On Brooks' Theorem For Sparse Graphs

Jeong Han Kim*†

Abstract

Let G be a graph with maximum degree $\Delta(G)$. In this paper we prove that if the girth $g(G)$ of G is greater than 4 then its chromatic number, $\chi(G)$, satisfies

$$\chi(G) \leq (1+o(1))\frac{\Delta(G)}{\log \Delta(G)}$$

where $o(1)$ goes to zero as $\Delta(G)$ goes to infinity. (Our logarithms are base e.)

More generally, we prove the same bound for the list-chromatic (or choice) number:

$$\chi_l(G) \leq (1+o(1))\frac{\Delta(G)}{\log \Delta(G)}$$

provided $g(G) > 4$.

1 Introduction

In this paper we are focusing on Vizing's question [29] concerning a possible "*Brooks' theorem for sparse graphs*":

Find a best possible upper bound for the chromatic number $\chi(G)$ of a graph G with girth $g(G)$ at least 4 in terms of the maximum degree $\Delta(G)$ of G,

where the *girth* $g(G)$ is the length of shortest cycles of G.

For general graphs G, $\Delta(G)+1$ is a trivial upper bound on $\chi(G)$. Brooks' Theorem [7] gives an exact description of the graphs achieving this bound (the connected ones are just the complete graphs and odd cycles). It is natural to expect that Brooks' bound is very weak for graphs without small cycles or large complete subgraphs, say for graphs of large degree without C_h or K_r-subgraphs (h, r fixed).

The first non-trivial result in this direction was discovered independently by Borodin and Kostochka [5], Catlin [8] and Lawrence [18]: For K_4-free G,

$$\chi(G) \leq (3/4)(\Delta(G)+2).$$

For triangle-free G (i.e. K_3-free), this was improved slightly (10 years later!) by Kostochka [17], who gave the bound

$$\chi(G) \leq (2/3)\Delta(G)+2. \tag{1}$$

---

*Partially supported by a DIMACS Graduate Assistantship and Sloan Foundation Dissertation Fellowship.

†Department of Mathematics, Rutgers University, New Brunswick NJ 08903 (current address: Room 2C-180, Mathematical Sciences Research Center, AT&T Bell Labs, Murray Hill, NJ 07974 USA; Email:jhkim@research.att.com Phone:908-582-2375 Fax:908-582-3340)

This remains the best upper bound known for Vizing's problem, a rather remarkable situation, since the bound (1) differs only by the factor 2/3 from the trivial upper bound.

On the other hand, it is now well-known (see e.g. [4]) that there are graphs G of arbitrarily large girth with

$$\chi(G) \geq C\frac{\Delta(G)}{\log \Delta(G)} \ , \tag{2}$$

where C is a constant. The best constant up to date is asymptotically 1/2 as $\Delta(G)$ goes to infinity. (Our logarithms are base e.)

We may consider how close the lower bound in (2) is to the truth. The situation here is quite analogous to that for independence number. (Recall that the *independence number* $\alpha(G)$ of a graph G is the maximum size of a set of pairwise nonadjacent vertices.) The independence and chromatic numbers are connected by the obvious relation

$$\chi(G) \geq |V(G)|/\alpha(G). \tag{3}$$

For independence number, the classic result of Turán [28] may be stated as

$$\alpha(G) \geq |V(G)|/(t+1),$$

where $t = t(G)$ is the *average* degree of G.

Turán's Theorem is sharp when G is the disjoint union of complete graphs of order $t+1$. On the other hand, Ajtai, Komlós and Szemerédi [2] (see also [1]) proved for triangle free G

$$\alpha(G) = \Omega(\frac{|V(G)|\log t}{t}), \tag{4}$$

and Shearer [24] improved this to

$$\alpha(G) \geq (1-o(t))\frac{|V(G)|\log t}{t}$$

(both bounds as t goes to infinity). These bounds are best possible up to the value of the constant since there are graphs G of arbitrarily large girth with

$$\alpha(G) \leq (2+o(t))\frac{|V(G)|\log t}{t} \ .$$

While the inequality (3) is very weak in general, it is close to the truth in many natural situations, suggesting again that the lower bound in (2) might give the correct order of growth for χ. (Note one cannot bound chromatic number in terms of average degree.)

Provided $g(G) \geq 5$, we prove that the lower bound in (2) gives the correct order of magnitude. In fact our result is more general. Define the *list-chromatic number* (or *choice number*) $\chi_l(G)$ of a graph G to be the minimum integer k such that for every assignment of a set $S(v)$ of k colors to every vertex v of G, there is a legal coloring of G that assigns to each vertex v a color from $S(v)$ (see e.g. [3], [10], or [30]).

Our main result is:

Theorem 1.1 *Let G be a graph. If $g(G) \geq 5$ then*

$$\chi_l(G) \leq (1+o(1))\frac{\Delta(G)}{\log \Delta(G)}$$

where $o(1)$ goes to zero as $\Delta(G)$ goes to infinity.

As a corollary of this theorem we have:

Corollary 1.2 *Let G be a graph. If $g(G) \geq 5$ then*

$$\chi(G) \leq (1 + o(1)) \frac{\Delta(G)}{\log \Delta(G)}$$

where $o(1)$ goes to zero as $\Delta(G)$ goes to infinity.

The basic approach is via the so-called "semirandom" method, some version of which seems to have been first used in [2]. Subsequent, more developed applications were in many papers, e.g. [16], [22] [11], [21] and [14]. See also [12] and [13] for fairly detailed discussions of these developments. The method here is close to that of [14].

In section 2 we sketch the proof of Theorem 1.1. In section 3 we introduce our basic parameters and algorithms, and prove Theorem 1.1 modulo the proof of our Main Lemma on the behavior of these parameters under a random coloring. The Main Lemma says roughly that the behavior of our basic parameters under an appropriate random coloring procedure is highly predictable. There are two parts to this: showing that expected values behave properly; and showing that the parameters are concentrated near their expectations.

Section 4 deals with the Main Lemma at the level of expectations. To prove high concentrations near means of the random variables (in Main Lemma), we develop Azuma-Hoeffding-type martingale inequalities in Section 5, which are thought to be of independent interest. Finally we prove the Main Lemma (the concentration results) in last two sections using these inequalities.

2 Sketch of Methods (Semirandom Methods)

In this section we give a rough idea of the proof of Theorem 1.1. Let G be a graph with girth at least 5 and maximum degree D. Further, suppose we have a set $S(v)$ of size $s \approx D/\log D$ assigned to every vertex v in G. We call $S(v)$ *the set of legal colors for v*. Our object is to find a S-legal coloring on $V(G)$, that is, a function from $V(G)$ to the set of all colors $\Gamma := \cup_{v \in V(G)} S(v)$ such that for all v, $\tau(v) \in S(v)$ and $\tau(v) \neq \tau(w)$ if $v \sim_G w$.

In each stage of our algorithm we will color some set, say X, of uncolored vertices so that the new set X together with the set of already colored vertices is legally colored. Our goal is to reach a situation in which the maximum degree of the graph induced by uncolored vertices is less than the minimum over uncolored v of $|S(v) \setminus \{\text{color of } w : \ w \sim v, w \ \text{ is colored}\}|$. Once we achieve this goal it is enough for us to color the uncolored vertices greedily.

Before telling how to choose such a set X and a legal coloring on it we would like to introduce the following notation: For $W \in V(G)$ and sets $S(w)$ of legal colors for $w \in W$, define for $v \in V(G)$

$$\begin{aligned} N_W(v) &= \{w \in W : w \sim_G v\}, & d_W(v) = |N_W(v)| \\ N_W(v;\gamma) &= \{w \in N_W(v) : \gamma \in S(w)\}, & d_W(v;\gamma) = |N(v;\gamma)| \ . \end{aligned} \tag{5}$$

Also for a set $A \subset V(G)$, we write

$$N_W(A) = \{w \in W : w \sim_G v \quad \text{for some } v \in A \ \} \ .$$

When $W = V(H)$ for an induced subgraph H of G we write $N_H(v)$ etc.. Usually we do not write the subscript W (or H) if the identity of W (or H) is obvious.

The induced subgraph of G on $W \subseteq V(G)$ is denoted by $G[W]$. For the rest of this section we use "$\approx$" to mean approximately equal, deferring precise statements to the next section.

We give a rough version of our coloring algorithm only for the "canonical case" that the graph G is D-regular and all $S(v)$ are the same. In general, the idea is similar, but we need some auxiliary structures (see the last part of this section) to make the evolution as in canonical case. (Note that it is no loss of generality to assume G is D-regular.)

Fix a small $\theta > 0$. First, we define parameters: $\alpha_0 = \beta_0 = 1$ and for $L = D/s \approx \log D$

$$\begin{aligned} \alpha_{i+1} &:= \exp(-\theta\beta_i e^{-\theta\beta_i})\alpha_i \\ \beta_{i+1} &:= (1-(\theta/L)e^{-\theta\beta_i})\beta_i \end{aligned} \tag{6}$$

$i = 0, 1, \cdots$.

Our first algorithm is:

Algorithm 1 (idea)

Initially we set $H_0 = G$, $T_0(v) = S(v)$, $t_0 = |T_0(v)| = s$ and $i = 0$.

(Step 1) In general at the beginning of each stage we will have H_i the subgraph of G induced by the set of uncolored vertices, a list $T_i(v)$ of still-legal colors for each $v \in V(H_i)$. The properties we seek to maintain are

$$\begin{aligned} d_i(v) &\approx \beta_i D \\ t_i(v) &\approx \alpha_i s \\ d_i(v;\gamma) &\approx \alpha_i\beta_i D \end{aligned}$$

for all $v \in V(H_i)$ and $\gamma \in T_i(v)$. (Note these are obvious initially, i.e. i=0.)

Assuming these properties hold, we define the random coloring τ_i according to

$$Pr(\tau_i(v) = \gamma) = \begin{cases} p_i := \theta/(\alpha_i D) & \text{if } \gamma \in T_i(v) \\ 1 - p_i|T_i(v)| & \text{if } \gamma = \Lambda \\ 0 & \text{otherwise} \end{cases}$$

(note that $p_i|T_i(v)| \approx \alpha_i s(\theta/\alpha_i D) \approx \theta/\log D < 1$) independently of all other colors $\tau_i(w)$, and set

$$X_i = \{v \in V(H_i) : \tau_i(v) \neq \Lambda, \quad v \sim_{H_i} w \Rightarrow \tau_i(v) \neq \tau_i(w)\} \ .$$

For the next stage, we should consider the induced subgraph $H_{i+1} := H_i[V(H_i) \setminus X_i]$ and the sets $T_{i+1}(v)$ of still legal colors for each $v \in V(H_i)$[1], defined in the obvious way:

$$T_{i+1}(v) = T_i(v) \setminus \{\tau_i(w) : w \in X_i, w \sim_{H_i} v\}.$$

Also let $t_{i+1}(v) = |T_{i+1}(v)|$.

[1]It is enough for us to consider these sets only for $v \in V(H_{i+1})$, but it is convenient to consider them for all $v \in V(H_i)$.

We then want

$$\begin{aligned} d_{i+1}(v) &\approx \beta_{i+1} D \\ t_{i+1}(v) &\approx \alpha_{i+1} s \\ d_{i+1}(v;\gamma) &\approx \alpha_{i+1}\beta_{i+1} D \end{aligned} \tag{7}$$

The definitions of α_{i+1} and β_{i+1} come from analyzing the (probable) behavior of the parameters under the random coloring specified above. Namely,

$$\alpha_{i+1}/\alpha_i \approx Pr(\gamma \in T_{i+1}(v)) \qquad (\gamma \in T_i(v)), \tag{8}$$

$$\beta_{i+1}/\beta_i \approx Pr(w \in V(H_{i+1})) \qquad (w \in V(H_i)). \tag{9}$$

(These are not hard to see, but for (8) we need the fact that the girth of H_i is at least 5.) Furthermore,

$$\alpha_{i+1}\beta_{i+1}/(\alpha_i\beta_i) \approx Pr(\gamma \in T_{i+1}(w), \quad w \in V(H_{i+1})), \quad (\gamma \in T_i(w)) \tag{10}$$

reflecting the idea that the events "$\gamma \in T_{i+1}(w)$" and "$w \in V(H_{i+1})$" are almost independent.

Once we have X_i and τ_i satisfying the properties (7), we proceed to

(Step 2) Set $i = i + 1$ and go to step 1.
The number of stages will be

$$a := \min\{i : \beta_i \leq D^{-\theta}/(2L)\} \tag{11}$$

(note that a is some power of $\log D$).

The goal of the above algorithm is to reach a situation in which each color degree $d(v;\gamma)$ is small enough relative to $t(v)$. (See (13) below.) To achieve this goal the role of θ is important though it is somewhat technical. Note that for $v \in V(H_i)$

$$Pr(v \in X) = Pr(\tau(v) \neq \Lambda)Pr(\tau(w) \neq \tau(v) \quad \forall w \sim v | \tau(v) \neq \Lambda).$$

and that as θ increases the first factor of the right hand side increases but the second factor decreases. Thus some optimization of θ is in order.

What is left now is to prove that the properties (7) are feasible, that is,

$$Pr(\text{"(7) happens"}) > 0. \tag{12}$$

To prove (12), we will consider the following three steps:

(a) Prove the properties (7) at the level of expectations.
(b) Prove that the random variables $d_{i+1}(v)$ etc. are highly concentrated near their means.
(c) Prove (12) using (b) and the Lovász Local Lemma. (Here it is very easy to show that we have enough independence for the local lemma.)

Parts (a) and (c) are not hard. The only hard part is (b). Though the martingale inequalities of [23], [15], and [14] are quite powerful, we cannot use them directly for $d'(v;\gamma)$. In Section 5 we will develop some martingale inequalities which are useful in our situation.

After running the above algorithm a times we will have

$$d_a(v;\gamma) \lesssim D^{-\theta} t_a(v)/2 \tag{13}$$

by the definition of a. We then run the following more efficient algorithm which prevents excessive error accumulation. Actually, we may not expect any nice behavior of $d_i(v;\gamma)$ $(i > a)$ since these might be too small to disregard error terms. Thus we need a new phase:

Algorithm 2 (idea)

We randomly color all remaining vertices as in Step 1 with $p_i = 1/t_i(v)$ $(i \geq a)$. (We may delete colors from the larger $T_i(v)$'s so that all $t_i(v)$'s are equal.) It turns out that in this phase the degrees will shrink rapidly while the numbers $t(v)$ remain almost constant.

More precisely, the properties we will have are:

$$d_i(v) \lesssim \frac{1}{2} D^{1-(i-a+1)\theta} \tag{14}$$

$$t_i(v) \approx \alpha_a s \ . \tag{15}$$

$i = a, ..., b$ where $b := a + \theta^{-1} + 3$. (Note that for $i = a$ these are obvious by the definition of a. Also, it turns out that we can not run this algorithm more than $\theta^{-1} + 3$ times since the expected degrees $E[d_{b+1}(v)]$, if possible, might be smaller than error terms.) To prove these we do not need any information about $d(v;\gamma)$ other than (13).

Assuming (14) and (15) it is clear that we can achieve our main goal (i.e $d_b(v) < t_b(v)$ for all uncolored v) provided

$$\alpha_b s \geq D^{2\theta} \ , \tag{16}$$

which is possible by choosing suitable θ.

In the general (i.e. non-canonical) case, we do not have (7). Instead, we will have

$$\begin{aligned} d_i(v) &\lesssim \beta_i D \\ t_i(v) &\gtrsim \alpha_i s \\ d_i(v;\gamma) &\lesssim \alpha_i \beta_i D \end{aligned} \tag{17}$$

for all $v \in V(H_i)$ and $\gamma \in T_i(v)$.

The first two properties are in our favor. For example, we may throw away some colors from $T_i(v)$ so that $t_i(v) \approx \alpha_i s$. But the last property may cause some trouble in the next stage. Roughly speaking, the reason is that we cannot control the $t_i(v)$'s well if some color degrees are small and the others are relatively big. To avoid such problems we add some new (artificial) vertices to H_i. These extra vertices are used to force the $t_i(v)$'s (for $v \in V(H_i)$) to behave as in the canonical case, and are then discarded before the beginning of the next stage.

For each $v \in V(H_i)$, $\gamma \in T_i(v)$ with $d_i(v;\gamma) < d_i$, we add $d_i - d_i(v;\gamma)$ new vertices $\{w_1, ..., w_{d_i - d_i(v,\gamma)}\} =: A(v;\gamma)$ all joined to v. (The precise value of $d_i \approx \alpha_i\beta_i D$ will be given below.) For each of these new vertices w_j, we add $d_i - 1$ more new vertices $\{u_1^{(j)}, ..., u_{d_i-1}^{(j)}\} =:$ $B(v;\gamma,w_j)$ all joined to w_j. Finally, set $T_i(z) = \{\gamma\}$ for all $z \in A(v;\gamma) \cup \cup_{j=1}^{d_i - d_i(v,\gamma)} B(v;\gamma,w_j)$. All sets $\{A(v;\gamma)\}_{(v,\gamma)}$ and $\{B(v;\gamma,w_j)\}_{(v,\gamma,j)}$ must be mutually disjoint.

From now on, we write $\hat{H}_i = \hat{H}_i(d_i)$ for the extended graph just defined. Also, we write $\hat{N}_i(v)$, $\hat{N}_i(v;\gamma)$ etc. for $N_{\hat{H}_i}(v)$, $N_{\hat{H}_i}(v;\gamma)$ etc. (see (5)). Note that if each $d_i(v;\gamma)$ is at most d_i then $\hat{d}_i(v;\gamma) = d_i$ for all $v \in V(H_i) \cup \hat{N}(V(H_i))$ with $\gamma \in T_i(v)$.

3 Main Lemma

In this section we define our parameters and algorithms precisely, and give the proof of Theorem 1.1 modulo our Main Lemma (Lemma 3.3) on the behavior of our random coloring procedure.

First, we need some parameters. Let $0 < \eta < 1$, and then choose $0 < \theta < 0.1$ with θ^{-1} an integer and δ such that

$$\frac{1}{2}(1 + \eta e^{\theta} + 2\theta) < \delta < 1 \ . \tag{18}$$

Set $\Delta(G) = D$ and $L = \eta \log D$. Also, let $\mu_0 = \nu_0 = 1$ and for $i = 0, 1, ...$

$$\begin{pmatrix} \mu_{i+1} \\ \nu_{i+1} \end{pmatrix} = \begin{pmatrix} 1 & \beta_i \\ 1/L & 1 + \beta_i/L \end{pmatrix} \begin{pmatrix} \mu_i \\ \nu_i \end{pmatrix}$$

(these parameters are to be used to control the error terms precisely), where as in (6), $\alpha_0 = \beta_0 = 1$ and

$$\begin{aligned} \alpha_{i+1} &:= \exp(-\theta \beta_i e^{-\theta \beta_i}) \alpha_i \\ \beta_{i+1} &:= (1 - (\theta/L) e^{-\theta \beta_i}) \beta_i \ . \end{aligned}$$

Furthermore, for notational convenience set

$$a := \min\{i : \beta_i \le D^{-\theta}/(2L)\} \ ,$$

and for $i = 0, 1, ...a$

$$\begin{aligned} \Delta_i &:= \beta_i(1 + \nu_i D^{\delta-1}) D \\ t_i &:= \alpha_i(1 - \mu_i D^{\delta-1}) D/L \\ d_i &:= \alpha_i \beta_i (1 + \nu_i D^{\delta-1}) D \end{aligned} \tag{19}$$

except

$$d_a := D^{-\theta} t_a \ . \tag{20}$$

As mentioned in the previous section, we use a two-part coloring procedure to prove that

$$\chi_l(G) \le \lfloor t_0 \rfloor \le D/L \ . \tag{21}$$

Notice that to prove Theorem 1.1 it is enough to prove this for each fixed η and large enough D.

Suppose we are given sets $S(v)$ of size t_0, $v \in V(G)$. (Of course, we should really write $\lfloor t_0 \rfloor$ here.) First we describe Algorithm 1 which colors many of the vertices of G and leaves an (induced) subgraph in which the color degrees are significantly smaller than the sizes of the sets of legal colors.

Algorithm 1

Initially we set $H_0 = G$, $T_0(v) = S(v)$, and $i = 0$. We run the following Steps a times.

(Step 1) Define the random coloring τ_i from $V(\hat{H}_i)$, $\hat{H}_i = \hat{H}_i(d_i)$, to the set of all colors according to

$$Pr(\tau_i(v) = \gamma) = \begin{cases} p_i := \theta/(\alpha_i D) & \text{if } \gamma \in T_i(v) \\ 1 - p_i |T_i(v)| & \text{if } \gamma = \Lambda \\ 0 & \text{otherwise} \end{cases}$$

independently of the other colors $\tau_i(w)$. Also set

$$
\begin{aligned}
X_i &= \{v \in V(\hat{H}_i) : \tau_i(v) \neq \Lambda, \quad v \sim w \ \text{ in } \hat{H}_i \quad \Rightarrow \tau_i(v) \neq \tau_i(w)\} \\
T_{i+1}(v) &= T_i(v) \setminus \{\tau_i(z) : z \in X_i, \quad z \sim v \ \text{ in } \hat{H}_i \ \} \ .
\end{aligned}
$$

and $H_{i+1} = H_i[V(H_i) \setminus X_i]$.

The properties we want are:

$$
\begin{aligned}
d_{i+1}(v) &\leq \Delta_{i+1} \\
t_{i+1}(v) &\geq t_{i+1} \\
d_{i+1}(v;\gamma) &\leq d_{i+1}
\end{aligned}
\tag{22}
$$

for all $v \in V(H_i)$ and $\gamma \in T_i(v)$ except

$$d_a(v;\gamma) \leq \alpha_a \beta_a (1 + \nu_a D^{\delta-1}) D.$$

Define an event $Q_i = \{$ (22) holds $\forall v \in V(H_i)$ and $\gamma \in T_i(v)$ $\}$. As mentioned, we need to show

$$Pr(Q_i) > 0 \ . \tag{23}$$

Supposing (23) is established, we choose τ_i so that (22) holds and proceed to Step 2.

(Step 2) Discard some colors, if necessary, from the sets $T_{i+1}(v)$ ($v \in V(H_{i+1})$) so that $|T_{i+1}(v)| = t_{i+1}$. (By this modification $d_{i+1}(v;\gamma)$ never increases.)

(Step 3) If $i < a-1$ then set $i = i+1$ (i.e. replace H_i by H_{i+1} etc.) and go to Step 1. Stop otherwise.

We will show below that values of μ_a, ν_a satisfy

$$\mu_a D^{\delta-1}, \nu_a D^{\delta-1} = o(1), \tag{24}$$

where $o(1)$ tends to zero as D tends to infinity. Thus by $\beta_a \leq D^{-\theta}/(2L)$ we have

$$\Delta_a \leq D^{-\theta}(1 + \nu_a D^{\delta-1}) D/(2L) \tag{25}$$

$$d_a \leq (2/3) D^{-\theta} t_a \qquad \text{(cf. (20))}. \tag{26}$$

We now continue with a modified algorithm better suited to the current values of our parameters. First, set $b := a + \theta^{-1} - 3$ and for $i = a, \cdots, b$

$$
\begin{aligned}
\Delta_{i+1} &= (1 + 1/\log D) D^{-\theta} \Delta_i \\
t_{i+1} &= (1 - 2D^{-\theta}) t_i \\
d_{i+1} &= D^{-\theta} t_{i+1} \ .
\end{aligned}
$$

We run the following steps $c := \theta^{-1} - 3$ times.

Algorithm 2

Initially, $i = a$.

(Step 1) Do step 1 of the first algorithm with $p_i = 1/t_i$. (Note $p_i t_i(v) = 1$ for $v \in V(H_i)$.) The properties we seek are:

$$\begin{array}{rcl} d_{i+1}(v) & \leq & \Delta_{i+1} \\ t_{i+1}(v) & \geq & t_{i+1} \\ d_{i+1}(v;\gamma) & \leq & d_{i+1} \end{array} \tag{27}$$

for all $v \in V(H_i)$ and $\gamma \in T_i(v)$. Note that the last inequality is trivial since by (26)

$$d_{i+1}(v;\gamma) \leq d_a(v;\gamma) \leq D^{-\theta} t_{i+1} \tag{28}$$

(because the number of stages is less than the fixed constant θ^{-1}). Define an event $Q_i = \{$ (27) holds $\forall v \in V(H_i)$ and $\gamma \in T_i(v)$ $\}$. Again, we need to show

$$Pr(Q_i) > 0 \ . \tag{29}$$

Supposing (29) is established, we choose τ_i so that (27) holds and proceed to Step 2.

(Step 2) As in Algorithm 1.

(Step 3) If $i < a + \theta^{-1} - 4$ then set $i = i + 1$ and go to step 1. Otherwise, stop.

Notice that once

$$d_b(v) < t_b(v) \qquad \text{for all } v \in V(H_b) \tag{30}$$

we may color the remaining vertices greedily. So to prove (21) (for large enough D), we just need to prove (23), (29), (24) and (30). We first dispose of the last two of these and then turn to the more difficult (23) and (29).

Lemma 3.1

$$\alpha_a \geq D^{-\eta e^{\theta}} \tag{31}$$

$$\max\{\mu_a, \nu_a\} = D^{o(1)}, \tag{32}$$

where o(1) goes to zero as D goes to infinity. In particular, we have (24).

Proof. Since

$$\alpha_i = \exp(-\theta\beta_{i-1} e^{-\theta\beta_{i-1}})\alpha_{i-1} \geq \exp(-\theta\beta_{i-1})\alpha_{i-1}$$

we have

$$\alpha_a \geq \exp(-\theta \sum_{i=0}^{a-1} \beta_i) \ .$$

On the other hand, since

$$\beta_i = (1 - (\theta/L)e^{-\theta\beta_{i-1}})\beta_{i-1} \leq (1 - (\theta/L)e^{-\theta})\beta_{i-1} \leq (1 - (\theta/L)e^{-\theta})^i \tag{33}$$

we have

$$\sum_{i=1}^{a-1} \beta_i \leq \sum_{i=0}^{\infty} (1 - \theta e^{-\theta}/L)^i = e^{\theta} L/\theta \ ,$$

which implies

$$\alpha_a \geq \exp(-\theta \sum_{i=0}^{a-1} \beta_i) \geq \exp(-\theta e^{\theta} L/\theta) = D^{-\eta e^{\theta}} \ .$$

To prove (32), let us define a_1 to be the maximum i such that $\beta_i > L^{-2}$. Then by (33), we have

$$a_1 \leq 2\theta^{-1} e^{\theta} L \log L \ .$$

Note that, trivially,

$$\begin{pmatrix} \mu_{a_1} \\ \nu_{a_1} \end{pmatrix} \leq \begin{pmatrix} 1 & 1 \\ 1/L & 1+1/L \end{pmatrix}^{a_1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(meaning, as usual, that μ_{a_1} (resp. ν_{a_1}) is at most the first (resp. second) component of the right hand side). Similarly we have

$$a \leq e^{\theta} L(\log D + \theta^{-1} \log(2L)) + 1 \ ,$$

and

$$\begin{pmatrix} \mu_{a} \\ \nu_{a} \end{pmatrix} \leq \begin{pmatrix} 1 & 1/L^2 \\ 1/L & 1+1/L^3 \end{pmatrix}^{a-a_1} \begin{pmatrix} \mu_{a_1} \\ \nu_{a_1} \end{pmatrix}$$

since $\beta_i \leq L^{-2}$ for $i > a_1$. Furthermore, the matrices

$$\begin{pmatrix} 1 & 1 \\ 1/L & 1+1/L \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1/L^2 \\ 1/L & 1+1/L^3 \end{pmatrix}$$

have diagonal Jordan forms with eigenvalues approximately $1 \pm 1/\sqrt{L}$, $1 \pm 1/(L\sqrt{L})$ respectively, and these with the above bounds on a, a_1 imply

$$\max\{\mu_{a_1}, \nu_{a_1}\} \leq 2\sqrt{L}(1 + 2/\sqrt{L})^{a_1} = D^{o(1)}$$

and

$$\max\{\mu_a, \nu_a\} \leq 2L(1 + 2/(L\sqrt{L}))^a D^{o(1)} = D^{o(1)}.$$

□

Proof of Theorem 1.1 Suppose now that we have run Algorithm 2 c times. Then by (24) and (25)

$$\Delta_b = (1 + 1/\log D)^c D^{-c\theta} \Delta_a \leq \exp(c/\log D) D^{-(c+1)\theta} D/L \leq D^{2\theta} \ .$$

On the other hand, by (24), (31) and (18) we have

$$t_b = (1 - 2D^{-\theta})^c t_a \geq (1 - 2D^{-\theta})^c \alpha_a D/(2L) \geq \frac{1}{3} D^{-\eta e^{\theta}} D/L > D^{2\theta} \ . \tag{34}$$

Thus we are done.

□

We have already mentioned in Section 2 the methods to be used in proving (23) and (29). The following lemmas are precise statements. We will prove them in last two sections.

From now on, we fix $i \in [b] := \{1, ..., b\}$ and for simplicity, we do not write the subscript i (i.e. $H = H_i, d(v) = d_i(v), \alpha = \alpha_i$ etc.). Also, we write H', α' etc. for H_{i+1}, α_{i+1} etc. (respectively).

Lemma 3.2 *For $v \in V(H)$ and $\gamma \in T(v)$,*

$$\begin{aligned} E[d'(v)] &= (1-pt(1-p)^d)d(v) \leq (1-pt(1-p)^d)\Delta, \\ E[t'(v)] &= (1-p(1-p)^d)^d t + O(1), \\ E[d'(v;\gamma)] &\leq (1-p(1-p)^d)^d(1-pt(1-p)^d)d + O(1). \end{aligned}$$

The proof of Lemma 3.2 is quite straight forward. Our main lemma is:

Lemma 3.3 (Main Lemma)

$$Pr(d'(v) - E[d'(v)] \geq \Delta^{1/2}\log\Delta) \leq \exp(-(\log\Delta)^2/4) \tag{35}$$

$$Pr(t'(v) - E[t'(v)] \leq -t^{1/2}\log t) \leq \exp(-(\log t)^2/2) \tag{36}$$

$$Pr(d'(v;\gamma) - E[d'(v;\gamma)] \geq d^{1/2}(\log d)^2) \leq 3D^2\exp(-\frac{1}{2}\log d\log\log d) \tag{37}$$

Our proof will give bounds on the probabilities in (35), (36) of other direction — e.g.

$$Pr(d'(v) - E[d'(v)] \leq -\Delta^{1/2}\log\Delta) \leq \exp(-(\log\Delta)^2/4)$$

— but we restrict the formal statement to the values we will actually use.

Once the above lemmas are proved, it is easy to prove (23) and (29). Before doing so, we summarize some inequalities already established. Here we write $x \ll y$ if there is a constant $\varepsilon > 0$ depending only on θ, δ and η such that $xD^\varepsilon \leq y$.

$$\delta - 1 > \frac{1}{2}(\eta e^\theta + 2\theta - 1) \qquad \text{by (18)} \tag{38}$$

$$t_j > D^{1-\eta e^\theta - o(1)} \gg D^{2\theta} \qquad \forall\ j \in [b] \qquad \text{by (34) and (18)} \tag{39}$$

$$\beta_j > D^{-\theta-o(1)} \qquad \forall\ j \in [a] \qquad \text{by the definition of } a \tag{40}$$

$$\frac{1}{\alpha_j D} < D^{\eta e^\theta - 1} \ll D^{\delta - 2\theta - 1} \qquad \forall\ j \in [a] \qquad \text{by (31) and (18)} \tag{41}$$

Moreover, by (40) and (39)

$$d_j \geq D^{-\theta-o(1)}t_j > D^{1-\eta e^\theta - \theta - o(1)} \gg D^\theta \qquad \forall\ j \in [b] \tag{42}$$

and by the definitions of Δ_{b-1}, Δ_a and $b = a + \theta^{-1} - 3$ for all $j = 1, 2, ..., b-1$

$$\Delta \geq \Delta_{b-1} \geq (1 + 1/\log D)^{b-1-a}D^{-\theta(b-1-a)}\Delta_a \geq D^{-\theta(\theta^{-1}-4)}D^{1-\theta-o(1)} \gg D^{2\theta}\ . \tag{43}$$

Proofs of (23) and (29). For each $v \in V(H)$ consider the event Q_v that we do not have the required properties for v, that is,

$$\begin{aligned} Q_v = &\{d'(v) > E[d'(v)] + \Delta^{1/2}\log\Delta\} \bigcup \{t'(v) < E[t'(v)] - t^{1/2}\log t\} \\ &\bigcup \{d'(v;\gamma) > E[d'(v;\gamma)] + d^{1/2}(\log d)^2 \ \text{ for some } \gamma \in T(v)\ \} \end{aligned}$$

Since (by (43), (39) and (42))

$$\min\{\Delta, t, d\} > D^\theta$$

Lemma 3.3 implies , e.g., (since D is large)

$$Pr(Q_v) \leq 3tD^2 \exp(-(\theta/3)\log D \log\log D) \leq D^3 \exp(-(\theta/3)\log D \log\log D)\ .$$

Furthermore, note that the event Q_v is independent of all events $\{Q_w\}$ for which the distance between v and w is more than 6 (since for all v, $d'(v), t'(v)$ and all $d'(v;\gamma)$'s are determined by the values of τ on vertices within distance 3 of v). Thus the Lovász Local Lemma [9], (see also [27]) together with the inequalities

$$4D^6 Pr(Q_v) \leq D^6 D^3 \exp(-(\theta/3)\log D \log\log D) < 1 \qquad \forall v \in V(H)$$

guarantees

$$Pr(\bigcap_{v\in V(H)} \bar{Q}_v) > 0\ .$$

Therefore, (using the values in Lemma 3.2) we can find a coloring τ on $V(H)$ such that for every v and $\gamma \in T'(v)$

$$\begin{aligned} d'(v) &\leq (1-pt(1-p)^d)\Delta + \Delta^{1/2}\log\Delta \\ t'(v) &\geq (1-p(1-p)^d)^d t - t^{1/2}\log t - O(1) \\ d'(v;\gamma) &\leq (1-p(1-p)^d)^d(1-pt(1-p)^d)d + d^{1/2}(\log d)^2 + O(1)\ . \end{aligned} \tag{44}$$

Thus to show (22), (27) we just have to show that the inequalities in (44) imply those in (22) if we are in Algorithm 1 and those in (27) if we are in Algorithm 2.

We analyze the two cases separately. In Algorithm 1 we have two kinds of error terms other than the trivial errors $O(1)$. The first kind is from accumulation of errors in the expectations. (Note that t and d already contain such error terms.) The other kind is, of course, from concentration errors ($\Delta^{1/2}\log\Delta$ etc.). As will appear below, we have chosen the parameters — see (18) — so that the errors of the first type dominate those of the second. Though not hard, the estimates are somewhat complicated and tedious. We will frequently use (41)-(43).

Suppose first that we are in Algorithm 1. Let us recall

$$pd = \theta\beta(1+\nu D^{\delta-1}) \leq 0.11, \quad pt = \theta(1-\mu D^{\delta-1})/L \leq 0.1 \tag{45}$$

We claim

$$(1-pt(1-p)^d) - (1-(\theta/L)e^{-\theta\beta}) \leq (\theta/L)(\mu+\theta\beta\nu)D^{\delta-1} + \theta\beta p \tag{46}$$

$$0 \leq \exp(-\theta\beta e^{-\theta\beta}) - (1-p(1-p)^d)^d \leq \theta\beta\nu D^{\delta-1} \tag{47}$$

For (46), since $1-p \geq e^{-p-p^2}$ we have

$$\begin{aligned} 1-pt(1-p)^d &\leq 1-pte^{-pd}e^{-p^2d} \\ &\leq 1-pte^{-pd}(1-p^2d) \qquad \text{by } e^{-p^2d} \geq 1-p^2d \\ &\leq 1-pte^{-pd}+\theta\beta p \qquad \text{by (45) and (24).} \end{aligned}$$

Now set

$$f(x,y) = 1-(\theta/L)(1-x)e^{-\theta\beta(1+y)}\ .$$

If $0 < x, y < 0.1$ then by Taylor's theorem

$$f(x,y) - f(0,0) \leq f_x(0,0)x + f_y(0,0)y = (\theta/L)e^{-\theta\beta}x + (\theta^2\beta/L)e^{-\theta\beta}y \leq (\theta/L)(x + \theta\beta y)$$

since all second order derivatives are non positive (for $0 < x, y < 0.1$). Setting $x = \mu D^{\delta-1}$ and $y = \nu D^{\delta-1}$, we have (46).

For the upper bound of (47), consider

$$\begin{aligned}
(1 - p(1-p)^d)^d &\geq (1 - pe^{-pd})^d \\
&\geq \exp(-pde^{-pd} - p^2de^{-2pd}) \\
&\geq (1 - p^2de^{-2pd})\exp(-pde^{-pd}) \\
&\geq \exp(-pde^{-pd}) - p \ . \qquad (48)
\end{aligned}$$

Set $h(y) = -\theta\beta(1+y)e^{-\theta\beta(1+y)}$. Then by the similar argument we have

$$h(y) - h(0) \geq h'(0)y = (-\theta\beta e^{-\theta\beta} + \theta^2\beta^2e^{-\theta\beta})y \geq -(\theta\beta - \theta^2\beta^2)y \ , \qquad (49)$$

for $0 < y < 0.1$. Moreover, we have by (40) and (41)

$$p \ll \theta^2\beta^2D^{\delta-1} \ , \qquad (50)$$

(note $p = \theta/(\alpha D)$ here). Again setting $y = \nu D^{\delta-1}$ we finally have

$$\begin{aligned}
(1 - p(1-p)^d)^d &\geq \exp(h(y)) - p && \text{by (48)} \\
&\geq \exp(h(0) - (\theta\beta - \theta^2\beta^2)y) - p && \text{by (49)} \\
&\geq \exp(-\theta\beta e^{-\theta\beta})(1 - (\theta\beta - \theta^2\beta^2)y) - p \\
&\geq \exp(-\theta\beta e^{-\theta\beta}) - (\theta\beta - \theta^2\beta^2)\nu D^{\delta-1} - p \\
&\geq \exp(-\theta\beta e^{-\theta\beta}) - \theta\beta\nu D^{\delta-1} && \text{by (50),}
\end{aligned}$$

which is exactly what we want for the upper bound.

Note that the upper bound is quite tight. Thus we may easily modify the estimation to show the lower bound. We leave this to the reader.

Now we claim the following to control the second kind of errors.

$$\begin{aligned}
\Delta^{1/2}\log\Delta + p\Delta &\leq (\theta/L)(\mu + \theta\beta\nu)D^{\delta-1}\Delta && (51) \\
t^{1/2}\log t + O(1) &\leq \theta\beta\nu D^{\delta-1}t && (52) \\
d^{1/2}(\log d)^2 + pd + O(1) &\leq (\theta/L)(\mu + \theta\beta\nu)D^{\delta-1}d \ . && (53)
\end{aligned}$$

We already saw p is small enough in (50). Thus it is enough for us to show

$$\max\{\Delta^{-1/2},\ \beta^{-1}t^{-1/2},\ d^{-1/2}\} \ll D^{\delta-1}.$$

(We can not disregard β here because it can be as small as $D^{-\theta}/(2L)$.) For (51), it is enough for us to note that by (40) and (38)

$$\Delta^{-1/2} \leq (\beta D)^{-1/2} \leq D^{(\theta-1)/2+o(1)} \ll D^{\delta-1} \ .$$

Similarly, we have by (39), (40) and (38)

$$\beta^{-1}t^{-1/2} \leq D^{(\eta e^{\theta}+2\theta-1)/2+o(1)} \ll D^{\delta-1} \ .$$

Finally, by (42) and (38)

$$d^{-1/2} \le D^{(\eta e^{\theta}+\theta-1)/2+o(1)} \ll D^{\delta-1} \ ,$$

which completes the proof of our claims.

Using the above claims and the fact that β/β' is almost 1, we have

$$\begin{aligned} d'(v) &\le (1-(\theta/L)e^{-\theta\beta})\Delta + 2(\theta/L)(\mu+\theta\beta\nu)D^{\delta-1}\Delta \\ &\le \beta'(1+\nu D^{\delta-1})(1+2(\beta/\beta')(\theta/L)(\mu+\theta\beta\nu)D^{\delta-1})D \\ &\le \beta'(1+(\nu+(3\theta/L)(\mu+\theta\beta\nu))D^{\delta-1})D \\ &\le \beta'(1+(\nu+(\mu+\beta\nu)/L)D^{\delta-1})D \\ &= \beta'(1+\nu' D^{\delta-1})D \ . \end{aligned}$$

Here we do not have to be so careful about the product of the error terms since we already know $\mu, \nu = D^{o(1)}$. Similarly,

$$\begin{aligned} t'(v) &\ge \alpha'(1-(\mu+\beta\nu)D^{\delta-1})D/L = \alpha'(1-\mu' D^{\delta-1})D/L \\ d'(v;\gamma) &\le \alpha'\beta'(1+(\nu+(\mu+\beta\nu)/L)D^{\delta-1})D = \alpha'\beta'(1+\nu' D^{\delta-1})D \ . \end{aligned}$$

Suppose now we are in Algorithm 2. Then since $(1-p)^d \ge 1-pd = 1-D^{-\theta}$ we have

$$1-pt(1-p)^d = 1-(1-p)^d \le D^{-\theta}$$

and

$$(1-p(1-p)^d)^d \ge (1-p)^d \ge 1-D^{-\theta} \ .$$

Since by (43)

$$\Delta^{-1/2} < D^{-(3\theta-o(1))/2} = D^{-\theta}D^{-\theta/2+o(1)} \ll D^{-\theta}/(\log\Delta \log D)$$

we have

$$\begin{aligned} d'(v) &\le D^{-\theta}\Delta + \Delta^{1/2}\log\Delta \\ &\le (1+D^{\theta}\Delta^{-1/2}\log\Delta)D^{-\theta}\Delta \\ &\le (1+1/\log D)D^{-\theta}\Delta = \Delta' (= \Delta_{i+1}) \ . \end{aligned}$$

Similarly, by (39), we have

$$\begin{aligned} t'(v) &\ge (1-D^{-\theta})t - t^{1/2}\log t \\ &= (1-D^{-\theta}-t^{-1/2}\log t)t \\ &\ge (1-2D^{-\theta})t = t' (= t_{i+1}) \ . \end{aligned}$$

□

4 Expectations

In this section we prove Lemma 3.2. Let us recall the lemma.

Lemma 3.2 (restatement) *For $v \in V(H)$ and $\gamma \in T(v)$,*

$$\begin{aligned} E[d'(v)] &= (1-pt(1-p)^d)d(v), & (54)\\ E[t'(v)] &= (1-p(1-p)^d)^d t + O(1), & (55)\\ E[d'(v;\gamma)] &\le (1-p(1-p)^d)^d(1-pt(1-p)^d)d + O(1). & (56) \end{aligned}$$

Proof. (a) For degrees,

$$E[d'(v)] = \sum_{w \in N(v)} (1 - Pr(w \in X)) \ .$$

But

$$\begin{aligned} Pr(w \in X) &= \sum_{\gamma \in T(w)} Pr(\tau(w) = \gamma, \tau(z) \neq \gamma \ \ \forall z \in \hat{N}(w;\gamma)) \\ &= tp(1-p)^d \ . \end{aligned}$$

Therefore, we have (54).

(b) For the number of legal colors,

$$E[t'(v)] = \sum_{\gamma \in T(v)} Pr(\gamma \in T'(v)).$$

On the other hand, for fixed v and $\gamma \in T(v)$, we have $\gamma \in T'(v)$ if and only if there is no $w \in \hat{N}(v)$ for which the event

$$A_w := \{\tau(w) = \gamma, \ \ \tau(z) \neq \gamma \quad \forall z \sim w\}$$

happens. If we condition on $\tau(v) \neq \gamma$, then, since $g(G) \geq 5$, the event A_w $(w \in \hat{N}(v;\gamma))$ are independent, and we have

$$\begin{aligned} Pr(\gamma \in T'(v) | \tau(v) \neq \gamma) &= \prod_{w \in \hat{N}(v;\gamma)} Pr(\bar{A}_w)|\tau(v) \neq \gamma) \\ &= (1-p(1-p)^{d-1})^d & (57) \end{aligned}$$

Thus since $Pr(\tau(v) = \gamma) = p$,

$$\begin{aligned} Pr(\gamma \in T'(v)) &= Pr(\tau(v) = \gamma) Pr(\gamma \in T'(v) | \tau(v) = \gamma) \\ &\quad + \ Pr(\tau(v) \neq \gamma)(1-p(1-p)^{d-1})^d \\ &= (1-p(1-p)^{d-1})^d + O(p) \\ &= (1-p(1-p)^d)^d + O(p) \ , & (58) \end{aligned}$$

which (since $pt \leq 1$) gives (55).

(c) For color degree,

$$
\begin{aligned}
E[d'(v;\gamma)] &= \sum_{w\in N(v;\gamma)} Pr(w\notin X, \gamma\in T'(w)) \\
&= \sum_{w\in N(v;\gamma)} \left(Pr(\gamma\in T'(w)) - Pr(\gamma\in T'(w), w\in X)\right) .
\end{aligned}
$$

We claim

$$
Pr(\gamma\in T'(w), w\in X) \geq pt(1-p)^d(1-p(1-p)^d)^d + O(p). \tag{59}
$$

Since we know $Pr(\gamma\in T'(w)) = (1-p(1-p)^d)^d + O(p)$ and $pd \leq 0.11$ (see (45)), (56) follows if we prove (59).

To do so, we need only consider the case $\tau(w)\neq\gamma$, since the other case has the probability p. First note that since $w\in X$ implies $\tau(w)\neq\Lambda$ we have

$$
\begin{aligned}
& Pr(\gamma\in T'(w), w\in X) \\
= & \sum_{\gamma'\in T(w)\setminus\{\gamma\}} Pr(\tau(w)=\gamma')Pr(\gamma\in T'(w), w\in X|\tau(w)=\gamma') + O(p) \\
= & \; p\sum_{\gamma'\in T(w)\setminus\{\gamma\}} Pr(w\in X|\tau(w)=\gamma')P(\gamma\in T'(w)|w\in X, \tau(w)=\gamma') + O(p) \\
= & \; p(1-p)^d \sum_{\gamma'\in T(w)\setminus\{\gamma\}} P(\gamma\in T'(w)|w\in X, \tau(w)=\gamma') + O(p).
\end{aligned}
$$

Thus it is enough to show that

$$
Pr(\gamma\in T'(w)|w\in X, \tau(w)=\gamma') \geq (1-(1-p)^d)^d + O(p) \tag{60}
$$

Without the extra condition "$w\in X$", we may easily prove (60) as in (57). On the other hand, the extra condition is nothing but $\tau(z)\neq\gamma'$ for all $z\in\hat{N}(w;\gamma')$ and does not affect the mutual independence of events "$\tau(z)=\gamma$". The only change required here is replacement of $p = Pr(\tau(z)=\gamma)$ by

$$
p(z) := Pr(\tau(z)=\gamma|\tau(z)\neq\gamma') = \begin{cases} p/(1-p) & \text{if } z\in\hat{N}(w;\gamma') \\ p & \text{if } z\notin\hat{N}(w;\gamma') \end{cases}
$$

Then as in (57)

$$
Pr(\gamma\in T'(w)|w\in X, \tau(w)=\gamma') = \prod_{z\in\hat{N}(w;\gamma)} (1-p(z)(1-p)^{d-1}).
$$

Since $p(z) = p + O(p^2)$ we have (60).

□

5 Martingales

In this section, we develop Azuma-Hoeffding-type martingale inequalities which form the basis for our proofs of high concentrations of the random variables $d'(v)$, $t'(v)$, and $d'(v;\gamma)$ near their expectations. For general probability theory and martingales, see e.g. [6], [19] and [27].

Here we define finite martingales briefly:

Let Y be a random variable and $\mathcal{B}_0, \mathcal{B}_1, ..., \mathcal{B}_n$ a non-decreasing sequence of σ-fields on a probability space, where $\mathcal{B}_0$ is the trivial σ-field (i.e. $\mathcal{B}_0 = \{\emptyset, \text{ Whole Set}\}$). Suppose Y is $\mathcal{B}_n$-measurable, that is,

$$E[Y|\mathcal{B}_n] = Y \ .$$

Then the martingale generated by Y with respect to $\{\mathcal{B}_i\}_{i=0}^n$ is the sequence

$$\{Y_i := E[Y|\mathcal{B}_i]\}_{i=0}^n \ .$$

Note that $Y_0 = E[Y]$, $Y_n = Y$ and

$$E[Y_i|\mathcal{B}_{i-1}] = Y_{i-1} \qquad \forall \ i = 1, 2, ..., n \tag{61}$$

(actually, (61) is the general definition of martingales). Also, we define martingale difference sequence

$$Z_k := Y_k - Y_{k-1} \qquad \text{for } k = 1, ..., n,$$

and set $Z := \sum_{k=1}^n Z_k = Y - E[Y]$.

From now on when we refer martingales we always assume that $\{\mathcal{B}_i\}$, Z_i's etc. are taken for granted. We first introduce the following lemma from [15].

Lemma 5.1 *Let $\{Y_i\}_{i=0}^n$ be a martingale. Suppose that*

$$E[e^{\omega Z_k}|\mathcal{B}_{k-1}] \leq C_k \qquad \forall k = 1, ..., n \tag{62}$$

for some positive ω and $C_1, ..., C_n$. Then

(a) $E[e^{\omega Z}] \leq \prod_{k=1}^{n} C_k$ and

(b) $Pr(Y - E[Y] \geq \lambda) \leq e^{-\omega\lambda} \prod_{k=1}^{n} C_k$

for all real number λ.

Proof. First, note that (a) implies (b) since $Z = Y - E[Y]$ and

$$Pr(Z \geq \lambda) = Pr(e^{\omega Z} \geq e^{\omega\lambda}) \leq e^{-\lambda\omega} E[e^{\omega Z}]$$

by Markov's inequality. For (a), we show

$$E[e^{\omega(Z_1 + \cdots + Z_k)}] \leq \prod_{l=1}^{k} C_l$$

for all $k = 1, ..., n$ by induction. If $k = 1$,

$$E[e^{\omega Z_1}] = E[E[e^{\omega Z_1}|\mathcal{B}_0]] \leq C_1$$

For $k > 1$ using the induction hypothesis,

$$
\begin{aligned}
E[e^{\omega(Z_1+\cdots+Z_k)}] &= E[E[e^{\omega(Z_1+\cdots+Z_k)}|\mathcal{B}_{k-1}]] \\
&= E[e^{\omega(Z_1+\cdots+Z_{k-1})}E[e^{\omega Z_k}|\mathcal{B}_{k-1}]] \\
&\leq E[e^{\omega(Z_1+\cdots+Z_{k-1})}C_k] \\
&\leq \prod_{l=1}^{k} C_l \ .
\end{aligned}
$$

□

As mentioned in Section 2, we need something a little more general than Lemma 5.1 which allows the bounds (62) to fail occasionally.

Lemma 5.2 *If there are $A_{k-1} \in \mathcal{B}_{k-1}$ such that*

$$
E[e^{\omega Z_k}|\mathcal{B}_{k-1}]1_{\bar{A}_{k-1}} \leq C_k \qquad \forall k = 1, 2, \cdots, n \tag{63}
$$

with $C_k \geq 1$ for all k, then

$$
Pr(Y - E[Y] \geq \lambda) \leq e^{-\lambda\omega} \prod_{k=1}^{n} C_k + Pr(\bigcup_{k=0}^{n-1} A_k).
$$

When the $Pr(A_k)$ is small enough we may roughly speak of C_k as an "essential upper bound" on $E[e^{\omega Z_k}|\mathcal{B}_{k-1}]$.

Proof. First we define a stopping time

$$
\sigma(x) = \begin{cases} \min\{k|x \in A_k\} & \text{if } x \in \bigcup_{k=0}^{n-1} A_k \\ n & \text{otherwise.} \end{cases}
$$

Then by the Optional Sampling Theorem (see e.g. [6]), the sequence $\{Y_{k\wedge\sigma}\}_{k=0}^{n}$ is a martingale, where, as usual, $k \wedge \sigma := \min\{k, \sigma\}$. In particular, we have for $Y' = Y_{n\wedge\sigma}$

$$
E[Y'|\mathcal{B}_k] = Y_{k\wedge\sigma} \qquad \forall k = 0, \cdots, n. \tag{64}
$$

In particular $E[Y'] = E[Y]$.

Furthermore, for $Z'_k := E[Y'|\mathcal{B}_k] - E[Y'|\mathcal{B}_{k-1}] = Y_{k\wedge\sigma} - Y_{(k-1)\wedge\sigma}$, we know

$$
Z'_k = \begin{cases} 0 & \text{if } \sigma \leq k-1 \\ Y_k - Y_{k-1} = Z_k & \text{if } \sigma \geq k \ . \end{cases}
$$

Thus we have

$$
e^{\omega Z'_k} = e^{\omega Z'_k}1_{\{\sigma\leq k-1\}} + e^{\omega Z'_k}1_{\{\sigma\geq k\}} = 1_{\{\sigma\leq k-1\}} + e^{\omega Z_k}1_{\{\sigma\geq k\}} \ .
$$

Since $\{\sigma \leq k-1\}, \{\sigma \geq k\} \in \mathcal{B}_{k-1}$, $\{\sigma \geq k\} \subseteq \bar{A}_{k-1}$ and $C_k \geq 1$,, we have

$$
E[e^{\omega Z'_k}|\mathcal{B}_{k-1}] = 1_{\{\sigma\leq k-1\}} + E[e^{\omega Z_k}|\mathcal{B}_{k-1}]1_{\{\sigma\geq k\}} \leq C_k \ .
$$

Therefore, by Lemma 5.1 we have

$$Pr(Y' - E[Y'] \geq \lambda) \leq e^{-\omega\lambda} \prod_{k=1}^{n} C_k \ ,$$

which implies the result since $E[Y'] = E[Y]$ and $Y' = Y_{n\wedge\sigma} = Y_n = Y$ except on $\{\sigma < n\} = \cup_{k=0}^{n-1} A_k$.

□

Of course if we know, say, $|Z_k| \leq c_k$ on $\bar{A}_{k-1}$ then we can take $C_k = e^{\omega c_k}$ or $e^{\omega^2 c_k^2/2}$ (by $E[Z_k|\mathcal{B}_{k-1}] = 0$) in (5.2). But if (on $\bar{A}_{k-1}$) Z_k is only rarely near its maximum, then we should be able to do better. A typical example for us (and also, e.g., in [14], [15]) is that Z_k takes only two values, say

$$Z_k = \begin{cases} c_k & \text{on } B_k \\ c'_k & \text{on } \bar{B}_k \end{cases}$$

for some low probability set $B_k \in \mathcal{B}_k$. In this case, if B_k is independent of $\mathcal{B}_{k-1}$ then $E[Z_k|\mathcal{B}_{k-1}] = 0$ implies that c'_k is small (no more than $c_k Pr(B_k)$ in absolute value). This situation is described in the next lemma.

Lemma 5.3 *Suppose that there is a set $I \subseteq [n]$, such that*

$$\begin{aligned} |Z_k| 1_{\bar{A}_{k-1}} &\leq c_k 1_{B_k} + c_k Pr(B_k), \qquad &&\forall k \in I \qquad (65) \\ Z_k 1_{\bar{A}_{k-1}} &\leq c_k &&\forall k \in J := [n] \setminus I \qquad (66) \end{aligned}$$

for some constants c_k, and some sets $A_{k-1} \in \mathcal{B}_{k-1}$ and B_k independent of $\mathcal{B}_{k-1}$. Then we have for all positive ω with $\omega \max_{k\in I}\{c_k\} \leq \frac{1}{6}$

$$Pr(Y - E[Y] \geq \lambda) \leq Pr(\bigcup_{k=0}^{n-1} A_k) + \exp(-\omega(\lambda - \sum_{k\in J} c_k) + 3\omega^2 \sum_{k\in I} c_k^2 Pr(B_k)).$$

Proof. By Lemma 5.2 it is enough to show that

$$E[e^{\omega Z_k}|\mathcal{B}_{k-1}] 1_{\bar{A}_{k-1}} \leq e^{3\omega^2 c_k^2 Pr(B_k)} \qquad \text{for } k \in I \qquad (67)$$

and

$$E[e^{\omega Z_k}|\mathcal{B}_{k-1}] 1_{\bar{A}_{k-1}} \leq e^{\omega c_k} \qquad \text{for } k \in J \ . \qquad (68)$$

Note that (68) is immediate from (66), we really only need to prove (67).

For (67), set $V = Z_k 1_{\bar{A}_{k-1}}$, $\mathcal{B} = \mathcal{B}_{k-1}$, $c_k = c$, $B_k = B$ and $b = Pr(B)$ (for fixed $k \in I$). Then

$$E[e^{\omega Z_k}|\mathcal{B}_{k-1}] 1_{\bar{A}_{k-1}} = E[1_{\bar{A}_{k-1}} e^{\omega Z_k}|\mathcal{B}_{k-1}] \leq E[e^{\omega V}|\mathcal{B}] \ .$$

Also we know

$$E[V|\mathcal{B}] = E[Z_k 1_{\bar{A}_{k-1}}|\mathcal{B}_{k-1}] = E[Z_k|\mathcal{B}_{k-1}] 1_{\bar{A}_{k-1}} = 0 \ . \qquad (69)$$

Thus by (69) and (65) we have

$$\begin{aligned}
E[e^{\omega V}|\mathcal{B}] &= \sum_{j=0}^{\infty} E[\omega^j V^j|\mathcal{B}]/j! \\
&\leq 1 + \omega E[V|\mathcal{B}] + \frac{1}{2}\sum_{j=2}^{\infty} \omega^j E[|V^j|\ |\mathcal{B}] \\
&\leq 1 + \frac{1}{2}\sum_{j=2}^{\infty} \omega^j c^j E[(1_B + b)^j|\mathcal{B}] \ .
\end{aligned}$$

On the other hand, since

$$E[(1_B)^{j-l}|\mathcal{B}] = \begin{cases} b & \text{if } l \neq j \\ 1 & \text{if } l = j \end{cases} \ .$$

(since B_k is independent of $\mathcal{B}_{k-1}$), we have

$$\begin{aligned}
E[(1_B + b)^j|\mathcal{B}] &= \sum_{l=0}^{j} \begin{pmatrix} j \\ l \end{pmatrix} E[b^l (1_B)^{j-l}|\mathcal{B}] \\
&= \sum_{l=0}^{j} \begin{pmatrix} j \\ l \end{pmatrix} b^{l+1} + (b^j - b^{j+1}) \\
&= b(1+b)^j + b^j(1-b) \ .
\end{aligned}$$

Furthermore, since $\omega c \leq 1/6$ and $b \leq 1$

$$\begin{aligned}
\sum_{j=2}^{\infty} \omega^j c^j E[(1_B + b)^j|\mathcal{B}] &= \sum_{j=2}^{\infty} \omega^j c^j (b(1+b)^j + b^j(1-b)) \\
&= b\sum_{j=2}^{\infty} \omega^j c^j (1+b)^j + (1-b)\sum_{j=2}^{\infty} \omega^j c^j b^j \\
&= \frac{b\omega^2 c^2 (1+b)^2}{1 - \omega c(1+b)} + \frac{(1-b)\omega^2 c^2 b^2}{1 - \omega c b}
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{j=2}^{\infty} \omega^j c^j E[(1_B + b)^j|\mathcal{B}] &\leq b\omega^2 c^2 \frac{(1+b)^2 + b(1-b)}{1 - \omega c(1+b)} \\
&= b\omega^2 c^2 \frac{1+3b}{1 - \omega c(1+b)} \\
&\leq 6b\omega^2 c^2 \ .
\end{aligned}$$

Therefore

$$E[e^{\omega V}|\mathcal{B}] \leq 1 + 3b\omega^2 c^2 \leq \exp(3b\omega^2 c^2) \ .$$

$\square$

6 More Lemmas

In the previous section, we developed martingale inequalities which are useful when we know nice (essential) upper bounds on $Z_k = E[Y|\mathcal{B}_k] - E[Y|\mathcal{B}_{k-1}]$. It is relatively easy to find nice upper bounds if the random variable Y has the typical form

$$Y = Y(\tau_1, \tau_2, ..., \tau_n)$$

where $\tau_1, \tau_2, ..., \tau_n$ are mutually independent random variables such that for every k the σ-field generated by $\tau_1, \tau_2, ..., \tau_k$ is exactly $\mathcal{B}_k$. As all examples we require will look like this, we restrict our attention to such Y's from now on.

For

$$\tau := (\tau_1, \tau_2, ..., \tau_n) \qquad \text{and} \qquad \tau' := (\tau'_1, \tau'_2, ..., \tau'_n),$$

define equivalence relations $\equiv_k$ by

$$\tau \equiv_k \tau' \quad \text{if and only if} \quad \tau_j = \tau'_j \quad \text{for all} \;\; j \in [n] \setminus \{k\}.$$

Lemma 6.1 *With the above notation, suppose for some $k \in [n]$ there is a random variable W such that*

$$|Y(\tau) - Y(\tau')| \leq W(\tau) + W(\tau') \qquad \textit{whenever} \quad \tau \equiv_k \tau'. \tag{70}$$

Then

$$|Z_k| \leq E[W|\mathcal{B}_k] + E[W|\mathcal{B}_{k-1}] \; .$$

(Recall $Z_k = E[Y|\mathcal{B}_k] - E[Y|\mathcal{B}_{k-1}]$.)

Proof. First note that for fixed $\kappa = (\kappa_1, ..., \kappa_n)$

$$E[Y|\mathcal{B}_{k-1}](\kappa) = \sum_{\gamma_k, \cdots, \gamma_n} Y(\kappa_1, \cdots, \kappa_{k-1}, \gamma_k, \cdots, \gamma_n) Pr(\tau_k = \gamma_k, \cdots, \tau_n = \gamma_n),$$

and

$$\begin{aligned} E[Y|\mathcal{B}_k](\kappa) &= \sum_{\gamma_{k+1}, \cdots, \gamma_n} Y(\kappa_1, \cdots, \kappa_k, \gamma_{k+1}, \cdots, \gamma_n) Pr(\tau_{k+1} = \gamma_{k+1}, \cdots, \tau_n = \gamma_n) \\ &= \sum_{\gamma_k, \cdots, \gamma_n} Y(\kappa_1, \cdots, \kappa_k, \gamma_{k+1}, \cdots, \gamma_n) Pr(\tau_k = \gamma_k, \cdots, \tau_n = \gamma_n) \end{aligned}$$

since $\sum_{\gamma_k} Pr(\tau_k = \gamma_k, \cdots, \tau_n = \gamma_n) = Pr(\tau_{k+1} = \gamma_{k+1}, \cdots, \tau_n = \gamma_n)$. Thus by (70) we have

$$\begin{aligned} |Z_k(\kappa)| &= |(E[Y|\mathcal{B}_k] - E[Y|\mathcal{B}_{k-1}])(\kappa)| \\ &\leq \sum_{\gamma_k, \cdots, \gamma_n} |Y(\kappa_1, \cdots, \kappa_k, \gamma_{k+1}, \cdots, \gamma_n) - Y(\kappa_1, \cdots, \kappa_{k-1}, \gamma_k, \cdots, \gamma_n)| \\ &\qquad \times Pr(\tau_k = \gamma_k, \cdots, \tau_n = \gamma_n) \\ &\leq \sum_{\gamma_k, \cdots, \gamma_n} (W(\kappa_1, \cdots, \kappa_k, \gamma_{k+1}, \cdots, \gamma_n) + W(\kappa_1, \cdots, \kappa_{k-1}, \gamma_k, \cdots, \gamma_n)) \\ &\qquad \times Pr(\tau_k = \gamma_k, \cdots, \tau_n = \gamma_n) \\ &= E[W|\mathcal{B}_k](\kappa) + E[W|\mathcal{B}_{k-1}](\kappa) \; . \end{aligned}$$

$\square$

Now we come back our own problem. Before developing some inequalities of the form (70), we introduce more convenient notation: For $V(\hat{H}) := \{v_1, v_2, ..., v_n\}$ we write $\tau_k := \tau(v_k)$, $k \in [n]$. We will specify the order of the vertices later depending on our purpose. From now on, $\mathcal{B}_k$ is the σ-field generated by $\tau_1, \cdots, \tau_k$ and B_0 is the trivial σ-field that consists of the empty set and the whole set. We also write

$$\hat{N}_k := \hat{N}(v_k), \quad T_k := T(v_k), \quad T'_k := T'(v_k) \quad \text{and} \quad \hat{N}_k^\gamma := \hat{N}(v_k; \gamma) \ .$$

(Notice that T_k is in fact $T_i(v_k)$.)

We define new random variables

$$Q_{jk}(\tau) = \begin{cases} 1 & \text{if } v_j \sim v_k \text{ and } \tau_j = \tau_k \neq \Lambda \\ 0 & \text{otherwise.} \end{cases}$$

and

$$R^\gamma_{jk}(\tau) = \begin{cases} 1 & \text{if (1) } \tau_k = \gamma, \text{ and (2) } v_j \sim v_k \text{ or } \exists\ v_l \in \hat{N}_j^\gamma \cap \hat{N}_k^\gamma \ \cdot \ni \cdot \ \tau_l = \gamma \\ 0 & \text{otherwise.} \end{cases}$$

Remark 1. If $j \neq k$ then $|\hat{N}_j^\gamma \cap \hat{N}_k^\gamma| \leq 1$ because $g(\hat{H}) \geq 5$. Thus the second condition of (2) is very strong in most cases.

2. We could replace the condition $v_l \in \hat{N}_j^\gamma \cap \hat{N}_k^\gamma$ by $v_l \in \hat{N}_j \cap \hat{N}_k$, since the requirement $\tau_l = \gamma$ then forces $v_l \in \hat{N}_j^\gamma \cap \hat{N}_k^\gamma$.

As we saw in Section 4, our random variables are sums of 0-1 random variables. We first consider the 0-1 random variables.

Lemma 6.2 *Suppose $\tau \equiv_k \tau'$. Then we have*

$$|1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau')| \leq Q_{jk}(\tau) + Q_{jk}(\tau') + 1_{\{j=k\}} \tag{71}$$

$$|1_{\{\gamma \in T'_j\}}(\tau) - 1_{\{\gamma \in T'_j\}}(\tau')| \leq R^\gamma_{jk}(\tau) + R^\gamma_{jk}(\tau') \tag{72}$$

and

$$\begin{aligned} &|1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau')| \\ &\qquad \leq Q_{jk}(\tau) + Q_{jk}(\tau') + R^\gamma_{jk}(\tau) + R^\gamma_{jk}(\tau') + 1_{\{j=k\}} \end{aligned} \tag{73}$$

for $\gamma \in T_j$.

Proof. (a) For (71) suppose

$$1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau') = 1 \ .$$

Then we claim

$$Q_{jk}(\tau) + 1_{\{j=k\}} \geq 1 \ ,$$

which means

$$1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau') \le Q_{jk}(\tau) + 1_{\{j=k\}} \ . \tag{74}$$

Proof of claim. First note that

$$\begin{aligned} 1_{\{v_j \notin X\}}(\tau) = 1 \quad &\Rightarrow \quad \tau_j = \Lambda \ \text{ or } \ \tau_j = \tau_l \qquad && \text{for some } v_l \in \hat{N}_j \quad \text{and} \\ 1_{\{v_j \notin X\}}(\tau') = 0 \quad &\Rightarrow \quad \tau'_j \neq \Lambda \ \text{ and } \ \tau'_j \neq \tau'_l, \qquad && \text{for all } v_l \in \hat{N}_j \ . \end{aligned}$$

We consider two cases.
(1) If $\tau_j \neq \tau'_j$ then $k = j$. Thus $1_{\{j=k\}} = 1$.
(2) Suppose $\tau_j = \tau'_j (\neq \Lambda)$. Then we know $\tau_j \neq \Lambda$ and there is $v_l \in \hat{N}_j$ such that $\tau_j = \tau_l \neq \tau'_l$. Thus $l = k$ and $\tau_j = \tau_k \neq \Lambda$ i.e. $Q_{jk}(\tau) = 1$.

Similarly, we may have

$$1_{\{v_j \notin X\}}(\tau') - 1_{\{v_j \notin X\}}(\tau) \le Q_{jk}(\tau') + 1_{\{j=k\}} \ ,$$

which completes the proof.

(b) For (72) suppose that

$$1_{\{\gamma \in T'_j\}}(\tau) - 1_{\{\gamma \in T'_j\}}(\tau') = 1 \ .$$

Then we claim

$$R^{\gamma}_{jk}(\tau) + R^{\gamma}_{jk}(\tau') \ge 1 \ .$$

Proof of claim. First we have

$$\begin{aligned} &1_{\{\gamma \in T'_j\}}(\tau) = 1 \Rightarrow \forall v_l \in A := \{v_l \sim v_j : \tau_l = \gamma\} \quad \exists \ v_q \sim v_l \ \ \cdot \ni \cdot \ \ \tau_q = \gamma \quad \text{and} \\ &1_{\{\gamma \in T'_j\}}(\tau') = 0 \Rightarrow \exists v_l \in A' := \{v_l \sim v_j : \tau'_l = \gamma\} \ \ \cdot \ni \cdot \ \ \tau'_q \neq \gamma \quad \forall \ \ v_q \sim v_l \ . \end{aligned}$$

We again consider two cases.
(1) If $A' \setminus A \neq \emptyset$ then it is clear by $\tau \equiv_k \tau'$ that $A' \setminus A = \{v_k\}$. Thus $v_j \sim v_k$ and $\tau'_k = \gamma$ by the definition of A'. This means $R^{\gamma}_{jk}(\tau') = 1$.
(2) Suppose $A' \subseteq A$. Then take $v_l \in A'$ such that $\tau'_q \neq \gamma$ for all $v_q \sim v_l$. Since v_l is also in A ($\Rightarrow \tau_l = \gamma$), we know there is $v_{q_0} \sim v_l$ such that $\tau_{q_0} = \gamma$. Thus it is clear to see that $q_0 = k$ and so $R^{\gamma}_{jk}(\tau) = 1$. (Note that this includes the case $k = j$.)

Similarly, we have the same claim when the other case happens, which completes the proof.

(c) The inequality (73) follows from (71) and (72) via the triangle inequality, since

$$\begin{aligned} &|1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau')| \\ &\qquad \le |1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau')| + |1_{\{\gamma \in T'_j\}}(\tau) - 1_{\{\gamma \in T'_j\}}(\tau')| \ . \end{aligned}$$

$\square$

Finally we have the following easy lemma.

Lemma 6.3 *If $v_j \sim v_k$ and $j > k$ then we have*

$$\begin{aligned} E[Q_{jk}|\mathcal{B}_k] &= p1_{\{\tau_k \in T_j\}} \\ E[Q_{jk}|\mathcal{B}_{k-1}] &= p^2|T_j \cap T_k| \ . \end{aligned}$$

Also, if all vertices in $\hat{N}_j$ follow v_k then we have

$$\begin{aligned} E[R^{\gamma}_{jk}|\mathcal{B}_k] &\leq p|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k|1_{\{\tau_k=\gamma\}} \quad (\leq 1_{\{\tau_k=\gamma\}}) \\ E[R^{\gamma}_{jk}|\mathcal{B}_{k-1}] &\leq p^2|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k|1_{\{\gamma\in T_k\}} \quad (\leq p1_{\{\gamma\in T_k\}}) \end{aligned}$$

with equality unless $j = k$.

Proof. Suppose $v_j \sim v_k$ and $j > k$. Then

$$E[Q_{jk}|\mathcal{B}_k] = Pr(\tau_k = \tau_j \neq \Lambda|\mathcal{B}_k) \ .$$

Since τ_j is independent of $\mathcal{B}_k$, we get

$$Pr(\tau_k = \tau_j \neq \Lambda|\mathcal{B}_k) = \begin{cases} p & \text{if } \tau_k \in T_j \\ 0 & \text{otherwise.} \end{cases}$$

And since τ_k is independent of $\mathcal{B}_{k-1}$, it is clear that

$$\begin{aligned} E[Q_{jk}|\mathcal{B}_{k-1}] &= E[E[Q_{jk}|\mathcal{B}_k]|\mathcal{B}_{k-1}] \\ &= pE[1_{\{\tau_k\in T_j\}}|\mathcal{B}_{k-1}] \\ &= p^2|T_j \cap T_k| \ . \end{aligned}$$

For the second part, suppose all vertices in $\hat{N}_j$ follow v_k, in particular $v_k \not\sim v_j$. Then

$$\begin{aligned} E[R^{\gamma}_{jk}|\mathcal{B}_k] &= Pr(\exists v_l \in \hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k \ \cdot\ni\cdot \ \tau_l = \gamma|\mathcal{B}_k)1_{\{\tau_k=\gamma\}} \\ &\leq p|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k|1_{\{\tau_k=\gamma\}} \end{aligned} \tag{75}$$

since

$$\begin{aligned} Pr(\exists v_l \in \hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k \ \cdot\ni\cdot \ \tau_l = \gamma|\mathcal{B}_k) &= Pr(\exists v_l \in \hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k \ \cdot\ni\cdot \ \tau_l = \gamma) \\ &\leq p|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k| \ . \end{aligned}$$

And

$$E[R^{\gamma}_{jk}|\mathcal{B}_{k-1}] = p^2|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k|1_{\{\gamma\in T_k\}} \ . \tag{76}$$

Furthermore, in (75), we have equality whenever $|\hat{N}^{\gamma}_j \cap \hat{N}^{\gamma}_k| = 0$ or 1, which happens unless $j = k$ (since $g(\hat{H}) \geq 5$).

□

In what follows we will treat concentrations of the random variables $d'(v)$, $t'(v)$ and $d'(v;\gamma)$ separately. Since we would like to apply Lemma 5.3 the main goal is to establish inequalities of the form (65) or (66). In most cases $A_k = \emptyset$ and $I = [n]$, but in the proof of the concentration result for $d(v;\gamma)$ we use Lemma 5.2 essentially (i.e. $A_k \neq \emptyset$ in some cases) and I is no longer $[n]$. In each case we first choose the order of vertices carefully. Next we apply lemmas 6.1 and 6.2, and analyze the resulting upper bounds case by case (using Lemma 6.3 in most cases). Again in the the proof of the concentration result for $d(v;\gamma)$, we need to consider R^{γ}_{jk} under more complicated conditions, which will be developed in Section 7.3.

In the following section we always assume

$$\tau \equiv_k \tau'$$

when k is clear.

7 Proof of the Main Lemma

In this section we prove (35), (36) and (37) in the Main Lemma.

7.1 Degrees

Fix $v_1 = v \in V(H)$. Since $\hat{N}(N(v)) \cap N(v) = \emptyset$ by $g(\hat{H}) \geq 5$, we may label all vertices so that

$$\hat{N}(N(v)) \setminus \{v\} = \{v_2, ..., v_{m-1}\} \quad \text{and} \quad N(v) = \{v_m, ..., v_n\}$$

(recall $N(v) = \{w \in V(H) : w \sim v\}$). Note that $v_j \neq v_k$ if $j \neq k$ since $g(\hat{H}) \geq 5$. Our random variable Y is, of course,

$$Y = d'(v) = \sum_{w \in N(v)} 1_{\{w \notin X\}} = \sum_{j=m}^{n} 1_{\{v_j \notin X\}} \ .$$

We do not even define the order of the other vertices because Y does not depend on their colors.

We look for inequalities of the form (70). For $\tau \equiv_k \tau'$ we easily see that by (71)

$$\begin{aligned} |Y(\tau) - Y(\tau')| &\leq \sum_{j=m}^{n} |1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau')| \\ &\leq \sum_{j=m}^{n} (Q_{jk}(\tau) + Q_{jk}(\tau') + 1_{\{j=k\}}) \ . \end{aligned}$$

and by Lemma 6.1 we have

$$|Z_k| \leq \sum_{j=m}^{n} (E[Q_{jk}|\mathcal{B}_k] + E[Q_{jk}|\mathcal{B}_{k-1}] + 1_{\{j=k\}}) \ . \tag{77}$$

Now we claim that

$$Pr(Y - E[Y] \geq \lambda) \leq \exp(-(\log \Delta)^2/4)$$

where $\lambda := \Delta^{1/2} \log \Delta$.

First recall

$$pt \le 1 \;\; \text{(by the definition of } p) \qquad pd \le 0.11 \;\; \text{(by (45) or (28))}. \tag{78}$$

We consider three cases to get inequalities of the form (65). In what follows, we always assume $m \le j \le n$.
(Case 1) $\;k = 1$
Then using Lemma 6.3, (77) and the fact that $|N(v;\gamma)| \le |\hat{N}(v;\gamma)| = d$ for all $\gamma \in T(v)$, we have

$$\begin{aligned} |Z_1| &\le p \sum_{j=m}^{n} (1_{\{\tau_1 \in T_j\}} + p|T_j \cap T_1| \\ &= p(|N(v;\tau_1)| + ptd) \\ &\le p(d + ptd) \le 2pd \end{aligned}$$

Therefore, we have

$$|Z_1| \le 2pd \le 1 = 1/2 + 1/2,$$

i.e.

$$c_1 = 1/2, \qquad Pr(B_1) = 1. \tag{79}$$

in terms of parameters in (65).

(Case 2) $\;2 \le k \le m-1$.
In this case there is only one j ($m \le j \le n$), say $j(k)$, such that $v_j \sim v_k$. By (77) and Lemma 6.3 we have

$$|Z_k| \le p1_{\{\tau_k \in T_{j(k)}\}} + p^2|T_k \cap T_{j(k)}| \; .$$

That is, for (65) we may take $B_k := \{\tau_k \in T_{j(k)}\}$ and

$$c_k = p \qquad \text{and} \qquad Pr(B_k) = p|T_k \cap T_{j(k)}| \; . \tag{80}$$

(Case 3) $\;m \le k \le n$,
Since $v_k \sim v$ and $v_j \sim v$ we know $v_k \not\sim v_j$. Thus all Q terms in (77) disappear. Therefore, we have

$$|Z_k| \le 1 \quad \text{i.e.} \quad c_k = 1/2 \qquad \text{and} \qquad Pr(B_k) = 1 \; . \tag{81}$$

Therefore, by (79), (80) and (81), we know that

$$3\omega^2 \sum_{k=1}^{n} c_k^2 Pr(B_k) = 3\omega^2 (\frac{1}{4} + p^3 \sum_{k=2}^{m-1} |T_k \cap T_{j(k)}| + \frac{1}{4}|N(v)|) \; .$$

Furthermore,

$$|N(v)| \le \Delta$$

and by (78)

$$\begin{aligned}
p^3 \sum_{k=2}^{m-1} |T_k \cap T_{j(k)}| &\leq p^3 \sum_{j=m}^{n} \sum_{v_k \in \hat{N}_j} |T_k \cap T_j| \\
&= p^3 \sum_{j=m}^{n} \sum_{v_k \in \hat{N}_j} \sum_{\gamma \in T_j} 1_{\{\gamma \in T_k\}} \\
&= p^3 \sum_{j=m}^{n} \sum_{\gamma \in T_j} \sum_{v_k \in \hat{N}_j} 1_{\{\gamma \in T_k\}} \\
&\leq p^3 \Delta t d < p\Delta \ . \qquad (82)
\end{aligned}$$

Finally, setting $\omega = \lambda/(2\Delta)$ and using Lemma 5.3 we have

$$Pr(Y - E[Y] \geq \lambda) \leq \exp(-\omega\lambda + \omega^2\Delta) = \exp(-(\log\Delta)^2/4) \ .$$

7.2 Sizes of sets of legal colors

We define an order similar to that of the previous section. Fix $v \in V(H)$ and set $v_1 = v$ and

$$\hat{N}^2(v) = \{v_2, ..., v_{m-1}\}, \quad \hat{N}(v) = \{v_m, ..., v_n\} \ ,$$

where, in general, for a subset (or vertex) A of $V(\hat{H})$

$$\hat{N}^0(A) = A \quad \text{and} \quad \hat{N}^j(A) := \hat{N}(\hat{N}^{j-1}(A)) \setminus \bigcup_{l=0}^{j-1} \hat{N}^l(A) \qquad \text{for } l = 1, 2, \cdots \ .$$

Notice that by the definition

$$\hat{N}^j(A) \cap A = \emptyset \quad \text{for all } j = 1, 2, ... \qquad (83)$$

We do not define any order on the other vertices because they are irrelevant.

If we set

$$Y = -t'(v_1) = -\sum_{\gamma \in T_1} 1_{\{\gamma \in T_1'\}} \ ,$$

then for $\tau \equiv_k \tau'$ we have by (72)

$$\begin{aligned}
|Y(\tau) - Y(\tau')| &\leq \sum_{\gamma \in T_1} |1_{\{\gamma \in T_1'\}}(\tau) - 1_{\{\gamma \in T_1'\}}(\tau')| \\
&\leq \sum_{\gamma \in T_1} (R_{1k}^{\gamma}(\tau) + R_{1k}^{\gamma}(\tau')) \ .
\end{aligned}$$

Hence by Lemma 6.1

$$|Z_k| \leq \sum_{\gamma \in T_1} (E[R_{1k}^{\gamma} | \mathcal{B}_k] + E[R_{1k}^{\gamma} | \mathcal{B}_{k-1}]) \ . \qquad (84)$$

We claim

$$Pr(Y - E[Y] \geq \lambda) \leq \exp(-(\log t)^2/2)$$

for $\lambda := t^{1/2} \log t$.

Again we first consider three cases.
(Case 1) $k = 1$
Then by (84), Lemma 6.3 and the fact that $|\hat{N}_k^\gamma| = d$, we have

$$|Z_1| \leq pd \sum_{\gamma \in T_1} (1_{\{\tau_1 = \gamma\}} + p) \leq 1$$

i.e.

$$c_1 = 1/2, \qquad Pr(B_1) = 1 \ . \tag{85}$$

in terms of the parameters in (65).
(Case 2) $2 \leq k \leq m-1$
Then there is only one element in $\hat{N}_1 \cap \hat{N}_k$, say $v_{j(k)}$. By (84) and Lemma 6.3 (using $j(k) > k$) we have

$$\begin{aligned} |Z_k| &\leq \sum_{\gamma \in T_1} (p|\hat{N}_1^\gamma \cap \hat{N}_k^\gamma| 1_{\{\tau_k = \gamma\}} + p^2 |\hat{N}_1^\gamma \cap \hat{N}_k^\gamma| 1_{\gamma \in T_k}) \\ &= \sum_{\gamma \in T_1} (p 1_{\{\gamma \in T_{j(k)}\}} 1_{\{\tau_k = \gamma\}} + p^2 1_{\{\gamma \in T_{j(k)}\}} 1_{\{\gamma \in T_k\}}) \\ &= p 1_{\{\tau_k \in T_1 \cap T_{j(k)}\}} + p^2 |T_1 \cap T_{j(k)} \cap T_k| \ . \end{aligned}$$

Thus we may say $B_k := \{\tau_k \in T_1 \cap T_{j(k)}\}$ and

$$c_k = p, \qquad Pr(B_k) = p|T_1 \cap T_{j(k)} \cap T_k|. \tag{86}$$

(case 3) $m \leq k \leq n$
Then by (84) and Lemma 6.3 we have

$$|Z_k| \leq \sum_{\gamma \in T_1} (1_{\{\tau_k = \gamma\}} + p 1_{\{\gamma \in T_k\}}) = 1_{\{\tau_k \in T_1\}} + p|T_1 \cap T_k| \ ,$$

that is, $B_k := \{\tau_k \in T_1\}$ and

$$c_k = 1, \qquad Pr(B_k) = p|T_1 \cap T_k| \ . \tag{87}$$

Now by (85), (86) and (87), we have

$$3\omega^2 \sum_{k=1}^{n} c_k^2 Pr(B_k) \quad = \quad 3\omega^2 (\frac{1}{4} + p^3 \sum_{k=2}^{m-1} |T_1 \cap T_{j(k)} \cap T_k| + p \sum_{k=m}^{n} |T_1 \cap T_k|) \ .$$

Moreover, by (78) we have

$$p \sum_{k=m}^{n} |T_1 \cap T_k| = pdt \leq 0.11t$$

and

$$\begin{aligned} p^3 \sum_{k=2}^{m-1} |T_1 \cap T_{j(k)} \cap T_k| &\leq p^3 \sum_{v_j \in \hat{N}_1} \sum_{v_k \in \hat{N}_j} \sum_{\gamma \in T_1} 1_{\{\gamma \in T_j \cap T_k\}} \\ &= p^3 \sum_{\gamma \in T_1} \sum_{v_j \in \hat{N}_1^\gamma} \sum_{v_k \in \hat{N}_j} 1_{\{\gamma \in T_k\}} \\ &= p^3 t d^2 \leq 1 \ . \end{aligned}$$

So setting $\omega = \lambda/t$ and using Lemma 5.3, we have

$$Pr(Y - E[Y] \geq \lambda) \leq \exp(-\omega\lambda + \omega^2 t/2) = \exp(-(\log t)^2/2) \ .$$

□

7.3 Color degrees

As we saw before, this case is a combination of the preceding two cases. One might guess that the upper bound we try to get is more and less the sum of the two previous upper bounds. However, our situation here is somewhat different so that we need a more subtle and complicated analysis. The reason will be briefly explained after we order vertices.

Fix $v \in V(H)$ and $\gamma \in T(v)$. Set

$$\begin{aligned} \{v_1, ..., v_{h-1}\} &= \hat{N}^2(N(v;\gamma)) \\ \{v_h, ..., v_{l-1}\} &= \hat{N}(N(v;\gamma)) \cap \{z \in V(\hat{H}) : \gamma \notin T(z)\} \\ \{v_l, ..., v_{m-1}\} &= \hat{N}(N(v;\gamma)) \cap \{z \in V(\hat{H}) : z \neq v, \gamma \in T(z)\} \\ \{v_m, ..., v_{n-1}\} &= N(v;\gamma) \end{aligned}$$

and $v_n = v$. Also set

$$Y = d'(v;\gamma) = \sum_{z \in N(v;\gamma)} 1_{\{z \notin X,\ \gamma \in T'(z)\}} = \sum_{j=m}^{n-1} 1_{\{v_j \notin X,\ \gamma \in T'_j\}} \ .$$

Then as in the previous sections for $\tau \equiv_k \tau'$ we have

$$\begin{aligned} |Y(\tau) - Y(\tau')| &\leq \sum_{j=m}^{n-1} |1_{\{v_j \notin X,\ \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X,\ \gamma \in T'_j\}}(\tau')| \\ &\leq \sum_{j=m}^{n-1} Q_{jk}(\tau) + Q_{jk}(\tau') + R^\gamma_{jk}(\tau) + R^\gamma_{jk}(\tau') + 1_{\{j=k\}} \ , \end{aligned}$$

and so by Lemma 6.1

$$|Z_k| \leq \sum_{j=m}^{n-1} E[Q_{jk}|\mathcal{B}_k] + E[Q_{jk}|\mathcal{B}_{k-1}] + E[R^\gamma_{jk}|\mathcal{B}_k] + E[R^\gamma_{jk}|\mathcal{B}_{k-1}] + 1_{\{j=k\}} \ . \qquad (88)$$

For the Q terms we may use the same estimation as in Section 7.1. However for the R terms we need new analysis. Shortly speaking, one (possibly main) reason is that we must take account into edges between vertices $U := \{v_l, \cdots, v_{m-1}\}$. For example, it may happen that there is a vertex v_k in U such that almost all vertices in N_k^γ are in U and precede v_k. Furthermore, it seems to be impossible to find a suitable order to avoid this kind of problem. Thus we are considering essential maximums. The next two lemmas are presented mainly for this purpose.

First we define new (random) sets

$$\begin{aligned} A_k^\gamma = A_k^\gamma(\tau) &:= \{v_i \in \hat{N}_k^\gamma : 1 \le i \le k-1, \quad \tau_i = \gamma\} \\ C_k^\gamma = C_k^\gamma(\tau) &:= \{v_i \in \hat{N}_k^\gamma : k \le i \le n, \quad \tau_i = \gamma\} \ . \end{aligned}$$

Then it easy to see that for $v_k \in \hat{N}_j$

$$R_{jk}^\gamma = 1_{\{\tau_k = \gamma\}} \tag{89}$$

and for $v_k \notin \hat{N}_j$

$$R_{jk}^\gamma \le (|\hat{N}_j^\gamma \cap A_k^\gamma| + |\hat{N}_j^\gamma \cap C_k^\gamma|)1_{\{\tau_k=\gamma\}} \ .$$

Furthermore, since $A_k^\gamma \in \mathcal{B}_{k-1} \subset \mathcal{B}_k$ and C_k^γ is independent of $\mathcal{B}_k$, we have

$$\begin{aligned} E[|\hat{N}_j^\gamma \cap A_k^\gamma| \, |\mathcal{B}_k] &= |\hat{N}_j^\gamma \cap A_k^\gamma| \\ E[|\hat{N}_j^\gamma \cap C_k^\gamma| \, |\mathcal{B}_k] &\le p|\hat{N}_j^\gamma \cap \hat{N}_k^\gamma| \ . \end{aligned}$$

Thus for $v_k \notin \hat{N}_j$, we have

$$E[R_{jk}^\gamma|\mathcal{B}_k] \le (|\hat{N}_j^\gamma \cap A_k^\gamma| + p|\hat{N}_j^\gamma \cap \hat{N}_k^\gamma|)1_{\{\tau_k=\gamma\}} \ . \tag{90}$$

The next lemma is easy to get using the above inequalities.

Lemma 7.1 *With the notation as above we have*

$$\sum_{j=m}^{n-1} E[R_{jk}^\gamma|\mathcal{B}_k] \le \begin{cases} c_k 1_{\{\tau_k=\gamma\}} & \textit{if } \ 1 \le k \le h \\ (2+|A_k^\gamma|)1_{\{\tau_k=\gamma\}} & \textit{if } \ h \le k \le m-1 \\ 1+pd & \textit{if } \ m \le k \le n-1 \end{cases}$$

where for $1 \le k \le h$

$$c_k = c_k^\gamma := p \ |\bigcup_{j=m}^{n-1} \hat{N}_j^\gamma \cap \hat{N}_k^\gamma| \ .$$

Proof. For $1 \le k \le h$ we know $\hat{N}_j^\gamma \cap A_k^\gamma = \emptyset$ since the all vertices in $\hat{N}_j^\gamma$ follow v_k. Also it is easy to see that

$$p\sum_{j=m}^{n-1} |\hat{N}_j^\gamma \cap \hat{N}_k^\gamma| = p|\bigcup_{j=m}^{n-1} \hat{N}_j^\gamma \cap \hat{N}_k^\gamma| = c_k \quad (\le pd) \tag{91}$$

because the sets in the sum are disjoint by $g(\hat{H}) \ge 5$. Thus by (90) we have

$$\sum_{j=m}^{n-1} E[R_{jk}^\gamma|\mathcal{B}_k] \le p\sum_{j=m}^{n-1} |\hat{N}_j^\gamma \cap \hat{N}_k^\gamma| 1_{\{\tau_k=\gamma\}} = c_k 1_{\{\tau_k=\gamma\}}$$

On the other hand, for $h \le k \le m-1$ there is only one j between m and $n-1$ such that $v_k \in \hat{N}_j$. Hence by (89), (90) and (91) we get

$$\begin{aligned}\sum_{j=m}^{n-1} E[R^{\gamma}_{jk}|\mathcal{B}_k] &\le (1+\sum_{j=m}^{n-1} |\hat{N}^{\gamma}_j \cap A^{\gamma}_k| + pd)1_{\{\tau_k=\gamma\}} \\ &\le (2+|A^{\gamma}_k|)1_{\{\tau_k=\gamma\}}\end{aligned}$$

again because of the disjointness of the sets.

Finally, for $m \le k \le n-1$ we know that if $j \ne k$ then $\hat{N}_j \cap \hat{N}_k = \{v_n\}$, which also means $\hat{N}_j \cap A^{\gamma}_k = \emptyset$. Thus by (90), we get

$$\sum_{j=m}^{n-1} E[R^{\gamma}_{jk}|\mathcal{B}_k] \le E[R^{\gamma}_{kk}|\mathcal{B}_k] + (d-1)p1_{\{\tau_k=\gamma\}} \le 1+pd\ .$$

□

In the above lemma, the size of A^{γ}_k can be as large as d. But the size is essentially small enough for our purpose. (Note that $E[|A^{\gamma}_k|] \le pd \le 0.11$.) The following lemma gives the exact meaning of this.

Lemma 7.2 *For all $\gamma_0 \in T_k$, we have*

$$Pr(|A^{\gamma_0}_k| \ge \log d) \le d\exp(-\log d \log\log d)$$

Proof. Set $Y' = |A^{\gamma_0}_k|$. For $\omega' = \log\log d$ we get

$$\begin{aligned}E[\exp(\omega' Y')] &\le E[\exp(\omega' \sum_{v_i \in \hat{N}^{\gamma_0}_k} 1_{\{\tau_i=\gamma_0\}})] \\ &= \prod_{v_i \in \hat{N}^{\gamma_0}_k} E[\exp(\omega' 1_{\{\tau_i=\gamma_0\}})] \\ &\le (1-p+pe^{\omega'})^d \\ &\le \exp(pde^{\omega'}) \\ &\le \exp(e^{\omega'}) = d\ .\end{aligned}$$

Thus using Markov inequality we have

$$\begin{aligned}Pr(Y' \ge \log d) &= Pr(\exp(\omega' Y') \ge \exp(\omega' \log d)) \\ &\le d\exp(-\omega' \log d)\ .\end{aligned}$$

□

Now we claim for $\lambda := d^{1/2}(\log d)^2$,

$$Pr(Y - E[Y] \ge \lambda) \le \exp(-\frac{1}{2}\log d \log\log d) \qquad (92)$$

using Lemma 5.3. That is, we first show that (65) and (66) with appropriate c_k's, B_k's, A_{k-1}'s which satisfy the conditions in Lemma 5.3.

We consider five cases. In what follows we always assume $m \le j \le n-1$.

(Case 1) $1 \le k \le h-1$
Note that $j \ne k$, and by (83) $v_j \not\sim v_k$ for all $m \le j \le n-1$. Thus all Q terms in (88) disappear as well as the term $1_{\{j=k\}}$. By (88) and Lemma 7.1, we have

$$|Z_k| \le c_{_k} 1_{\{\tau_{_k}=\gamma\}} + pc_{_k} 1_{\{\gamma \in T_k\}} \ .$$

(Case 2) $h \le k \le l-1$
By $\gamma \notin T_k$, all R terms in (88) disappear. Furthermore because there is only one j, say $j(k)$, such that $v_{_k} \sim v_{_j}$, we have

$$|Z_k| \le p1_{\{\tau_k \in T_{j(k)}\}} + p^2|T_{j(k)} \cap T_k| \ \ (\le 2p) \ . \tag{93}$$

as in the Case 2 of Section 7.1.

Hence $B_k := \{\tau_k \in T_{j(k)}\}$ and

$$c_k = p, \qquad Pr(B_k) = p|T_{j(k)} \cap T_k| \ . \tag{94}$$

(Case 3) $l \le k \le m-1$
Let $j(k)$ as in (Case 2). Then we have the same bound in (93) for Q terms. Now we set

$$A_{k-1} := \{\tau : |A^{\gamma}_{_k}(\tau)| \ge \log d\} \in \mathcal{B}_{k-1} \ .$$

Then by (88) and Lemma 7.1 we have

$$|Z_k|1_{\bar{A}_{k-1}} \le 2p + (2+\log d)1_{\{\tau_{_k}=\gamma\}} + p(2+\log d) \le (4+\log d)1_{\{\tau_{_k}=\gamma\}} + p(4+\log d) \ .$$

Hence we may say that $B_k := \{\tau_{_k} = \gamma\}$ and

$$c_{_k} = 4+\log d, \qquad Pr(A_{k-1}) \le a_{_k}, \qquad Pr(B_k) = p \tag{95}$$

where $a_k := \exp(-d\log d \log\log d)$ (see Lemma 7.2).

(Case 4) $m \le k \le n-1$
Note that $v_j \notin \hat{N}_k$ and for $k \ne j$, $\hat{N}_j \cap \hat{N}_k = \{v_n\}$ $(m \le j \le n-1)$. So all Q terms disappear. Therefore, by (88) and Lemma 7.1, we get

$$|Z_k| \le 2 + 2pd + 1 \le 4 \tag{96}$$

That is, $c_{_k} = 2$ and $Pr(B_k) = 1$.

(Case 5) $k = n$
For

$$M_n(\tau) := \max_{\gamma_0 \in T_n} \{|A^{\gamma_0}_{_n}(\tau)|\} \ ,$$

we define

$$A_{n-1} := \{\tau : M_n(\tau) \ge \log d\} \in \mathcal{B}_{n-1} \ .$$

Then it is easy to check by Lemma 7.2 that

$$Pr(A_{n-1}) \le td\exp(-\log d \log\log d) \ .$$

We now claim

$$Z_n 1_{\bar{A}_{n-1}} \leq 2 + \log d \ ,$$

that is, $J = \{n\}$ and

$$c_n = 2 + \log d \tag{97}$$

in terms of parameters in Lemma 5.3.

Proof of claim. For Q terms, note that

$$\sum_{j=m}^{n-1} Q_{jn}(\tau) = \sum_{j=m}^{n-1} 1_{\{\tau_j = \tau_n \neq \Lambda\}}(\tau) \leq M_n(\tau)$$

and

$$\sum_{j=c}^{m} E[Q_{jn}|\mathcal{B}_{n-1}] \leq pd \leq 1$$

Hence by (88) we have

$$\begin{aligned} |Z_n| 1_{\overline{A}_{n-1}} &\leq \log d + 1 + \sum_{j=m}^{n-1} (1_{\{\tau_n = \gamma\}} + p) \\ &\leq \log d + 1 + d 1_{\{\tau_n = \gamma\}} + pd \\ &= 2 + \log d + d 1_{\{\tau_n = \gamma\}} \end{aligned} \tag{98}$$

If $\tau_n \neq \gamma$ then we get

$$|Z_n| 1_{\overline{A}_{n-1}} \leq 2 + \log d \ .$$

When $\tau_n = \gamma$, the upper bound in (98) is no longer good. Actually the (essential) maximum of $|Z_n|$ is quite big. (Note that p is not so small.) But we can find a nice essential upper bound of Z_n. To do so we need a lemma, which is to be proved later. Our result is an easy corollary of the lemma.

Recall that it is enough for us to consider only the case $\tau_n = \gamma$.

Lemma 7.3 *With the same notation as above, suppose $\tau \equiv_n \tau'$ and $\tau_n = \gamma$. Then for $m \leq j \leq n-1$*

$$1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau') \leq 1_{\{\tau_j = \gamma\}}(\tau) \ . \tag{99}$$

Corollary 7.4 *If $\tau_n = \gamma$ then*

$$Z_n 1_{\overline{A}_{n-1}} \leq \log d \ .$$

Proof. We use the same method in the proof of Lemma 6.1. For $\tau = (\tau_1, ..., \tau_{n-1}, \gamma)$ we know

$$\begin{aligned} Z_n(\tau) &= Y(\tau) - E[Y|\tau_1, ..., \tau_{n-1}] \\ &= \sum_{\gamma' \in T_n \cup \{\Lambda\}} (Y(\tau) - Y(\tau')) Pr(\tau_n = \gamma') \\ &= \sum_{\gamma' \in T_n \cup \{\Lambda\}} \sum_{j=m}^{n-1} (1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau')) Pr(\tau_n = \gamma') \end{aligned}$$

where $\tau' = (\tau_1, ..., \tau_{n-1}, \gamma')$. Thus by Lemma 7.3 we have

$$\begin{aligned}
Z_n 1_{\overline{A}_{n-1}} &\leq \sum_{\gamma' \in T_n \cup \{\Lambda\}} \sum_{j=m}^{n-1} 1_{\{t_j = \gamma\}} 1_{\overline{A}_{n-1}} Pr(\tau_n = \gamma') \\
&= \sum_{j=m}^{n-1} 1_{\{t_j = \gamma\}} 1_{\overline{A}_{n-1}} \sum_{\gamma' \in T_n \cup \{\Lambda\}} Pr(\tau_n = \gamma') \\
&= \sum_{j=m}^{n-1} 1_{\{\tau_j = \gamma\}} 1_{\overline{A}_{n-1}} \leq \log d \ .
\end{aligned}$$

$\square$

We now have

$$\begin{aligned}
3\omega^2 \sum_{k=1}^{n-1} c_k^2 Pr(B_k) &= 3\omega^2 (p \sum_{k=1}^{h-1} c_k^2 1_{\{\gamma \in T_k\}} + p^3 \sum_{k=h}^{l-1} |T_{j(k)} \cap T_k| \\
&\quad + p \sum_{k=l}^{m-1} (4 + \log d)^2 + \sum_{k=m}^{n-1} 4) \ .
\end{aligned}$$

Also, it is easy to check that

$$p^3 \sum_{k=h}^{l-1} |T_{j(k)} \cap T_k| \leq p^3 t d^2 \leq 1, \quad p \sum_{k=l}^{m-1} (4 + \log d)^2 \leq p d^2 (4 + \log d)^2 \leq 0.12 d (\log d)^2 \ ,$$

and

$$\begin{aligned}
p \sum_{k=1}^{h-1} {c_k}^2 1_{\{\gamma \in T_k\}} &= p^3 \sum_{k=1}^{h-1} | \bigcup_{j=m}^{n-1} \hat{N}_j^\gamma \cap \hat{N}_k^\gamma |^2 1_{\{\gamma \in T_k\}} \\
&\leq p^3 d \sum_{k=1}^{h-1} | \bigcup_{j=m}^{n-1} \hat{N}_j^\gamma \cap \hat{N}_k^\gamma | 1_{\{\gamma \in T_k\}} \\
&\leq p^3 d d^2 d = p^3 d^4
\end{aligned}$$

since the last sum is less than the number of edges between $\cup_{j=m}^{n-1} \hat{N}_j^\gamma$ and its neighbors v_k with $\gamma \in T_k$.

Hence setting $\omega = d^{-1/2}$ and using Lemma 5.3 (recall $\lambda = d^{1/2} (\log d)^2$) we have

$$\begin{aligned}
Pr(Y - E[Y] \geq \lambda) &\leq \exp(-\omega(\lambda - 2 - \log d) + 3\omega^2 (p^3 d^4 + 1 + 0.12 d (\log d)^2 + 4d)) \\
&\quad + Pr(\bigcup_{k=l}^{m-1} A_{k-1} \cup A_{n-1}) \\
&\leq \exp(-\frac{1}{2} (\log d)^2) + (d^3 + td) \exp(-\log d \log \log d) \\
&\leq \exp(-\log d \log \log d / 2) \ .
\end{aligned}$$

$\square$

We complete the proof of the Main Lemma by proving Lemma 7.3.

Proof of Lemma 7.3. First recall $v_n \sim v_j$. We consider two cases.

If $v_n \in X(\tau)$ (i.e. $1_{\{v_n \in X\}}(\tau) = 1$) then since $\tau_n = \gamma$, we have $\gamma \notin T'_j(\tau)$ (i.e. $1_{\{\gamma \in T'_j\}}(\tau) = 0$), which implies

$$1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) = 0 \ .$$

Thus the left hand side of (99) is less than 0 while $1_{\{\tau_j = \gamma\}} \geq 0$.

If $v_n \notin X(\tau)$ then it is easy to see

$$\begin{aligned} \gamma \notin T'_j(\tau) \quad & \text{if and only if} \quad \exists v_i \in \hat{N}_j \cap X(\tau) \ \ s.t. \ \ \tau_i = \gamma \\ & \text{if and only if} \quad \exists v_i \in \hat{N}_j \cap X(\tau') \ \ s.t. \ \ \tau'_i = \gamma \\ & \text{if and only if} \quad \ \ \gamma \notin T'_j(\tau') \end{aligned}$$

because $\tau \equiv_n \tau'$ and $g(\hat{H}) \geq 5$. That is, $1_{\{\gamma \in T'_j\}}(\tau) = 1_{\{\gamma \in T'_j\}}(\tau')$.

Thus by (74) we have

$$\begin{aligned} 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau) - 1_{\{v_j \notin X, \gamma \in T'_j\}}(\tau') \ &= \ (1_{\{v_j \notin X\}}(\tau) - 1_{\{v_j \notin X\}}(\tau'))1_{\{\gamma \in T'_j\}}(\tau) \\ &\leq \ Q_{jn}(\tau) = 1_{\{\tau_j = \gamma\}} \ . \end{aligned}$$

□

8 Further discussion

Our result (Theorem 1.1) gives the correct order of magnitude for both chromatic and list-chromatic numbers (cf. (2)). However, the original question regarding triangle-free graph (i.e. girth at least 4) is still open. Here we (J. Kahn and the author) would like conjecture that the same result holds for girth 4:

Conjecture 8.1 *Let G be a graph. If $g(G) \geq 4$ then*

$$\chi_l(G) \leq (1 + o(1)) \frac{\Delta(G)}{\log \Delta(G)}$$

where $o(1)$ goes to zero as $\Delta(G)$ goes to infinity.

Remark Recently, R. Häggkvist said that A. Johansson and S. McGuiness had just (independently) proved our result and were pretty sure that for girth 4 they could show $\chi(G) = O(\Delta(G)/\log \Delta(G))$ and $\chi_l(G) = o(\Delta(G))$.

Acknowledgments The author is very grateful to Professor J. Kahn for Lemma 5.2 which was jointly obtained with him.

References

[1] Ajtai, M., Komlós J., Szemerédi E. (1980), A note on Ramsey numbers, *J. Combinatorial Th. (A)* **29**, 354-360.

[2] Ajtai, M., Komlós J., Szemerédi, E. (1981), A dense infinite Sidon sequence, *European J. Combin.* **2**, 1-15.

[3] Alon, N., Restricted coloring of graphs, *Proc. 14th British Cobinatorial Conf.*, to appear.

[4] Bollobás, B. (1978), Chromatic number, girth and maximal degree, *Discrete Math.* **24**, 311-314.

[5] Borodin, O., Kostochka, A., (1977) On an upper bound of a graph's chromatic number depending on the graph's degree and density, *J. Combin. Theory Ser. B* **23**, 247-250.

[6] Breiman, L. (1968), *Probability*, Addison-Wesley.

[7] Brooks, R. L. (1941), On colouring the nodes of a network, *Proc. Cambridge Philos. Soc.* **37**, 194-197.

[8] Catlin, P. (1978), A bound on the chromatic number of a graph, *Discrete Math.* **22**, 81-83.

[9] Erdős, P., Lovász L. (1974) Problems and results on 3-chromatic hypergraphs and some related questions, *Colloq. Math. Soc. Janos Bolyai* **10**, 609-627.

[10] Erdős, P., Rubin, A. L., Taylor, H. (1979), Choosability in graphs, *Proc. West Coast Conf. on Combinatorics, Graph Theory and Computing, Congressus Numerantium* **26**, 125-157.

[11] Frankl, P., Rödl, V. (1985), Near-perfect coverings in graphs and hypergraphs, *Europ. J. Combinatorics* **6**, 317-326.

[12] Füredi, Z. (1988) Mathchings and Covers in Hypergraphs, *Graphs and Combinatorics* **4**, 115-206.

[13] Kahn, J. (1991), Recent results on some not-so-recent hypergraph matching and covering problems, *Proc. 1st Int'l Conference on Extremal Problems for Finite Sets, Visegrád, June*, to appear.

[14] Kahn, J. (1993), Asymptotically good list colorings, submitted to Journal of Combinatorial Theory, Series A.

[15] Kahn, J., Szemerédi, E. (1988), The second eigenvalue of a random regular graph, *manuscript*, (appeared as first chapter in: J. Friedman, J. Kahn and E. Szemerédi, On the second eigenvalue of a random regular graph, Proc. 21st STOC, ACM, 1989).

[16] Komlós, J., Pintz, J., Szemerédi, E. (1982), On Heilbronn's triangle problem, *J. London Math. Soc.* **25**, 13-24.

[17] Kostochka, A., *Letter to B. Toft.*

[18] Lawrence, J. (1978), Covering the vertex set of a graph with subgraphs of smaller degree, *Discrete Math.* **21**, 61-68.

[19] Milman, V., Schechtman, G. (1980), *Asymptotic Theory of Finite Dimensional Normed Spaces*, Springer, Berlin.

[20] N. Pippenger, unpublished.

[21] Pippenger, N., Spencer, J. (1989), Asymptotic behavior of the chromatic index for hypergraphs, *J. Combinatorial Th. (A)* **51**, 24-42.

[22] Rödl, V. (1985), On a packing and covering problem, *Europ. J. Combinatorics* **5**, 69-78.

[23] Shamir, E., Spencer, J. (1987), Sharp concentration of the chromatic number on random graphs $G_{n.p}$, *Combinatorica* **7**, 121-129.

[24] Shearer, J. (1983), A note on the independence number of triangle-free graphs, *Discrete Math.* **46**, 83-87.

[25] Shearer, J. (1991), A note on the independence number of triangle-free graphs, II, *J. Combin. Theory Ser. B* **53**, 300-307.

[26] Spencer, J. (1977), Asymptotic lower bounds for Ramsey functions, *Discrete Math.* **20**, 69-76.

[27] Spencer, J. (1987), *Ten Lectures on the Probabilistic Method*, Society for Industrial and Applied Mathematics.

[28] Turán, P. (1941), Egy gráfelméleti szélsőértékfeladatról, *Mat. Fiz. Lapok* **48** 436-452; see also: On the theory pf graphs, *Colloq. Math.* **3** (1954), 19-30.

[29] Vizing, V. G. (1968), Some unsolved problems in graph theory (Russian), *Uspehi Mat. Nauk.* **23**, 117-134. English Translation in *Russian Math. Surveys* **23**, 125-141.

[30] Vizing V. G. (1967), Coloring the vertices of a graph in prescribed colors (Russian), *Diskret. Analiz. No. 29, Metody Diskret. Anal. v. Teorii Kodov i Shem* **101**, 3-10.