{1018}{1000}\cdot\frac{9231}{10000}\right\} < \frac{94}{100}.$$

Case 2: D_2 is positive In this case

$$\alpha(\mathbf{f},\mathbf{y}) = \Phi(F_1)F_1\left(M\cdot 2D_2+\sum_{j=2}^{4}\frac{F_j}{(1-y_j)\Phi(y_j)}\right).$$

As we did in Case 1, we first consider when $y_j = \frac{1}{2}$ for some $j \in \{2,3,4\}$. We similarly obtain

$$\alpha(\mathbf{f},\mathbf{y}) = 2Mf_2\cdot\Phi(F_1)F_1\left(1-\sum_{\substack{k=1\\k\neq 2}}^{4}\frac{F_k}{1-f_k}\right)$$
$$= \frac{4Mf_2(A-3+y_3+y_4)}{A-2}.$$

Notice that $\alpha(\mathbf{f},\mathbf{y})$ is monotonically increasing on y_3 and y_4, so we take $y_3 = y_4 = \frac{1}{2}$ which yields

$$\alpha(\mathbf{f},\mathbf{y}) \leq 4Mf_2 \leq 2M.$$

Now we once more assume $y_j \leq \frac{6}{13}$ for all $j \in \{2,3,4\}$. Recall $\lambda = \frac{9996}{10000}$, we now prove that

$$\alpha(\mathbf{f},\mathbf{y}) = \frac{\sum_{j=2}^{4}(1-f_j)y_j(\frac{1}{2}-y_j)+2Mf_2(A-3+y_3+y_4)}{\frac{A}{2}-1} < \lambda.$$

Since the denominator of $\alpha(\mathbf{f},\mathbf{y})$ is positive, $\alpha(\mathbf{f},\mathbf{y}) < \lambda$ is equivalent to $G \triangleq \sum_{j=2}^{4}(1-f_j)y_j(\frac{1}{2}-y_j)+2Mf_2\left(A-3+y_3+y_4\right)-\lambda\left(\frac{1}{2}A-1\right) < 0$.

Note that G is quadratic on y_3, we can write it as $G = -(1-f_3)y_3^2 + \left(2Mf_2+\frac{1}{2}(1-f_3)+\frac{1}{2}\lambda(1-f_3)-2Mf_2(1-f_3)\right)y_3 + C = (1-f_3)\left(-y_3^2+\left(\frac{2Mf_2}{1-f_3}+\frac{1+\lambda}{2}-2Mf_2\right)y_3\right)+C$, where C is a polynomial containing no y_3.

Therefore, G is increasing in $[-\infty,x_0]$ where $x_0 = \frac{Mf_2}{1-f_3}+\frac{1+\lambda}{4}-Mf_2 \geq \frac{1+\lambda}{4} \geq \frac{6}{13}$. Since y_3 and y_4 are symmetric, the same argument holds for y_4.

We only need to prove that $G' \triangleq G|_{y_3=y_4=\frac{6}{13}} < 0$. Applying $f_2+f_3+f_4 = 1$, a direct calculation yields $G' = \frac{2}{13}Mf_2^2(13y_2-6)+\frac{1}{338}(6-91\lambda+169y_2+169\lambda y_2-338y_2^2)+\frac{1}{338}f_2\left((13y_2-6)(26y_2-1-52M-13\lambda)\right)$.

Since $y_2 \leq \frac{6}{13}$, G' is increasing in $[-\infty,x_1]$ where $x_1 = \frac{1+52M+13\lambda-26y_2}{104M} > \frac{1}{2}$. Therefore, we only need to prove that $G'' \triangleq G'|_{f_2=\frac{1}{2}} = \frac{9}{338} + \frac{3M}{13} - \frac{2\lambda}{13} + \left(\frac{1}{4} - \frac{M}{2} + \frac{\lambda}{4}\right) y_2 - \frac{y_2^2}{2} < 0$, which holds for $y_2 \in \left[0, \frac{6}{13}\right]$.

This concludes that $\alpha(\mathbf{f}, \mathbf{y}) \le \max\{2M, \lambda\} = \lambda$.

The case $d_2 = 0$. At last we come to the discussion for $d_2 = 0$. In this case y_1 is not necessarily 0, but all y_j's are accurately computed. Redefine

$$F_1 = \frac{1 - y_1}{1 - y_1 + \sum_{j=2}^{4}(1 - f_j)(1 - y_j)},$$
$$A = 1 - y_1 + \sum_{j=2}^{4}(1 - f_j)(1 - y_j).$$

As we did before, we shall discard the derivatives of y_j's and assume D_2 has the opposite sign against others. Now

$$\begin{aligned}
\alpha(\mathbf{f}, \mathbf{y}) &= 2Mf_2 \cdot \Phi(F_1)F_1 \left| 1 - \sum_{\substack{k=1\\k\neq 2}}^{4} \frac{F_k}{1 - f_k} \right| \\
&= \frac{4Mf_2 \left| A - (1 - y_1) - (1 - y_3) - (1 - y_4) \right|}{A - 2(1 - y_1)} \\
&= \frac{4Mf_2 \left| \sum_{j=2}^{4}(1 - f_j)(1 - y_j) - 2 + y_3 + y_4 \right|}{\sum_{j=2}^{4}(1 - f_j)(1 - y_j) - (1 - y_1)}
\end{aligned}$$

is monotonically decreasing on y_1. So we can take $y_1 = 0$ and this is reduced to a situation we have discussed before.

To summarize the analysis in section 6.2.2, we have

$$\alpha(\mathbf{f}, \mathbf{y}) \le \max\left\{\frac{94}{100}, \lambda\right\} = \lambda.$$

So far we have exhausted all possible cases when $\mathsf{deg}_G(v) = 2$. Putting together the conclusions of section 6.1 and section 6.2, we can finish the proof of Lemma 6.2.

6.3 Proof of Theorem 6.1 By the discussion on cases in 6.1 and 6.2, we have finished the proof of Lemma 6.2 so far.

Thus we can prove Theorem 6.1 now.

Proof. [Proof of Theorem 6.1] Let $\lambda = \frac{9996}{10000}$ be constant.

We first claim that if a vertex v satisfies $\mathsf{deg}_G(v) \le 2$ and $|L(v)| \ge \mathsf{deg}_G(v) + 2$, then one of the following statements holds:

- $P(G, L, v, i, D) = \mathbf{Pr}_{G,L}[c(v) = i]$;
- $|\varphi(P(G, L, v, i, D)) - \varphi(\mathbf{Pr}_{G,L}[c(v) = i])| \le C_1 \cdot \lambda^{D-2}$, where $\varphi(x) = 2\ln x - 2\ln\left(\frac{1}{2} - x\right)$ and $C_1 > 0$ is a constant.

Given the claim, we have for some constant $C_2 > 0$, it holds that

$$\begin{aligned}
&|P(G, L, v, i, D) - \mathbf{Pr}_{G,L}[c(v) = i]| \\
&= \frac{1}{\Phi(\tilde{x})} \cdot |\varphi(P(G, L, v, i, D)) - \varphi(\mathbf{Pr}_{G,L}[c(v) = i])| \\
&\le C_2 \cdot \lambda^D,
\end{aligned}$$

where $\Phi(x) \triangleq \varphi'(x) = \frac{1}{x\left(\frac{1}{2} - x\right)}$ and $\tilde{x}$ is some real between $\varphi(P(G, L, v, i, D))$ and $\varphi(\mathbf{Pr}_{G,L}[c(v) = i])$.

Now assume $(G = (V, E), L)$ satisfies $|L(v)| \ge \mathsf{deg}_G(v) + 1$ for every $v \in V$. Let $v \in V$ be an arbitrary vertex and consider the computation tree of $P(G, L, v, i, D)$. According to the construction in Section 2, all the smaller instances $P(G', L', v', i', D')$ called by the procedure satisfy $|L(v)| \ge \mathsf{deg}_{G'}(v') + 2$ and $\mathsf{deg}_{G'}(v') \le 2$, i.e., the condition specified in the above claim. Further note that in all cases, the 1-norm of the gradients of our recursions

- $F(x, y, z) = \frac{1-x}{3-x-y}$, if $\mathsf{deg}_G(v) = 1$ and $|L(v)| = 2$;
- $F(x, y) = \frac{1-x}{2+y}$, if $\mathsf{deg}_G(v) = 1$ and $|L(v)| = 3$;
- $F(x) = \frac{1-x}{3}$, if $\mathsf{deg}_G(v) = 1$ and $|L(v)| = 4$;
- $F(\mathbf{f}, \mathbf{y}) = \frac{(1-f_i)(1-y_i)}{\sum_{j\in L(v)}(1-f_j)(1-y_j)}$, if $\mathsf{deg}_G(v) = 2$;
- $F(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \frac{(1-x_i)(1-y_i)(1-z_i)}{\sum_{j\in L(v)}(1-x_j)(1-y_j)(1-z_j)}$, if $\mathsf{deg}_G(v) = 3$,

are bounded above by some constants for parameters in the range $[0, \frac{1}{2}]$. Therefore it follows from the mean value theorem and the claim that

$$|P(G, L, v, i, D) - \mathbf{Pr}_{G,L}[c(v) = i]| \le C \cdot \lambda^D.$$

for some constant $C > 0$.

It remains to prove the claim. We apply induction on D. The base case is that $D = 2$. It follows from Theorem 5.1 and Lemma 5.2 that if $\mathbf{Pr}_{G,L}[c(v) = i]$ is 0 or $\frac{1}{2}$, then the algorithm return the correct value, i.e., $P(G, L, v, i, D) = \mathbf{Pr}_{G,L}[c(v) = i]$. Otherwise, the function $\varphi(\cdot)$ is bounded from above and thus the claim holds. For $D > 2$, the claim follows from the induction hypothesis and Lemma 6.1.

7 Proof of the Main Theorem

In this section, we prove Theorem 1.1. We start the proof by first analyzing the running time of Algorithm 1.

Let $G = (V, E)$ be a graph with $|V| = n$, L be its color lists, $v \in V$ be a vertex, $i \in \{1, 2, 3, 4\}$ be a color and D be nonnegative integer. Let $\tau(G, L, v, i, D)$ denote the running time of the procedure $P(G, L, v, i, D)$, then we have:

LEMMA 7.1. *If* $\mathsf{deg}_G(v) \le 2$*, then* $\tau(G,L,v,i,D) = O\left(n^3 12^D\right)$.

Proof. We apply induction on n to show that for some constant $C \ge 0$, $\tau(G,L,v,i,D) \le C \cdot n^3 12^D$. The base case is that $n = 1$, then the algorithm terminates in constant time.

For general n, we need to analyze cases $\mathsf{deg}_G(v) = 1, 2$ respectively.

Case $\mathsf{deg}_G(v) = 1$**:** Algorithm 2 contains two subcases. We use an adjacency matrix to represent a graph. Thus we can construct in n^2 time the graph G_v which contains $n-1$ vertices. We then have the following recursions for the two cases respectively (assuming notations in the description of Algorithm 2):

$$\begin{aligned}&\tau(G,L,v,i,D) \le \tau(G_v, L_{1,i}, v_1, i, D-1) + n^2;\\&\tau(G,L,v,i,D) \le \tau(G_v, L_{1,i}, v_1, i, D-1) +\\&\qquad\qquad\qquad \tau(G_v, L_{1,j}, v_1, j, D-1) + n^2.\end{aligned}$$

Then the lemma follows from the induction hypothesis.

Case $\mathsf{deg}_G(v) = 2$**:** Algorithm 3 has at most 12 branches, we have (assuming notations in the description of Algorithm 3):

$$\begin{aligned}\tau(G,L,v,i,D) \le &\sum_{k\in d_1}\sum_{w\in L(v_1)} \tau(G_{v,v_1}, L'_{k,w}, w, D-1)\\&+ \sum_{j\in L(v)} \tau(G_v, L_{2,j}, j, D-1) + n^2.\end{aligned}$$

Then the lemma follows from the induction hypothesis.

If $\mathsf{deg}_G(v) = 3$, then the algorithm $P(G,L,v,i,D)$ will call $P3(G,L,v,i,D)$ described in Algorithm 4. However, since the maximum degree of G is at most three hence in further recursion call to Algorithm 1, the degree of a vertex decreases by at least one. Therefore Algorithm 4 can be called at most once. Combining Lemma 7.1, we have

LEMMA 7.2. $\tau(G,L,v,i,D) = O\left(n^3 12^D\right)$.

Now we prove Lemma 3.1.

Proof. [Proof of Lemma 3.1] First, we need to bound the value $\mathbf{Pr}_{G,L}[c(v) = i]$ on the computation tree. If $\mathbf{Pr}_{G,L}[c(v) = i] = 0$ then it is clear to see $P(G,L,v,i,D) = 0$ thus we are done. Otherwise we have $\mathbf{Pr}_{G,L}[c(v) = i] \ge \frac{1}{13}$ if (G,L,v) is a *reachable* instance. In previous discussion we know that (G,L,v) is on the root of our computation tree if this instance is not *reachable*. In this case

$$\mathbf{Pr}_{G,L}[c(v) = i] = \frac{\prod_{k=1}^{d}(1 - \mathbf{Pr}_{G_v,L_{k,i}}[c(v_k) = i])}{\sum_{j\in L(v)}\prod_{k=1}^{d}(1 - \mathbf{Pr}_{G_v,L_{k,j}}[c(v_k) = j])},$$

where $d = \mathsf{deg}_G(v) \le 3$ and $|L(v)| \le 4$. It yields

$$\mathbf{Pr}_{G,L}[c(v) = i] \ge \frac{\left(1 - \frac{1}{2}\right)^3}{\left(1 - \frac{1}{2}\right)^3 + 1 + 1 + 1} = \frac{1}{25}.$$

Combining with the bound of reachable cases it implies $\mathbf{Pr}_{G,L}[c(v) = i] \ge \frac{1}{25}$ for all instances in computation tree.

By Theorem 6.1, there exists constants $\lambda = \frac{9996}{10000}$ and $C > 0$ such that for every list-coloring instance (G,L) satisfying conditions in the statement of the Lemma, it holds that

$$|P(G,L,v,i,D) - \mathbf{Pr}_{G,L}[c(v) = i]| \le C \cdot \lambda^{D-3}$$

for all $D \ge 3$.

For any $0 < \varepsilon < 1$, let t be the smallest integer such that $C \cdot \lambda^{t-3} \le \frac{\varepsilon}{25}$ and let $\hat{p} = P(G,L,v,i,t)$. We can show that Algorithm 1 up to depth t is the algorithm outputs $\hat{p}$ such that

$$(1-\varepsilon)\hat{p} \le \mathbf{Pr}_{G,L}[c(v) = i] \le (1+\varepsilon)\hat{p}$$

in time $\mathrm{poly}(|V|, \frac{1}{\varepsilon})$.

Theorem 6.1 implies

$$\mathbf{Pr}_{G,L}[c(v) = i] - \frac{\varepsilon}{25} \le \hat{p} \le \mathbf{Pr}_{G,L}[c(v) = i] + \frac{\varepsilon}{25}$$

and thus by the bound of $\mathbf{Pr}_{G,L}[c(v) = i]$ above it holds that

$$(1-\varepsilon)\mathbf{Pr}_{G,L}[c(v) = i] \le \hat{p} \le (1+\varepsilon)\mathbf{Pr}_{G,L}[c(v) = i]\ .$$

So

$$(1-\varepsilon)\hat{p} \le \frac{\hat{p}}{1+\varepsilon} \le \mathbf{Pr}_{G,L}[c(v) = i] \le \frac{\hat{p}}{1-\varepsilon} \le (1+\varepsilon)\hat{p}\,.$$

Next we show that Algorithm 1 up to depth t is a polynomial time algorithm with respect to $|V|$ and $\frac{1}{\varepsilon}$. By Lemma 7.2, $\tau(G,L,v,x,t) = O\left(n^3 12^t\right)$. Since t is the smallest integer such that $C \cdot \lambda^{t-3} \le \frac{\varepsilon}{25}$, we have

$$t - 4 \le \log_\lambda \frac{\varepsilon}{25C} \le t - 3\,,$$

which implies $\tau(G,L,v,x,t) = O\left(n^3 12^{\log_\lambda \frac{\varepsilon}{25C}}\right) = O\left(n^3 \left(\frac{25C}{\varepsilon}\right)^{-\log_\lambda 12}\right)$. λ and C are constants, so $\tau(G,L,v,x,t) = \mathrm{poly}\left(|V|, \frac{1}{\varepsilon}\right)$.

Finally, combining Lemma 3.1 and Lemma 2.1 completes the proof of Theorem 1.1.

8 Computer Assisted Proofs

We utilize some *Mathematica* codes to assist our proof. Due to the limit of space, we omit the codes in the current version. Please refer to the full version of the paper.

Acknowledgements

We would like to thank Jingcheng Liu for insightful discussion.

References

[1] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science (FOCS'97)*, pages 223–231. IEEE, 1997.

[2] Russ Bubley, Martin Dyer, Catherine Greenhill, and Mark Jerrum. On approximately counting colorings of small degree graphs. *SIAM Journal on Computing*, 29(2):387–400, 1999.

[3] Martin Dyer, Abraham D Flaxman, Alan M Frieze, and Eric Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 29(4):450–465, 2006.

[4] Martin Dyer and Alan Frieze. Randomly coloring graphs with lower bounds on girth and maximum degree. *Random Structures & Algorithms*, 23(2):167–179, 2003.

[5] Martin Dyer, Alan Frieze, Thomas P Hayes, and Eric Vigoda. Randomly coloring constant degree graphs. *Random Structures & Algorithms*, 43(2):181–200, 2013.

[6] Andreas Galanis, Daniel Štefankovic, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic ising and hard-core models. *arXiv preprint arXiv:1203.2226*, 2012.

[7] David Gamarnik and Dmitriy Katz. Correlation decay and deterministic FPTAS for counting colorings of a graph. *Journal of Discrete Algorithms*, 12:29–47, 2012.

[8] Thomas P Hayes. Randomly coloring graphs of girth at least five. In *Proceedings of the 35th Annual ACM Symposium on Symposium on Theory of Computing (STOC'03)*, pages 269–278. ACM, 2003.

[9] Thomas P Hayes and Eric Vigoda. A non-markovian coupling for randomly sampling colorings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS'03)*, pages 618–627. IEEE, 2003.

[10] Thomas P Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *The Annals of Applied Probability*, 16(3):1297–1318, 2006.

[11] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures and Algorithms*, 7(2):157–166, 1995.

[12] Johan Jonasson. Uniqueness of uniform random colorings of regular trees. *Statistics & Probability Letters*, 57(3):243–248, 2002.

[13] Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *Proceedings of the 23th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*, pages 922–940. SIAM, 2012.

[14] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'13)*, pages 67–84. SIAM, 2013.

[15] Jingcheng Liu and Pinyan Lu. FPTAS for counting monotone CNF. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'15)*, pages 1531–1548. SIAM, 2015.

[16] Pinyan Lu and Yitong Yin. Improved FPTAS for multi-spin systems. In *Proceedings of APPROX-RANDOM*, pages 639–654. Springer, 2013.

[17] Michael Molloy. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM Journal on Computing*, 33(3):721–737, 2004.

[18] Jesús Salas and Alan D Sokal. Absence of phase transition for antiferromagnetic potts models via the dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86(3-4):551–579, 1997.

[19] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state antiferromagnetic spin systems on bounded degree graphs. *Journal of Statistical Physics*, 155(4):666–686, 2014.

[20] Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*, pages 361–369. IEEE, 2012.

[21] Eric Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3):1555–1569, 2000.

[22] Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'06)*, pages 140–149. ACM, 2006.