Cited

THE CHOICE NUMBER OF SPARSE GRAPHS

ANDERS JOHANSSON

1. Introduction

1.1. Definitions and notations.

1.1.1. *General.* All graphs considered are, unless stated otherwise, simple and finite. For a graph G, we often write $v \in G$ and $uv \in G$ instead of $v \in V(G)$ and $uv \in E(G)$, when no confusion may result. For a vertex $v \in G$ we write G_v for the graphs spanned by the neighbourhood to v, i.e. $G_v := G[N(v, G)]$.

1.1.2. *Colourings.* For a graph $G = (V, E)$ and a set Γ of "colours", let $\mathcal{P} = V \times \Gamma$ be the set of *pairs*. For $\mathcal{S} \subset \mathcal{P}$ we write $1_{\mathcal{S}}$ for the characteristic function of $\mathcal{S}$ and

$$\mathcal{S}_v = \mathcal{S} \cap \mathcal{P}_v = \mathcal{S} \cap \{v\} \times \Gamma.$$

A colouring σ of G is a map $\sigma : V \to \Gamma$ and a partial colouring is a colouring of an induced subgraph $G[I]$, $I \subset V$ of G. The *"graph"* $\mathcal{G}_\sigma$ of a partial colouring σ is the set of pairs

$$\mathcal{G}_\sigma = \{(v, \gamma) : \sigma(v) = \gamma\}.$$

A colouring σ is *proper* if, for each edge uv, $\sigma(u) \neq \sigma(v)$. The set of proper colourings of a graph G is denoted by Σ_G and the set of proper partial colourings are denoted by Σ'_G.

1.1.3. *The choice number.* If $\mathcal{S} \subset \mathcal{P}$, a (partial) colouring σ of G is said to be *$\mathcal{S}$-legal* if it is *proper*, i.e. $\sigma(u) \neq \sigma(v)$ for all edges $uv \in G$, and

$$\mathcal{G}_\sigma \subset \mathcal{S}.$$

The choice number $\chi_l(G)$ of G is the smallest integer t such that there exists an $\mathcal{S}$-legal colouring whenever

$$|\mathcal{S}_v| \geq t \qquad \forall v \in G.$$

If we restrict attention to the case when $\mathcal{S} = \mathcal{P}$, the same definition gives the *chromatic number* $\chi(G)$ of G.

Let μ denote a probability distribution generating a random *independent* subset S of $V(G)$. Then the *fractional* chromatic number of G is given by

$$\chi^*(G) := \min_\mu \max \{1/\Pr_\mu[v \in S] : v \in G\}.$$

Date. August 28, 1996.

1.1.4. *Asymptotic notation.* We will use standard asymptotic notation. For a relevant limit taken, we write $f = O(g)$ for $\limsup |f/g| < \infty$, $f = \Omega(g)$ for $\liminf |f/g| > 0$ and $f = \Theta(g)$ for $\Omega(g) < f < O(g)$, and also $f = o(g)$ for $|f/g| \to 0$ and $f = \omega(g)$ for $|f/g| \to \infty$. Implicitly, it will often be assumed, that the limit taken will be the maximum degree Δ tending to infinity.

If an asymptotic expression is given for some *local* quantity, say, defined for some vertex in the graph or a pair, it should be understood that the implied constants and convergence-rates are independent of the individual vertex or pair.

1.1.5. *Probabilistic notation and terminology.* For a random variable X defined on some probability space $(\Omega, \mathcal{F}, \Pr[\cdot])$, we write $\mathrm{E}[X]$ for the expectation. For the purposes of this paper it will not matter if we assume that Ω is always *finite.* Moreover, Ω is often implicit in the discussion.

A *filtration* is a sequence of increasing sub-sigma-algebras of $\mathcal{F}$. The algebra *generated* by some set $\{X_1, X_2, \ldots\}$ of random variables is written $\langle X_1, X_2, \ldots \rangle$.

For an event $A \in \mathcal{F}$ we write $\mathrm{I}(A)$ for the corresponding indicator. We also use the notation $\mathrm{M}(A)$ for the *mean-one indicator* of A, defined by

$$\mathrm{M}(A) := \mathrm{I}(A)/\Pr[A],$$

the value being zero when probability is. For a sub-algebra $\mathcal{M}$ we define similarily

$$\mathrm{M}(A \mid \mathcal{M}) := \mathrm{I}(A)/\Pr[A \mid \mathcal{M}].$$

1.2. Present result. The main purpose of this paper is to relate some asymptotic Brook's type bound, i.e. bounds on the choice number in terms of the maximimum degree. The graphs considered will have certain conditions on their first neighbourhood, G_v for $v \in G$. These include K_r-free graphs and graphs with low *local* chromatic number, i.e. the maximum chromatic number of any neighbourhood. We do this by proving a fairly general theorem, using a variant of the nibble method. See [Kah93].

1.3. K_r-free graphs. With a K_r-free graph is understood one that contains no complete graph on r vertices or more as a subgraph.

This paper is closely related to [?] where the main theorem is the following.

Theorem 1. *For a triangle free graphs G the choice number satisfies*

$$\chi_l(G) \leq O\left(\frac{\Delta}{\log \Delta}\right)$$

where Δ denotes the maximum degree.

Constructionsby,e.g. KostochkaandMazurovain[?], showsthattheboundistightwithina multiplicative constant.

The present results on K_r-graphs extends this resultandit is perhapsthemaintheoremofthispaper.

Theorem 2. *There is a $C > 0$ such that if G is a K_r-free graph then the choice number*

$$\chi_l(G) \leq C\frac{\Delta}{\log \Delta}(1 + (r-3)\mathrm{loglog}\Delta)$$

provided the maximum degree Δ is large enough.

Remark 1.1. The value of the constant obtained is of the order 200, although this can certainly be improved.

It should also be pointed out that the bound in Theorem 2, in contrast with Theorem 1, is not known to be tight. Perhaps the $1/\text{loglog}\Delta$ factor can be removed.

The previous best known Brooks'-type bound for the chromatic number of K_r-free graphs is by Borodin and Kostochka [?], Catlin [?] and Lawrence [?] who all independently proved the following theorem.

Theorem 3. *If G is a K_r-free graph where $r \geq 4$ then*

$$\chi(G) \leq \frac{k-1}{k}(\Delta + 2).$$

Let G be a K_r-free graph with *average degree* t. By removing at most $n/3$ vertices we can obtain a graph with maximum degree smaller than $3t$. Hence, Theorem 2 implies the following.

Corollary 4. *For every r for some constant $c_r > 0$*

$$\alpha(G) \geq c_r \frac{n}{t} \frac{\log t}{\text{loglog} t}.$$

for every K_r-free graph G on n vertices with average degree t.

Incidentally, this is the same result as J. Shearer recently obtained in [She95], using quite different (and much slicker) methods. The previously best known bound was,

$$\alpha(G) \geq c_k \frac{n}{t} \text{loglog} t,$$

by Ajtai, Erdős, Komlós and Szemerédi in [AEKS81]. They also conjecture that the $\text{loglog}t$-factor could be replaced by a $\log t$-factor.

1.4. Graphs with small local chromatic number. Triangle free graphs can, besides being graphs with no K_3, also be thought of as graphs where each induced neighbourhood G_v has chromatic number 1. The following theorem is thus a generalization of Theorem 1 in a somewhat different direction.

Theorem 5. *For any G be a graph with the property that*

$$\chi^*(G_v) \leq k \qquad \forall v \in G,$$

then

$$\chi_\ell(G) \leq O\left(\frac{\Delta}{\log \Delta} \log(k+1)\right).$$

Actually, the bound holds under slightly more general assumptions. For a graph H let the number $\chi^{**}(H)$ denote the following relaxation of the fractional chromatic number: Let μ denote a probability distribution generating a random subset S of $V(H)$ satisfying for every edge uv

$$\Pr_\mu[\{u,v\} \subset S] \leq \frac{1}{2}\Pr_\mu[u \in S]\Pr_\mu[v \in S]$$

and let

$$\chi^{**}(H) := \min_\mu \max_v \{1/\Pr_\mu[v \in S]\}.$$

Then obviously, $\chi^*(H) \leq \chi^{**}(H) \leq \chi(H)$, so the previous theorem follows from the one below.

Theorem 6. *Given a graph G. Let $k := \max_v \chi^{**}(G_v)$. Then*

$$\chi_\ell(G) = O\left(\frac{\Delta}{\log \Delta}\log(k+1)\right).$$

Both theorems Theorem 6 and Theorem 2 will follow from a more general theorem: Theorem 7 in section §3. This is proved in section §4. But before stating this we give an heuristic outline of the the method of proof. This has some bearing on the formulation of Theorem 7, which has a rather unpleasing "operational" formulation.

2. An heuristic outline

A "greedy colouring" means constructing an $\mathcal{S}$-legal colouring of a graph by assigning colours to vertices sequentially, and never backtrack to change any assignment. Assuming that the number of colours prescribed by $\mathcal{S}$ to any vertex exceeds the degree, we see that this works, since we always have at least one colour left to assign. This is fundamental argument behind e.g. Brook's theorem. If the graph is *locally sparse* and far from being a complete graph, one may suspect that this worst case analysis could be improved. In particular, randomizing the procedure may help.

2.1. Colouring processes. Given a graph $G = (V, E)$ and a colouring scheme $\mathcal{S}$, it is therefore natural to consider a stochastic process $\sigma^t \in \Sigma'_G$, is a $\mathcal{S}-$*legal* partial colouring of G. We thus assume a probability space $(\Omega, \mathcal{F}, \Pr[\cdot])$ with filtration $\mathcal{F}^t$.

We also assume the process is *greedy* in the sense that, the value of an assigned colour is never changed. That is, for $s > t$, σ^s is $\mathcal{S}$-legal that extension of σ^t. Take σ^0 to be the empty colouring. The set of vertices coloured at time t is denoted by I^t. The graph that remains to be coloured at time t is thus $G^t := G[V \setminus I^t]$. Denote by σ the final colouring σ^∞.

Such a process can be thought of as generalizing the *greedy* colouring algorithm and it is "successful" if it with high probability colours all vertices. In the real proof we will truncate the process at a time T, when the rest of the graph can be coloured greedily.

A base example can be defined as follows in continuous time: For each vertex v pick an exponentially distributed colouring time $t_v \sim Exp(1)$. Choose the value of $\sigma^s(v)$ for $s > t_v$ uniformly at random from the set $\mathcal{S}_v^{t_v}$ of colours not previously assigned to any neighbour of v. That is, from the set

$$\mathcal{S}^t = \mathcal{S} \setminus \left\{(v,\gamma) \in \mathcal{P} : \exists u \sim v, \sigma^t(u) = \gamma\right\}. \tag{2.1}$$

We will refer to this process as the *uniform* greedy process.

Remark 2.1. For the heuristic treatment in this section we will let t be a continuous time parameter. The author feels that using a fake continuous process simplifies the understanding of the overall dynamics. Se Spencer [?] or Kim [?] for more information on related continuous time processes. However, the actual proof will, for reason explained below, be carried out in a discrete time.

Moreover, the heuristic character of the treatment in this section should be stressed: we will naively use concepts not properly defined and properties thereof not necessarily true if they were.

2.2. Intensities and states. We call generally a mapping $p : \mathcal{P} \to [0, \infty]$ a *state* [1]. An important example is the following state

$$q^t(v,\gamma) := \Pr\left[\sigma(v) = \gamma \mid \mathcal{F}^t\right],$$

giving the marginal state of the final colouring at each time.

Furthermore, we assume that the process has a well-defined state of *intensities*, meaning that for all uncoloured vertices $v \in G^t$

$$p^t(v,\gamma) := \lim_{h \downarrow 0} \Pr\left[\sigma^{t+h}(v) = \gamma \mid \mathcal{F}^t\right] \tag{2.2}$$

holds for some state of intensities $\{p^t(v,\gamma) \geq 0 : (v,\gamma) \in \mathcal{P}\}$.

The state of intensities defined in (2.2) above must thus satisfy

$$\operatorname{supp} p^t \subset \mathcal{S}^t.$$

The process is *intensity-governed* if the colouring-process σ^t is *defined* by the relation (2.2) above. As an example: the uniform greedy process defined above may be also defined as the process governed by the intensities

$$p^t(v,\gamma) := 1_{\mathcal{S}^t}/|\mathcal{S}^t_v|, \quad \text{for } v \in G^t, \tag{2.3}$$

whenever $\mathcal{S}^t_v \neq \emptyset$.

We will generally assume the intensity-state p to be almost *normalized* in the sense that for $v \in G^t$ we have

$$p^t(\mathcal{P}_v) := \sum_\gamma p^t(v,\gamma) \approx 1,$$

provided $(\operatorname{supp} p)_v$ is non-empty.

The colouring times

$$t_v := \inf\left\{t : v \in I^t\right\} \cup \{\infty\},$$

are, provided the process is "successful", therefore roughly $Exp(1)$ distributed.

Notation. For a subset $\mathcal{B}$ of $\mathcal{P}$ and a state p we write $p(\mathcal{B})$ for the sum $\sum_{(v,\gamma)\in\mathcal{B}} p(v,\gamma)$. Such sums will loosely be referred to as "mass", since we think of p as assigning weights to the pairs.

We also give $\mathcal{P}$ its natural graph structure, i.e. we let (v,γ_1) be adjacent to (u,γ_2) iff $v \sim u$ and $\gamma_1 = \gamma_2$,i.e. we have $|\Gamma|$ disjoint copies of G. Thus we will sometimes talk about adjacent pairs and the neighbourhood of a pair.

2.3. The martingale property. We also prescribe for the process we consider the following "martingale property" on the state of intensities: For $v \in G^s$

$$\mathrm{E}\left[p^s(v,\gamma) \mid \mathcal{F}^t\right] = p^t(v,\gamma). \tag{2.4}$$

We can call such a process a *martingale* greedy process. Note that this does not include the uniform greedy. The idea behind (2.4) was picked up in the survey paper [Kah90] by Jeff Kahn, where the author gave a preliminary sketch of a proof for the choice number of certain line graphs.

A consequence of (2.4) is that we can rely on *concentration properties* to keep the state sufficently normalized. Thus obtaining "colour independence" for the intensities on the remaining graph.

[1] This is one of many examples where the choice of terms is discutable.

Moreover, assuming the p^t's are defined to have support on all of V instead of only U^t, they may for a martingale process be thought of as the Doob martingale of the marginals, i.e.

$$p^t :\approx \mathrm{E}\big[\, 1_{\mathcal{G}_\sigma} \mid \mathcal{F}^t \,\big],$$

where $\sigma \in \Sigma'$ has the same distribution μ as σ^∞.

Given an increasing random sequence I^t of V, such that $I^0 = \emptyset$ and $I^\infty = V$, and assuming that $\mathcal{F}^t$ is being generated of what has "happened" on the "seen" vertex-set I^t, i.e.

$$\mathcal{F}^t \approx \langle \sigma^t \rangle := \langle \sigma|_{I^t} \rangle,$$

where $\sigma \sim \mu^2$, one may in general think of a martingale process as giving the conditional expectations of such a vertex exposure filtering.

It is also worth noting that this filtering I^t of the vertex-set V can be made in different ways. If the process is governed by a normalized state of intensities, each vertex move to the set I^t with intensity one, making the colour-times t_v exponentially $Exp(1)$ distributed.

But there is nothing sacred about this construction. For instance when colouring the edges of a multigraph: the great idea in [?] is to use a "hard-core" measure on the set of matchings to decide the structure of $I^{t+dt} \setminus I^t$ [3].

2.4. Some important statistics. For the type of processes we consider, we now introduce some time dependent statistics that are particularly important.

2.4.1. *The degrees.* First of all we study the degrees $d(v, G^t)$. Since vertices are coloured with intensity $p^t(\mathcal{P}_v) \approx 1$ we have that

$$\frac{d^{\mathrm{E}}}{dt} d(v, G^t) \approx -d(v, G^t).$$

where we use the notation

$$\frac{d^{\mathrm{E}}}{dt} X^t := \lim_{h \downarrow 0} \frac{1}{h} \left(\mathrm{E}\big[\, X^{t+h} \mid \mathcal{F}^t \,\big] - X^t \right)$$

Hence, assuming the variance of $d(v, G^t)$ is negligible, the degree should decrease exponentially

$$d(v, G^t) \approx e^{-t} d(v, G^0) \leq e^{-t} \Delta. \tag{2.5}$$

Remark 2.2. In general, since the number of colours will be large and the process will run fairly independently between different colours, statistics that are not defined by an individual colour, are *concentrated.* That is, at time $t + dt$, they are with high probability within a factor $(1 + o(dt))$ of the value expected at time t.

In the heuristic discussion we therefore substitute "expected behaviour" (Which in a strict sense often is not the expected behaviour at all.) with actual behaviour. Of course, the real problem is to prove these statements for all vertices. To this end we will use the "nibble method". (See §2.8 below.)

---

[2] In the modified model we will use to proveTheorem 7 we add, for reasons explained below, moreover some elementary events apart from the colouring events. But the general argument with a vertex-filtering can still be made.

[3] The proof in [?] has in other respects different characteristics than outlined here, basically it is a "uniform" model, but I think it should be possible to adapt the basic idea — using hard-core measures to construct the nibble — to the setting described in this paper.

$$\frac{d}{dt}\left(x \log x\right) = x'(\log x + 1)$$

2.4.2. *The entropy.* Moreover, we define for each $v \in G$ the *local entropy* of p^t by

$$h(v, p^t) := \sum_\gamma -p^t(v,\gamma) \log p^t(v,\gamma). \tag{2.6}$$

The importance of this statistic is seen from

$$h(v, p^t) \le \log |\mathcal{S}_v^t|. \tag{2.7}$$

Hence any lower bound on the entropy gives a lower bound on the number of available colours. It will be more convenient to work with the local entropy than the sizes of the lists.

2.4.3. *The energy.* Let $\mathcal{P}^t$ denote $V(G^t) \times \Gamma$ and denote by $N^t(v,\gamma)$ the neighbourhood in $\mathcal{P}^t$ of (v,γ), i.e. $N^t(v,\gamma) := \{(u,\gamma) : u \in N(v, G^t)\}$.

Another important quantity will be the *local energy*:

$$e(v, p^t) := \sum_{u \sim v} \sum_\gamma p^t(v,\gamma) p^t(u,\gamma) = \sum_\gamma p^t(v,\gamma) p^t(N^t_{(v,\gamma)}). \tag{2.8}$$

This give the expected rate with which the mass $p^t(\mathcal{P}_v)$ should decrease were it not adjusted to stay normalized.

We write also for $uv \in G$

$$e(uv, p^t) := \sum_\gamma p^t(v,\gamma) p^t(u,\gamma),$$

for the edge-wise energy. Generally, when talking about the "energy" in a subgraph of G we mean the edge-wise energy summed over the edges in this subgraph.

Remark 2.3. "Energy" is perhaps an unhappy choice of words. Less overused terms would pehaps be "clash" or "tension". However, the choice is not entirely out of the blue: If we regard, in the obvious way, the state p^t as a probability distribution on the set of colourings of G^t, then the "energy", as defined above, corresponds to the natural energy functional (or "cost") for such distributions.

2.5. Energy as the rate of entropy loss. The importance of the local energy stems from the fact that, in general,

$$\frac{d^{\mathrm{E}}}{dt} h(v, p^t) \ge e(v, p^t), \qquad ? \leqslant -$$

since $p(v,\gamma)$ must be set to zero whenever some neighbour $u \sim v$ get assigned the colour γ.

Assuming the reverse inequality holds within some factor C independent of t, we have in expectation

$$h(v, p^t) \ge h(v, p^0) - C \int_0^t e(v, p^s) ds. \tag{2.9}$$

This give us a way to bound the *loss in entropy* — the loss of colours available — by bounding the energy $e(v, p^t)$.

In particular, if for some T

$$h(v, p^0) - C \int_0^T e(v, p^s) ds \ge \log d(v, G^T), \tag{2.10}$$

we expect the number of available colours to exceed the degree at each vertex. The process may then be terminated, since the rest of the graph can be coloured greedily.

2.6. A construction for triangle free graphs. Perhaps, the most likely choice of a martingale greedy process is the following: Define, heuristically,

$$p^{t+dt}(v,\gamma) := p^t(v,\gamma)\cdot \mathrm{M}\big((v,\gamma)\in S^{t+dt}\mid \mathcal{F}^t\big) \tag{2.11}$$

where, as above, for an event A

$$\mathrm{M}(A\mid\mathcal{F}^t) := \mathrm{I}(A)/\Pr\big[A\mid\mathcal{F}^t\big]\,.$$

This means we "kill" a pair (v,γ) only if a neighbour becomes coloured, and adjust the weight to compensate for this event. It follows then that we obtain the bound in (2.9) with $C\approx 1$.

In the case of a triangle free graph the above procedure works quite well and is the method used in [?]. The special useful property of triangle free graphs is the following: Since the neighbourhoods of any two adjacent vertices u and v are disjoint the values $p^{t+dt}(v,\gamma)$ and $p^{t+dt}(u,\gamma)$ are, given $\mathcal{F}^t$, almost independent. Thus, for $uv\in G^{t+dt}$,

$$\mathrm{E}\big[\,e(uv,p^{t+dt})\mid\mathcal{F}^t\,\big]\approx e(uv,p^t). \tag{2.12}$$

We will loose a negligible amount by conditioning on $uv\in G^{t+dt}$.

At the same time, for any vertex $v\in G^{t+dt}$, a small fraction ($\approx dt$) of the neighbours will disappear since they become coloured. Therefore

$$\frac{d^{\mathrm{E}}}{dt}e(v,p^t)\approx -e(v,p^t) \tag{2.13}$$

thus giving an exponential decrease in the energy:

$$\mathrm{E}\big[\,e(v,p^t)\big]\approx e^{-t}e(v,p^0). \tag{2.14}$$

We assume the initial energy and entropy have some suitable values, i.e. if the initial lists satisfy

$$|\mathcal{S}_v|\geq \Delta/K\quad\text{where}\quad K=k\log\Delta,$$

we have $e(v,p^0)=K=k\log\Delta$. For a small choice of k the integral in (2.9) stays sufficiently small ($\leq K$) and we obtain (2.10) by choosing T large enough.

2.7. The modified approach. However, to deal with graphs more general than triangle free, the approach must be modified somewhat. Since (2.12) does not hold, one must find some way to counteract the potentially superexponential increase of the energy that may result. This would otherwise effectively "kill" the process.

The modified approach is roughly the following: Instead of colouring $\sigma^t(v)=\gamma$ with intensity $p^t(v,\gamma)$ we say that the pair (v,γ) gets *excited* with intensity $p^t(v,\gamma)$. Then, when a pair (v,γ) gets excited, we throw a throw a fair coin. If heads come up, we colour v with colour γ and, for all pairs (u,γ) with $u\sim v$, we set $p^{t+dt}(u,\gamma)=0$. If instead tails come up: we set for $u\sim v$

$$p^{t+dt}(u,\gamma)=2p^t(u,\gamma)M_v(u,\gamma)$$

where $M_v(u,\gamma)$, for $u\sim v$ is a random variable having expectation 1. This implies (2.4). The exponential rate of decrease in the degree is now $\approx 1/2$, but the overall argument stays unchanged.

The random vector $M_v(\cdot,\gamma)$ will be called a "modifier". The concept of such local modifiers will be used in the formulation of Theorem 7 below. The modifiers are constructed so that they counteract the increase of the energy. For instance: If each neighbourhood has chromatic number k we can define M as follows. Choose at random, among the k given, a colour class $V_\alpha \subset N(v,G^t)$ and set

$$M_v(u,\gamma) = \mathbf{I}(u \in V_\alpha)\,/\,\mathrm{Pr}[u \in V_\alpha]\,.$$

By construction, $M_v(\cdot,\gamma)$ thus shuffles all the mass $p(N(v,\gamma))$ to the independent set V_α. This clearly means that the energy of the neighbourhood $N(v,G^t)$ will not increase as a result of (v,γ) being excited.

However, using such local modifiers M implies that the decrease of the local entropy is enhanced by a factor $C \approx \log 2 + \kappa$ where

$$\kappa = \max_{v,u} \mathrm{E}[\,M_v(u,\gamma)\log M_v(u,\gamma)\,]$$

then res the "kindness" of M But, using sufficently large values for the initial lists, $(k = \Omega(1/C))$, the integral in (2.9) will still be small enough.

2.8. The nibble method. However, since we want the constructed colouring to be good at all corners of a possibly huge graph, we will not use a continuous process at all. It is possible that a modified version of the continuous process described above could work for moderately sized graphs. But in the general case we must use the Local Lemma, or some other sieve, to "rectify" the process globally. That is, making the expected evolution the norm for *all* vertices.

To this end, we look instead at a small time-interval $[t,t+\theta)$, $t = i\theta$, where θ is sufficiently small so that the process within this time interval can be approximated by a probability space Ω^t — the "nibble space" — which is made up of a set of mutually independent variables.

At the same time the time slice θ is sufficiently big so that large deviation results can be used together with the Local Lemma to choose some outcome where the relevant quantities stay close to their expected value globally: That is, we will maintain inequalities corresponding to (2.5), (2.9) and (2.14) above. This way of going from a continuous process to a discrete version in order to obtain good global behaviour can be thought of as using the "nibble method".

In general, the nibble method is a way to iteratively extend a partial structure — a covering, a colouring, and independent set e.t.c — by using a random construction with sufficently small steps. The main iteration is often followed by a greedy step to take care of the rest. For a survey: see [Kah90]. Although the method had been used before, the term nibble (or "Rödl nibble") was introduced to describe the central idea in the proof [?] by V. Rödl of the Erdős-Hanani conjecture. The method, has since been used with great success on many hard combinatorial problems.

2.9. The truncation. We also take care to "truncate" the state p so that pairs are active only if $p(v,\gamma) < \hat{p}$ for some small number $\hat{p} = \Delta^{-\Omega(1)}$. In this way the relevant quantities will have "bounded differences" relative the natural filtration of Ω^t. Standard large deviation results, e.g. Bernstein's inequality [?], then implies sufficient concentration, to allow the use of the Local Lemma.

The reason this truncation scheme works is that the bounds on the entropy given by (2.9) above implies that the pairs $\mathcal{I} \subset \mathcal{P}$ discarded by the truncation is a relatively small set with small local mass $p(\mathcal{I}_v) < \varepsilon$, for all $v \in G$.

3. A MORE GENERAL THEOREM

In this section we state Theorem 7 and give the proofs that this theorem implies Theorem 2 and Theorem 6. But, first we define some crucial concepts used in the proof.

3.1. Mean one variables. For our purposes, a *mean one* variable M is a non-negative random variable with expectation $\mathrm{E}[M] = 1$ and such that $M = 0$ or $M \geq 1$. We will use mean one variables to "modify" the weights $p(v,\gamma)$, in order to keep the "energy" low. The following identity will give the corresponding loss of "entropy": For $p > 0$ and a mean one variable M we clearly have

$$\mathrm{E}[(pM)\log(pM)] = p\log p + p\kappa(M). \tag{3.1}$$

where

$$\kappa(M) := \mathrm{E}[M\log M] \leq \log \mathrm{E}[M^2].$$

Note that for an event A we have

$$\kappa(\mathrm{M}(A)) = \log(1/\Pr[A])$$

and for independent m.o. variables M and N we have

$$\kappa(MN) = \kappa(M) + \kappa(N).$$

3.2. The stopped product. In the proof we will crucially use the following construction. Let Y_i, $i = 1,\dots,$ be a set of independent m.o. variables. Let X_i, $i = 0,\dots,n$ be the martingale defined by $X_0 = 1$ and

$$X_i := \prod_{j=1}^{i} Y_j, \quad i \geq 1.$$

For a given threshold $a > 0$ define the *stopped product* $\hat{X}_a$ by

$$\hat{X}_a := X_{n\wedge\tau_a} \tag{3.2}$$

where τ_a is the stopping time

$$\tau_a = \min\{i : X_i \geq a\} \cup \{\infty\}.$$

Then $\|\hat{X}_a\|_\infty \leq \hat{\kappa}\cdot a$ where $\hat{\kappa} = \max_i \|X_i\|_\infty$.

Furthermore, $\kappa(\hat{X}_a) = 0$ if $a \leq 1$ and it is also easy to see that

$$\kappa(\hat{X}_a) \leq \sum_{i=1}^{n} \kappa(Y_i). \tag{3.3}$$

3.3. Modifiers. Given a graph $G = (V,E)$. A vector $N = (N(v) : v \in G)$, such that each $N(v)$ is a mean one variable, is called a *modifier* for G. Set

$$\kappa(N) := \max_v \kappa(N(v))$$

and denote by $\hat{\kappa}(M)$ the maximal outcome of M, i.e.

$$\hat{\kappa}(M) := \max_v \|M(v)\|_\infty.$$

Given a modifier M of G we say that M is *contracting* if for all edges $uv \in G$

$$\mathrm{E}[M(u)M(v)] \leq \frac{1}{2}. \tag{3.4}$$

Given a vertex weighting $p : V \to [0, \infty]$ and a $c > 0$. We say that M is (p, c)-*trimming* if for every $v \in G$

$$\sum_{u \sim v} p(v)\, p(u)\, (2\mathrm{E}[M(v)M(u)] - 1) \leq c \cdot p(v). \tag{3.5}$$

A contracting modifier is thus clearly (p, c)-trimming for any p and c.

3.4. The general theorem. We may now state following theorem which is the central theorem of this paper, giving an asymptotic Brook's bound for graphs with a local "trimming" property.

Theorem 7. *Let G be a graph with the following property: For each weighting p of G such that for all $v \in G$*

$$p(G_w) \leq 100 \log \Delta \quad \textit{and} \quad \max_{v \in G_w} p(v) < \Delta^{-3/4},$$

one can construct a $(p, 1/4)$-trimming modifier $N_w(\cdot)$ on G_w with

$$\kappa(N_w) \leq \kappa \quad \textit{and} \quad \hat{\kappa}(N_w) \leq \hat{\kappa} \leq \Delta^{1/100}$$

Then

$$\chi_\ell(G) = O\,(\Delta / \log \Delta)(\log 2 + \kappa).$$

Remark 3.1. The value of the implicit constant obtained is approximately 800/3. This could certainly be improved by a factor 10 if one was more careful.

Remark 3.2. The use of the number 100 at two places is of course ad hoc: Any pair of sufficiently large numbers would do. Also, the trimming constant 1/4 is quite arbitrarily chosen, it is enough to ask for a c strictly smaller than 1/2. Besides, the choice of c is obviously related to the bounds on p.

Remark 3.3. The theorem can be generalized along the lines indicated for the main theorem in [?]. In particular, we can start with smaller lists provided these have sufficiently small intersection for edge-wise pairs.

Remark 3.4. To make the proof go through it is actually enough to ask that, unless it is zero, $N_w(\cdot, \gamma) \geq 1/2$ instead of 1.

The following corollary is perhaps more pleasing than Theorem 7.

Corollary 8. *Let G be a graph with the following property: For each w one can construct a contracting modifier N_w such that*

$$\kappa(N_w) \leq \kappa \quad \textit{and} \quad \hat{\kappa}(N_w) \leq \hat{\kappa} \leq \Delta^{1/100}$$

Then

$$\chi_\ell(G) = O\,(\Delta / \log \Delta)(\log 2 + \kappa).$$

Remark 3.5. The proof of Theorem 7 can be slightly simplified and the constant considerably improved if we restrict the attention to contracting modifiers.

The stated theorems implies both Theorem 6 and Theorem 2 and the rest of this section is devoted to showing these implications. The first implication is immediate from Theorem 8: By assumption we have a random subset $S \subset V(G_w)$ such that for any $uv \in G_w$ the probability $\Pr[uv \in G_w[S]]$ is at most one half of $\Pr[u \in S] \Pr[v \in S]$ and such that $\max_{u \in G_w} \{1/\Pr[u \in S]\}$ is less than k. But these assumptions states $N_w(v) := \mathrm{M}(v \in S)$ is a contracting modifier with κ equal to $\log k$, and $\hat{\kappa}$ equal to k, where k may be assumed to be smaller than $\Delta^{1/100}$.

3.5. Proof that Theorem 7 implies Theorem 2. The construction is here a bit more intricate. The idea is to randomly partition the vertices of each neighbourhood into K_{r-2}-free graphs and a rest-part and then, by a random choice of one of the parts, "shuffle" all the mass into the part chosen. If that part is one of the K_{r-2}-free graph we obtain the trimming property by induction on r. The vertices in the rest-part will by construction have neighbours with low total weight, giving the trimming property for all vertices.

For simplicity, let (G,p) denote one of the weighted induced neighbourhoods in a K_{r+1} free graph. Hence $G=(V,E)$ is K_r-free and has at most Δ vertices. Let $1>c>0$ and assume that the weighting p satisfies:

$$p(V) := \sum_{v\in V} p(v) = O(\log \Delta). \tag{3.6}$$

and $\forall\, v \in V$

$$p(v) \le (256/c)^{-r^2} (p(V))^{-r}. \tag{3.7}$$

For the given p we shall define a modifier M on G with the property:

$$\sum_{u\sim v} 2\, p(v)p(u)\mathrm{E}[\, M(v)M(u)\,] \le cp(v). \tag{3.8}$$

This property is thus somewhat stronger than the (p,c)-trimming property.

The constructed M will satisfy the following bounds

$$\kappa(M) \le \kappa_r(t,c) := (r-2)\log t + (r-2)^2 \log(256/c) \tag{3.9}$$

and

$$\hat{\kappa}(M) < \hat{\kappa}_r(t,c) = (256/c)^{r^2} t^r. \tag{3.10}$$

where $t := p(V)$. One can easily check that this is enough for Theorem 2. Inductively, we may assume that, for all c, such an M can be defined for all weighted K_{r-1}-free graphs satisfying (3.6) and (3.7) above. (For $r=2$ there is nothing to prove: we can set $M(v)\equiv 1$.)

As the first step: for each set $X=\{x_1,x_2,\dots,x_m\}\subset V$ with $m<32t/c$ we construct a modifier M_X as follows. Let $V_0 := V\setminus N(X)$ and partition $N(X)$ into subset $V_1,\dots,V_m$ such that $V_i\subset N(x_i)$. Thus each V_i, $i\ge 1$, spans a K_{r-1}-free graph. Let $A := \{i\in[1,m] : p(V_i)>c\}$ and let $B := [1,m]\setminus A$.

Define p' by setting $p' = p(v)/w_i$ on V_i, where

$$w_i := \begin{cases} 1/2 & \text{if } i=0 \\ p(V_i)/4\sum_A p(V_i) & \text{if } i\in A \\ 1/4|B| & \text{if } i\in B. \end{cases}$$

for $i=0,\dots,m$. Note that, since $m<32t/c$ and $\sum_{i\in A} p(V_i)\le t$ the construction implies that

$$w_i > c/128t \quad\text{and}\quad p'(V_i) < 128t. \tag{3.11}$$

Let $M_0'(v)\equiv 1$ and for each $i\ge 1$ we can , by the induction hypothesis, choose a modifier M_i' on $G[V_i]$ that satisfies the stated properties for $c=c/2$ and p'. Pick at

random an index α among $[0, m]$ according to the probabilities $w_0, \ldots, w_m$ defined above and let the modifier M_X be defined by

$$M_X(v) := \mathrm{M}(v \in V_\alpha)\, M'_\alpha(v)$$

From (3.11) it follows that

$$\begin{aligned}\kappa(M_X) &= \max_i\{\log(1/w_i) + \kappa(M'_i)\}\\ &\le \log(128t/c) + \kappa_{r-1}(128t, c/2)\\ &< (r-2)\log t + (r-2)^2\log(256/c)\end{aligned}$$

and similarily that

$$\hat{\kappa}(M_X) \le \max_i\{\hat{\kappa}(M'_i)\cdot(1/w_i)\} < (256/c)^{r^2} t^r.$$

Furthermore, for $v \in V_i$, where $i \neq 0$, i.e. if $v \in N(X)$, the modifier M_X satisfies

$$\sum_{u\sim v} 2\, p(v)p(u)\mathrm{E}[\, M_X(v)M_X(u)\,] \le (c/2)p(v) \tag{3.12}$$

since, the left hand side equals

$$w_i \sum_{\substack{u\sim v\\ u\in V_i}} 2p'(v)p'(u)\mathrm{E}[\, M'_i(v)M'_i(u)\,] \le w_i(c/2)p'(v) = (c/2)p(v).$$

Choose finally Y' as a random subset by picking $v \in V$ independently at random with probability $8p(v)/c$ — this is smaller than 1 by (3.7). Define the random variable Y as Y' *conditioned* on the event that $|Y'| < 32t/c$, i.e. for $|X| < 32t/c$

$$\Pr[Y = X\,] = \Pr[Y' = X\,]\,/\Pr[\,|Y'| < 32t/c\,].$$

With the random subset Y as above set finally $M = M_Y$ as the modifier for G. By the construction above, we have that M_Y satisfies the stated bounds (3.9) and (3.10) for each Y of positive probability. Thus, it only remains to verify that M satisfies (3.8).

We have by (3.12) that

$$\sum_{u\sim v} 2\, p(v)p(u)\mathrm{E}[\, M(v)M(u) \mid v \in N(Y)\,] \le (c/2)p(v).$$

On the other hand

$$\mathrm{E}[\, M(v)M(u) \mid v \notin N(Y)\,] = 2\Pr[\, u \notin N(Y) \mid v \notin N(Y)\,]$$

and therefore it is enough to show

$$4p(v)m(v)\Pr[\, v \notin N(Y)\,] \le \frac{c}{2}p(v), \tag{3.13}$$

where $m(v) = \sum_{u\sim v} p(u)$.

We find that

$$\Pr[\, v \notin N(Y)\,] \le e^{-(8/c)m(v)}/(3/4)$$

since $\Pr[\,|Y'| \le 32t/c\,] \ge 3/4$ by Chebyshev's inequality. Hence,

$$4p(v)m(v)\Pr[\, v \notin N(Y)\,] \le 4p(v)\cdot(c/8)xe^{-x}\frac{4}{3}.$$

where $x = (8/c)m(u)$. But $e^{-x}x < 1/e$ giving (3.13). □

Remark 3.6. The given construction could certainly be improved and simplified. It would be interesting to know if one can do considerably better by constructing a trimming modifier instead of one satisfying the stronger property (3.8).

4. The proof of Theorem 7

4.1. General outline. We will now give the proof. It is divided into three sections, where this first section contains the overall structure up to the main lemma: Lemma 9. The proof of this is the content of the two following sections. The main point being the argument in §5 bounding the expected energy.

4.1.1. *Parameters.* The proof will follow the heuristic outline given above: We start with a graph G^0 satisfying the conditions of the theorem, and a state p^0 defined by

$$p^0(v,\gamma) := 1_{\mathcal{S}}(v,\gamma)/|\mathcal{S}_v|,$$

where $\mathcal{S} \subset \mathcal{P}$ is the supplied colouring scheme. We tacitly assume that Δ is large enough to allow the estimates to be carried out.

Let $\varepsilon = 1/100$ and let

$$\hat{p} := \Delta^{-\alpha}, \quad \text{where } \alpha = 3/4 + 5\varepsilon.$$

Then, let $c = 1/4$ be the trimming constant from the theorem. Let, $a = 1/2 - 3\varepsilon$ and $b = a - c$. These two numbers will give the exponential rate of decrease for the degree and the energy, respectively. Finally, set

$$K = \frac{(1-\varepsilon)^2\varepsilon(1-\alpha)b}{C}\log\Delta, \tag{4.1}$$

where

$$C := \log 2 + \kappa.$$

By assumption we have that $|\mathcal{S}_v| > \Delta/K$ for all $v \in G^0$. Hence,

$$e(v,p^0) \le K \quad \text{and} \quad h(v,p^0) \ge \log\Delta - \log K. \tag{4.2}$$

Remark 4.1. It will be apparent that we make no real effort to compute the best possible constants. Besides, the constant $\varepsilon = 1/100$ can be thought of as a generic sufficiently small parameter.

4.1.2. *The nibble.* We let θ be given by

$$\theta := \frac{\Delta^{-1-\varepsilon}}{\hat{\kappa}^2\hat{p}} \ge \Delta - 1/4 + 2\varepsilon. \tag{4.3}$$

and for $t = i\theta$ $i = 0, 1, \dots$ we will construct a probability space Ω^t, that carries a random proper partial colouring σ' defined on a subset $I' \subset V(G^t)$. All probabilistic notation will refer to the space Ω^t unless otherwise stated. Let $G' := G^t[V(G^t)\setminus I']$ be the remaining graph. The main point in the proof is to construct the residual state p', which for convenience has support on G^t instead of G'.

We use the Local Lemma to get a *good outcome* in Ω^t and for this outcome we set $\sigma^{t+\theta} = \sigma^t \cup \sigma'$ and

$$G^{t+\theta} := G' \quad \text{and} \quad p^{t+\theta} := p'|_{G'},$$

where

$$p|_{G'} := p \cdot 1_{\{(v,\gamma):\, v\in G'\}}.$$

We iterate this construction up to some suitable time $t = T$, when we stop and colour the remaining graph greedily.

Notation. For convenience we will drop time indices when referring to the "present" time t. That is, let $p = p^t$ and $G = G^t = (V, E)$. When referring to vertices we will mostly use the letters v, u, w, and, if nothing else is said, these symbols refer to some arbitrary vertex in $G = G^t$. We will also implicitly assume that γ always stands for some arbitrary colour in Γ.

4.2. The active state. We will not use the whole state p: Some exceptional pairs will be excluded from consideration. Given the state p, define the set

$$\mathcal{I} := \{(v,\gamma) : p(v,\gamma) > \hat{p}\}.$$

These pairs will be permanently discarded. Define the *active* state p_a and the *excitable state* p_c by

$$p_a := p \cdot 1_{\mathcal{P}\setminus\mathcal{I}} \quad \text{and} \quad p_c := p \cdot 1_{\mathcal{P}\setminus\mathcal{I}\setminus\mathcal{J}}.$$

where

$$\mathcal{J} := \{(v,\gamma) : p(N_{(v,\gamma)}) \geq (1/\varepsilon)K\}.$$

Thus we "temporarily" deactivate the pairs in $\mathcal{J}$ where the neighbourhood mass is to big to allow the modifiers to be defined. We will see that these two states are almost equal to p: The exceptional pairs in $\mathcal{I}$ and $\mathcal{J}$ will be rare and not carry much mass.

4.3. The excited set and the colouring. The main properties of the probability space Ω^t are given by the following construction: Define a sparse random subset of *excited* atoms $\mathcal{E}' \subset \mathcal{P}$ as follows: For $(v,\gamma) \in \mathcal{P}$ let

$$\epsilon'(v,\gamma) := \mathbf{I}((v,\gamma) \in \mathcal{E}') \sim B(\theta p_c(v,\gamma))$$

be a set of independent Bernouilli variables each with expectation $\theta p_c(v,\gamma)$.

We assume that V is *ordered* and, for $\mathcal{B} \subset \mathcal{P}$, let $N^-(\mathcal{B})$ ($N^+(\mathcal{B})$) denote the set of pairs adjacent to and preceding (succeding) some pair in $\mathcal{B}$. Define the proper part $\tilde{\mathcal{E}}'$ of $\mathcal{E}'$ by setting

$$\tilde{\mathcal{E}}' := \mathcal{E}' \setminus N^-(\mathcal{E}').$$

That is, remove all excited pairs that have some *preceding* neighbour that is also excited.

Moreover, for each pair $(v,\gamma) \in \mathcal{P}$: throw a fair coin $\eta'(v,\gamma) \sim B(1/2)$ and construct the "colouring" $\mathcal{C}' \subset \tilde{\mathcal{E}}'$ by

$$\mathbf{I}((v,\gamma) \in \mathcal{C}') := \mathbf{I}\Big((v,\gamma) \in \tilde{\mathcal{E}}'\Big) \cdot \eta'(v,\gamma).$$

If $\mathcal{C}'_v$ is non-empty we can take any element of $\mathcal{C}'_v$ for $\sigma'(v)$. Thus

$$\mathbf{I}(v \in G') = \mathbf{I}(\mathcal{C}'_v = \emptyset).$$

4.4. The residual state. We now give the construction of the residual state p'. The central idea is to define, for each γ, a modifier $M'(\cdot,\gamma)$ on G and then set

$$p'(v,\gamma) := p(v,\gamma) \cdot M'(v,\gamma) \tag{4.4}$$

Crucially, we then get the martingale property

$$\mathbf{E}[p'(v,\gamma)] = p(v,\gamma). \tag{4.5}$$

The construction will imply that all pairs (v,γ) which are neighbours to a pair $(w,\gamma) \in \mathcal{C}'$ will be "killed", meaning that the value of the new active state $p'_a(v,\gamma)$ will be zero. Here,

$$p'_a := p' \cdot 1_{\mathcal{P}\setminus\{(v,\gamma):p'(v,\gamma)>\hat{p}\}}.$$

Since, $M(v,\gamma) \geq 1$ unless it is zero, we have that $p'(v,\gamma) \geq p(v,\gamma)$, whenever $p'vg > 0$. This implies that the sequential colouring stays proper, since pairs in $\mathcal{I}$ will never become active again.

To define M' note that for $p_c(w,\gamma) > 0$ we have

$$\sum_{u \in G_w} p_a(u,\gamma) \leq (1/\varepsilon)K \leq (1/\varepsilon)\log\Delta.$$

Hence, the assumptions of the theorem allow us to define for each (w,γ) a modifier $N_w(\cdot,\gamma)$ on G_w which is $(p_a(\cdot,\gamma),c)$-trimming. Each $N_w(\cdot,\gamma)$ being an "independent copy".

Then, for each pair (w,γ) with $p_c(w,\gamma) > 0$, define $M'_w(\cdot,\gamma)$ by

$$M'_w(v,\gamma) = \begin{cases} 1 & \text{if } (w,\gamma) \notin \mathcal{E}' \\ \mathrm{M}(\eta'(w,\gamma) = 0) \cdot N_w(v,\gamma) & \text{if } (w,\gamma) \in \mathcal{E}'. \end{cases} \tag{4.6}$$

Finally, we define $M'(v,\gamma)$ as the stopped product (See (3.2)) of the $M'_w(v,\gamma)$ for $w \sim v$ with respect to the threshold $\hat{p}/p(v,\gamma)$. (The ordering used is arbitrary.) This implies that the value of $M'(v,\gamma)$ never exceeds $\hat{p}/p(v,\gamma)$ with more than a factor of $2\hat{\kappa}$. Therefore we have the following bounds

$$p'(v,\gamma) \leq 2\hat{\kappa}\hat{p} \quad \text{and} \quad p'_a(v,\gamma) \leq \hat{p}. \tag{4.7}$$

4.5. Independence properties. The probability space Ω^t is thus constructed from the set

$$\{\epsilon'(v,\gamma), \eta'(v,\gamma), N_v(\cdot,\gamma) : (v,\gamma) \in \mathcal{P}\}$$

of mutually independent random variables. (We look upon $N_w(\cdot,\gamma)$ as a vector-valued random variable.)

From the construction it clear that the random variable $d(v,G')$ is determined by the outcome of $\epsilon'(w,\gamma)$ and $\eta'(w,\gamma)$ with w within distance 2 of v. Similarily, for fixt v and γ we have that $p'(v,\gamma)$ is determined by $\epsilon'(w,\gamma)$, $\eta'(w,\gamma)$ and $N'_w(v,\gamma)$ with w is at distance 1 from v.

A consequence is that $p'(v,\gamma_1)$ and $p'(u,\gamma_2)$ are independent random variables whenever

(1) $\gamma_1 \neq \gamma_2$ or
(2) $v \neq u$ and $N(u,G) \cap N(v,G) = \emptyset$.

This give us enough independence to carry out the proof of the concentration properties in §6 relatively easy.

4.6. Behaviour of local quantities. The main point is now to maintain the following inequalities bounding the evolution of the crucial parameters. Basically, the inequalities state that the quantities should not deviate to much in the wrong

direction from their expected values:

(N) $$p^t(\mathcal{P}_v) = 1 + \Delta^{-\Omega(1)},$$

(D) $$d(v, G^t) < \Delta e^{-at},$$

(E) $$e(v, p_a^t) < e^{-bt}K,$$

(H) $$h(v, p^t) > h(v, p^0) - (1 + o(1)) \cdot C \cdot K \int_0^t e^{-bs} ds.$$

To see where this leads us, note first that (H) and (4.1) implies that

$$\begin{aligned} h(v, p^t) &> h(v, p^0) - \varepsilon(1-\varepsilon)(1-\alpha)\log\Delta \\ &> [1 - \varepsilon(1-\varepsilon/2)(1-\alpha)]\log\Delta. \end{aligned} \tag{4.8}$$

since $h(v, p^0) = (1 - o(1))\log\Delta$ and the integral in (H) is smaller than $1/b$.

4.7. The active mass. A consequence of (4.8) is then that

$$p(\mathcal{I}_v) < \varepsilon, \tag{4.9}$$

i.e. almost all mass is active at each vertex.

The reason is the following: We can assume that $p(v, \gamma) > 1/\Delta$ unless $p(v, \gamma) = 0$. From (N) and (4.8) we then get

$$\begin{aligned} p(\mathcal{I}_v) \cdot \min\{\log(\Delta \cdot p(v, \gamma)) : (v, \gamma) \in \mathcal{I}\} &\le \sum_\gamma log(\Delta \cdot p(v, \gamma)) \\ = (1 \pm \Delta^{-\Omega(1)})(\log\Delta - h(v, p)) &\le \varepsilon(1-\varepsilon/2)(1-\alpha)\log\Delta \end{aligned}$$

But, by definition of $\mathcal{I}$, the factor $\min\{\log(\Delta \cdot p(v, \gamma)) : (v, \gamma) \in \mathcal{I}\}$ is greater than $(1-\alpha)\log\Delta$, implying (4.9).

In the same manner, we can deduce from (E) that

$$p_a(\mathcal{J}_v) < e^{-bt}\varepsilon$$

since

$$e^{-bt}K \ge e(v, p_a) \ge \sum_{g \in \mathcal{J}_v} p_a(v, \gamma) p(N_{(v,\gamma)}) > p_a(\mathcal{J}_v) \cdot (1/\varepsilon)K.$$

Hence, for all $v \in G$

$$p_c(\mathcal{P}_v) = p(\mathcal{P}_v) - p(\mathcal{I}_v) - p_a(\mathcal{J}_v) > 1 - 2\varepsilon. \tag{4.10}$$

4.8. The termination. From (4.9) and (4.8) we may then deduce that the logarithm of the number of available colours at a vertex v exceeds

$$\begin{aligned} \log|(\operatorname{supp} p_a)_v| &\ge (1-\varepsilon)\log|(\operatorname{supp} p)_v| \\ &\ge (1-\varepsilon)(1-\varepsilon(1-\alpha))\log\Delta > (1-\varepsilon)\log\Delta \end{aligned}$$

since by removing the pairs (v, γ) in $\mathcal{I}$ we remove the pairs of highest weight $p(v, \gamma)$ and thus, in numbers, a smaller fraction than ε.

Thus, if we set

$$T := \frac{\varepsilon}{a}\log\Delta$$

and assume that we have managed to maintain the inequalities (N)— (H) up to time $t = T$. Then, by (D)

$$\log d(v, G^T) \le \log\Delta - aT \le (1-\varepsilon)\log\Delta,$$

i.e. the number of available colours exceeds the degree. Hence, we may terminate the process and colour the rest greedily.

4.9. The main lemma. First, we introduce the following statistic:

$$e'_v := \sum_{u \sim v} e(uv, p'_a)\, \mathbf{I}(u \in G').$$

Note that e'_v gives the energy $e(v, p'_a|_{G'})$ at v should v stay uncoloured. It will be more convenient to use e'_v and avoid conditioning on $v \in G'$.

The rest of the proof consist in showing the following lemma.

Lemma 9 (The Main Lemma). *Assume that* (N)—(H) *above hold. Then there is an outcome of Ω^t such that the following local events hold* globally, *i.e. for all vertices v in G:*

$$\text{(N')} \qquad p'(\mathcal{P}_v) = p(\mathcal{P}_v) \pm o(\delta_N),$$

$$\text{(D')} \qquad d(v, G') \le (1 - a_1\theta) d(v, G^t) + o(\delta_D),$$

$$\text{(E')} \qquad e'_v \le (1 - b_1\theta) e(v, p_a) + o(\delta_E)$$

$$\text{(H')} \qquad h(v, p') \ge h(v, p) - \theta C e(v, p_a) - o(\delta_E).$$

Here $a_1 > a$ and $b_1 > b$ are fixed constants, and the error-terms are given by

$$\delta_N = \theta \Delta^{-1/10}, \quad \delta_E = \theta e^{-bT} K \quad \textit{and} \quad \delta_D = \theta e^{-aT} \Delta.$$

It is easy to see that by induction this lemma implies the inequalites (N)—(H).

To prove Lemma 9 we will make use of the Local Lemma. (See e.g. [AS92].) The independence properties stated in §4.5 implies that the dependence degree of the set of events stated in Lemma 9 is bounded by a polynomial in Δ (actually $O(\Delta^4)$). It is therefore more than enough to show that each local event has a probability of order $1 - \Delta^{-\omega(1)}$ to occur. In that case, we say that the events have *very high probability.*

To prove the lemma we will show in the next section that the inequalities are true in expectation. Then we prove in the section that follows, that the variables are sufficiently concentrated, so that they with high probability are within the stated error terms from their expected values.

5. Expectations

Thus, we have to show the following bounds on the expectations:

$$\text{(5.1)} \qquad \mathrm{E}[p'(\mathcal{P}_v)] = p(\mathcal{P}_v),$$

$$\text{(5.2)} \qquad \mathrm{E}[d(v, G')] < (1 - a_1\theta) d(v, G),$$

$$\text{(5.3)} \qquad \mathrm{E}[e'_v] \le (1 - b_1\theta) e(v, p_a)$$

$$\text{(5.4)} \qquad \mathrm{E}[h(v, p')] \ge h(v, p) - \theta C e(v, p_a)$$

Note that (5.1) follows directly from the martingale property (4.5). The central point is to show (5.3): It is here we use the "trimming property" of the modifiers.

5.1. Proof of (5.2). . Since, $d(v,G') = \sum_{u\sim v} \mathbf{I}(u \in G')$ it is clearly enough to show that

$$\Pr[v \in G'] \le 1 - \theta a - \Omega(\theta). \tag{5.5}$$

We have by definition that

$$\mathbf{I}(v \in G') = \prod_{\gamma} (1 - \epsilon'(v,\gamma)\eta'(v,\gamma)\zeta'(v,\gamma)) \tag{5.6}$$

where

$$\zeta'(v,\gamma) := \prod_{\substack{w\sim v \\ w<v}} (1 - \epsilon'(w,\gamma)).$$

Taking expectations we get, since $\theta \ll 1$, that

$$\begin{aligned} \Pr[v \in G'] &= \prod_{\gamma} (1 - \frac{\theta}{2} p_c(v,\gamma)\mathbf{E}[\zeta'(v,\gamma)]) \\ &\le 1 - \sum_{\gamma} \frac{\theta}{2} p_c(v,\gamma)\mathbf{E}[\zeta'(v,\gamma)] + O(\theta^2). \end{aligned} \tag{5.7}$$

But

$$\mathbf{E}[\zeta'(v,\gamma)] \ge 1 - \theta \sum_{\substack{w\sim v \\ w<v}} p_c(w,\gamma),$$

and hence

$$\Pr[v \in G'] \le 1 - \frac{\theta}{2} \sum_{\gamma} p_c(v,\gamma) + \theta^2 e(v,p_c) + O(\theta^2). \tag{5.8}$$

Since, $p_c(\mathcal{P}_v) > 1 - 2\varepsilon = 2a + \varepsilon$ by (4.10), we get

$$\Pr[v \in G'] \le 1 - \theta a + \theta\varepsilon + O(\theta^2 K). \tag{5.9}$$

The second term is clearly negligible here and (5.5) follows. □

5.2. Proof of (5.4). For a stopped pair $(v,\gamma) \in \mathcal{I}$ it is clear that

$$\mathbf{E}[-p'(v,\gamma)\log[p'(v,\gamma)]] = -p(v,\gamma)\log p(v,\gamma).$$

From (3.1) it therefore follows that

$$\mathbf{E}[h(v,p')] \ge h(v,p) - \sum_{\gamma} p_a(v,\gamma)\kappa(M'(v,\gamma)). \tag{5.10}$$

From (3.3) and the assumptions on N_w it follows that for fixt (v,γ)

$$\begin{aligned} \kappa(M'(v,\gamma)) &\le \sum_{w\sim v} \kappa(M'_w) = \sum_{w\sim v} \theta p_c(w,\gamma)(\log 2 + \kappa(N'_w)) \\ &\le \theta \sum_{w\sim v} p_c(w,\gamma)(\log 2 + \kappa). \end{aligned} \tag{5.11}$$

By summing over γ, we see that we have shown (5.4). □

5.3. The proof for (5.3). First, we show that the possible increase in the energy is counteracted by the decrease resulting from the removal of vertices. That is, we show that

$$\mathbf{E}[e'_v] \le (1 - a\theta - \Omega(\theta))\mathbf{E}[e(v, p'_a)]. \tag{5.12}$$

Proof of (5.12). Since, $\mathbf{I}(u \in G') = \mathbf{I}(C'_u = \emptyset)$ and

$$\mathbf{I}(C'_u = \emptyset) = \mathbf{I}(C'_u \subset \{\gamma\}) - \mathbf{I}(C'_u = \{\gamma\}),$$

we have

$$\begin{aligned} e'_v &:= \sum_{u \sim v} e(uv, p'_a)\,\mathbf{I}(u \in G') \\ &= \sum_{u \sim v}\sum_{\gamma} p'_a(v,\gamma)p'_a(u,\gamma)(\mathbf{I}(C'_u \subset \{\gamma\}) - \mathbf{I}(C'_u = \{\gamma\})) \qquad (5.13) \\ &= \sum_{u \sim v}\sum_{\gamma} p'_a(v,\gamma)p'_a(u,\gamma)\,\mathbf{I}(C'_u \subset \{\gamma\}) \end{aligned}$$

since $p'_a(v,\gamma) = 0$ if $C'_u = \{\gamma\}$.

By colour independence we have

$$\mathbf{E}[p'_a(v,\gamma)p'_a(u,\gamma)\,\mathbf{I}(C'_u \subset \{\gamma\})] = \mathbf{E}[p'_a(v,\gamma)p'_a(u,\gamma)]\,\Pr[C'_u \subset \{\gamma\}],$$

and moreover

$$\Pr[C'_u \subset \{\gamma\}] - \Pr[u \in G'] \le \theta p_c(u,\gamma) \le th\hat{p}.$$

Hence,

$$\Pr[C'_u \subset \{\gamma\}] \le \Pr[u \in G'] - \theta\hat{p} \le 1 - a\theta - \Omega(\theta).$$

by (5.5). The statement (5.12) now follows. □

We note for further use that it also follows that

$$\mathbf{E}[e(uv, p'_a)\,\mathbf{I}(u \in G')] = (\Pr[u \in G'] - \theta\hat{p})\mathbf{E}[e(uv, p'_a)]. \tag{5.14}$$

Hence, by (5.12), to prove (5.3) it is enough to prove the bound

$$\mathbf{E}[e(v, p'_a)] \le (1 + o(1)) \cdot (1 + c\theta)e(v, p_a) \tag{5.15}$$

since $a - b = c$.

5.3.1. *Proof of* (5.15). Fix an edge uv and a colour γ. Note that

$$p'_a(v,\gamma)p'_a(u,\gamma) \le p_a(v,\gamma)p_a(u,\gamma)\left(\prod_{w \sim v} M'_w(v,\gamma)\right)\left(\prod_{w \sim u} M'_w(u,\gamma)\right) \tag{5.16}$$

since the r.h.s is zero unless both sides are equal. (The state p'_a contains no non-zero stopped pairs.)

Furthermore, since M'_{w_1} and M'_{w_2} are independent whenever $w_1 \neq w_2$, we get by taking expectations that

$$\mathbf{E}[p'_a(v,\gamma)p'_a(u,\gamma)] = p_a(v,\gamma)p_a(u,\gamma) \prod_{uvw \exists uv} \mathbf{E}[M'_w(u,\gamma)M'_w(v,\gamma)]. \tag{5.17}$$

the product taken over all triangles containing the edge uv.

By the definition of M'_w in (4.6) we have

$$\mathrm{E}[\, M'_w(u,\gamma)M'_w(v,\gamma) \mid (w,\gamma) \notin \mathcal{E}' \,] = 1$$

and

$$\mathrm{E}[\, M'_w(u,\gamma)M'_w(v,\gamma) \mid (w,\gamma) \in \mathcal{E}' \,] = 2\mathrm{E}[\, N'_w(u,\gamma)N'_w(v,\gamma) \,].$$

It follows that

$$\mathrm{E}[\, M'_w(u,\gamma)M'_w(v,\gamma) \,] = 1 + \theta p_c(w,\gamma) f_w(uv,\gamma) \tag{5.18}$$

where

$$f_w(uv,\gamma) := 2\mathrm{E}[\, N'_w(u,\gamma)N'_w(v,\gamma) \,] - 1.$$

Note that, by (4.3), we have

$$\sum_{uvw \ni uv} |\theta p_c(w,\gamma) f_w(uv,\gamma)| \le \Delta\theta\hat{p}\hat{k}^2 = \Delta^{-\varepsilon}. \tag{5.19}$$

Then, by Taylor approximation, we get that $\mathrm{E}[p'_a(v,\gamma)p'_a(u,\gamma)] - p_a(v,\gamma)p_a(u,\gamma)$ is smaller than

$$(1+o(1))\theta \sum_{uvw \ni uv} p_a(v,\gamma)p_a(u,\gamma)p_c(w,\gamma)f_w(uv,\gamma).$$

Summing over $u \sim v$ gives, since $p_c \le p_a$, that

$$\mathrm{E}[p'_a(v,\gamma)p'_a(N_{(v,\gamma)})] - p_a(v,\gamma)p_a(N_{(v,\gamma)}),$$

is less than

$$\begin{aligned}
&(1+o(1))\theta \sum_{u\sim v} \sum_{uvw \ni uv} p_a(u,\gamma)p_a(v,\gamma)p_a(w,\gamma)f_w(uv,\gamma) \\
&\quad \le (1+o(1))\theta \sum_{w\sim v} p_a(v,\gamma)p_a(w,\gamma) \sum_{uvw \ni vw} p_a(u,\gamma)f_w(uv,\gamma) \\
&\quad \le (1+o(1))\theta c p_a(v,\gamma)p_a(N_{(v,\gamma)}),
\end{aligned}$$

where the last inequality is from the trimming property. Finally, summing this over γ gives (5.15). □

Remark 5.1. The definition of θ in (4.3) implies the bound (5.19) which is necessary for the Taylor approximation above. A more elaborate analysis could probably allow for a much greater value for θ, since the bound $\hat{p}\Delta$ is much to big compared to a "typical" value of $p(N_{(v,\gamma)})$. A bigger θ would help to reduce the size of the constants. Note also that the Taylor approximation in unnecessary if the modifier is contracting.

6. Concentration

We can now conclude the proof of Lemma 9 by showing the following statements about concentration:

$$\text{with v.h.p:} \quad \left|\mathrm{Z}[p'(\mathcal{P}_v)]\right| \le o(\delta_N), \tag{6.1}$$

$$\text{with v.h.p:} \quad \left|\mathrm{Z}[d(v,G')]\right| \le o(\delta_D), \tag{6.2}$$

$$\text{with v.h.p:} \quad \left|\mathrm{Z}[e(v,p'_a)]\right| \le o(\delta_E), \tag{6.3}$$

$$\text{with v.h.p:} \quad \left|\mathrm{Z}[h(v,p')]\right| \le o(\delta_E), \tag{6.4}$$

where $\mathrm{Z}[X]$ stands for $X - \mathrm{E}[X]$.

The main property we shall use in proving (6.1) – (6.4) is the independence between distinct colours. This allows us to use Bernstein's theorem (see [?] or [?]), which has the following simple corollary from [?].

Lemma 10. *Assume that $\{Y_j : j \in J\}$ is a finite set of independent* non-negative *random variables such that $Y_j \le c$. If X is the random variable $X = \sum_j Y_j$, with expectation $\mathrm{E}[X] = O(\mu)$ and*

$$\lambda^2/(c\mu) = \omega(\log \Delta) \quad \textit{and} \quad \lambda/c = \omega(\log \Delta)$$

then it holds with v.h.p that

$$|Z[X]| \le \lambda.$$

Notation. It is convenient to write $f = \omega_c(g)$ if $f = \omega(\log \Delta \cdot g)$ and $f = o_c(g)$ if $g/\log \Delta = o(f)$. For given μ, c and λit thus suffice to check that

$$\lambda = \omega_c(\sqrt{c\mu} + c) \quad \text{or} \quad \sqrt{c\mu} + c = o_c(\lambda)$$

We list some asymptotic bounds on the parameters which are used below. They can all be easily checked from the definitions. The "limiting" bound is (6.10) below.

$$\sqrt{2\hat{\kappa}\hat{p} \cdot 1} + 2\hat{\kappa}\hat{p} = o_c(\delta_N) \tag{6.5}$$

$$\sqrt{2\hat{\kappa}\hat{p} \cdot (\log \Delta)^2} + \log \Delta \cdot 2\hat{\kappa}\hat{p} = o_c(\delta_E) \tag{6.6}$$

$$\sqrt{\theta\Delta} + 1 = o_c(\delta_D) \tag{6.7}$$

$$2\theta\hat{p}K = o(\delta_E) \tag{6.8}$$

$$\sqrt{2\theta\hat{p}K} + \hat{p} = o_c(\delta_E) \tag{6.9}$$

$$\sqrt{2K\Delta} \cdot \hat{p} + \Delta \cdot \hat{p}^2 = o_c(\delta_E). \tag{6.10}$$

Proof of (6.1) *and* (6.4). Both $\sum_\gamma p'(v,\gamma)$ and $\sum_\gamma -p'vg \log p'(v,\gamma)$ are obviously sums of independent terms, so the lemma applies.

In (6.1) we have $\mu = 1$ and $c = 2\hat{\kappa}\hat{p}$, hence it follows from (6.5) and Lemma 10. For the entropy we have $\mu = O(\log \Delta)$ and $c = 2\hat{\kappa}\hat{p}\log \Delta$. Thus (6.6) implies (6.4). □

Hence, it only remains to prove (6.2) and (6.3), where we must work a bit harder since the sums involved do not have independent terms. However, the following lemma help us to overcome this obstacle.

Lemma 11. *Let the random variable X be defined by*

$$X := \sum_{v \in G} c_v \, \mathbf{I}(v \in G')$$

for constants c_v with $|c_v| \le c$. Then $|Z(X)| < \lambda$ with very high probability if

$$\frac{\lambda^2}{c\mu} = \omega(\log \Delta) \quad \textit{and} \quad \lambda/c = \omega(\log \Delta)$$

where

$$\mu := \theta \sum_v c_v$$

6.1. The proof of (6.2) and (6.3). We use Lemma 11 to finish the proof:

Proof of (6.2). The lemma immediately implies (6.2) by taking

$$c_v := 1_{N(v,G)} \leq 1.$$

Since $c = 1$ and $\mu = \theta\Delta$ (6.2) follows from (6.7). □

Proof of (6.3). Let v be fixt and let $c'_u = e(uv, p'_a)$ for $u \sim v$, $c_u := \mathrm{E}[c'_u]$ and

$$X' := e'_v = \sum_{u\sim v} \mathbf{I}(u \in G')\, c'_u$$
$$X := \sum_{u\sim v} \mathbf{I}(v \in G')\, c_u.$$

Then

$$\begin{aligned} |\mathrm{Z}[X']| &\leq |X' - X| + |X - \mathrm{E}[X]| + |\mathrm{E}[X] - \mathrm{E}[X']| \\ &\leq \sum_{u\sim v} |c'_u - c_u| + |\mathrm{Z}[X]| + |\mathrm{E}[X] - \mathrm{E}[X']|. \end{aligned} \tag{6.11}$$

The last term is smaller than

$$\theta\hat{p} \cdot \sum_{u\sim v} c_u \leq \theta\hat{p} \cdot 2K = o(\delta_E)$$

by (6.8), (5.15) and (5.14).

It is therefore enough to show that

$$\textit{with v.h.p:} \qquad \sum_{u\sim v} |c'_u - c_u| = o(\delta_E) \tag{6.12}$$

and that with *v. h. p*

$$|\mathrm{Z}[X]| = o(\delta_E).$$

The last statement follows from Lemma 11: We have $c = \hat{p}$, $\mu = \theta \sum_{u\sim v} c_u \leq \theta 2K$ and

$$\sqrt{\hat{p}\theta 2K} + \hat{p} = o_c(\delta_E).$$

To prove (6.12), we use Lemma 10: Each c'_u is a sum of independent terms bounded by $\hat{p}^2$ and expectation c_u. We can deduce that

$$\textit{with v.h.p:} \qquad |c'_u - c_u| \leq \omega_c \left(\sqrt{\hat{p}^2 c_u} + \hat{p}^2\right).$$

By Cauchy-Schwarz

$$Q := \sum_{u\sim v} (\hat{p}\sqrt{c_u} + \hat{p}^2) \leq \sqrt{\Delta} \cdot \hat{p} \cdot \sqrt{2K} + \Delta\hat{p}^2.$$

But, since the intersection of at most Δ events of *v. h. p* is an event of *v. h. p*, the sum $\sum_{u\sim v} |c'_u - c_u|$ is therefore with *v. h. p* smaller than $o(\delta_E) = \omega_c(Q)$, by (6.10). □

6.2. The proof of Lemma 11. To make the proof complete, we shall prove Lemma 11. Following the standard procedure (See e.g. [?] or [Kah93].) it is enough to show the following bound on the Laplace transform

$$\mathrm{E}\left[e^{\xi Z}\right] \le e^{\xi^2 Q}, \quad \text{for } \xi < 1/c \quad , \tag{6.13}$$

for $Z = \pm \mathrm{Z}[X]$ and $Q = O(c\mu) \ge 0$.

To see how (6.13) implies the lemma: Note that from Chebyshev's inequality follows

$$\Pr[Z > \lambda] = \Pr\left[e^{\xi Z} > e^{\xi\lambda}\right] \le e^{\xi^2 Q - \xi\lambda}.$$

With $\xi = \lambda/2Q$, assuming $\lambda \le 2Q$, the right hand side is $e^{-\lambda^2/4Q} = e^{-\omega(\log \Delta)}$. This implies the lemma if $\lambda \le 2Q/c$. Should this not be the case, we repeat the argument with $Q' = c\lambda > Q$, and $\xi = \lambda/2Q' = 1/2c$ instead. Finally, the lemma follows by repeating the argument with $-Z$ instead of Z.

In order to prove (6.13), we use Lemma 12 below which is also more or less standard. But before stating this, we introduce some notation: Given a probability space $(\Omega, \mathcal{M}, \Pr[\cdot])$ and a filtration $\mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_n$ of sub-(sigma)-algebras of $\mathcal{M}$, let $\mathrm{Z}_i[\cdot] := \mathrm{Z}_{i,i-1}[\cdot]$ where

$$\mathrm{Z}_{i,j}[\cdot] := \mathrm{E}[\cdot \mid \mathcal{M}_i] - \mathrm{E}[\cdot \mid \mathcal{M}_{i-1}].$$

Then we have the following.

Lemma 12. *If X is a r.v. such that for each i for some ξ*

$$\mathrm{E}\left[e^{\xi \mathrm{Z}_i[X]} \mid \mathcal{M}_{i-1}\right] \le 1 + \xi^2 Q_i. \tag{*}$$

Then

$$\mathrm{E}\left[e^{\xi \mathrm{Z}_{0,n}[X]} \mid \mathcal{M}_0\right] \le e^{\xi^2 \sum_i Q_i}.$$

Furthermore, if for each i

$$|\mathrm{Z}_i[X]| \le c_i \le c$$

and

$$\mathrm{E}\left[\left|\mathrm{Z}_i[X]\right| \mid \mathcal{M}_{i-1}\right] \le b_i$$

then (*) holds with $Q_i = c_i b_i$ and $|\xi| \le 1/c$.

Proof. The first statement follows from

$$\begin{aligned} \mathrm{E}\left[e^{\xi \mathrm{Z}_{n,0}[X]} \mid \mathcal{M}_0\right] &= \mathrm{E}\left[\mathrm{E}\left[e^{\xi \mathrm{Z}_n[X]} \mid \mathcal{M}_{n-1}\right] e^{\xi \mathrm{Z}_{n-1,0}[X]} \mid \mathcal{M}_0\right] \\ &\le e^{\xi^2 Q_n} \cdot \mathrm{E}\left[e^{\xi \mathrm{Z}_{n-1,0}[X]} \mid \mathcal{M}_0\right] \end{aligned} \tag{6.14}$$

and induction.

The second statement follows readily from a Taylor approximation:

$$e^x \le 1 + x + x^2 \quad \text{if } |x| < 1.$$

□

6.3. The proof of (6.13). We assume that $V = [1, N]$ and that the ordering used in the definition of $\tilde{\mathcal{E}}'$ is the natural one. Let $\mathcal{M}_0 = \emptyset$ and for $v \in V$ define the algebra

$$\mathcal{M}_v := \{\epsilon'(w,\gamma), \eta'(w,\gamma) : \gamma \in \Gamma, w \le v\}.$$

Denote by $\mathrm{E}_v[\cdot]$ the conditional expectation $\mathrm{E}[\cdot \mid \mathcal{M}_v]$ relative the algebra $\mathcal{M}_v$.

Let $Y_v = c_v \mathbf{I}(v \in G')$ and write $T_v = \sum_{u>v} Y_u$. Then

$$\mathrm{Z}_v[X] = \mathrm{Z}_v[Y_v] + \mathrm{Z}_v[T_v],$$

since the value of Y_u, for $u < v$, is determined by $\mathcal{M}_{v-1}$. Furthermore, since

(6.15)
$$\mathrm{E}_{v-1}\left[e^{\xi \mathrm{Z}_v[X]}\right] \le \frac{1}{2}\left(\mathrm{E}_{v-1}\left[e^{2\xi \mathrm{Z}_v[Y_v]}\right] + \mathrm{E}_{v-1}\left[e^{2\xi \mathrm{Z}_v[T_v]}\right]\right).$$

it is enough to show the following two inequalities:

(6.16)
$$\mathrm{E}_{v-1}\left[e^{\xi \mathrm{Z}_v[Y_v]}\right] \le 1 + \xi^2 \theta c_v^2$$

and

(6.17)
$$\mathrm{E}_{v-1}\left[e^{\xi \mathrm{Z}_v[T_v]}\right] \le 1 + \xi^2 R_v$$

where

$$R_v = 2K\theta^3 \sum_{u>v} c_u^2 \cdot e(uv, p_c).$$

This implies (6.13) since summing the R_v's gives

$$\sum_v R_v \le 2K\theta^3 \cdot \left(\sum u \in V c_u^2 \cdot \sum_{v<u} e(uv, p_c)\right)$$
$$\le O\left(K^2\theta^2\right) \cdot c\mu = o(c\mu).$$

6.3.1. *Proof of* (6.16). Clearly, $|\mathrm{Z}_v[Y_v]| \le |c_v| \le c$ and for any indicator variable I we have

$$\mathrm{E}[|\mathrm{Z}[I]|] \le 2\min\{\mathrm{E}[I], 1 - \mathrm{E}[I]\},$$

and hence that $\mathrm{E}_{v-1}\left[\,|\mathrm{Z}_v[Y_v]|\,\right] \le 2\theta c_v$. By Lemma 12, this proves (6.16). □

6.3.2. *Proof of* (6.17). In order to show (6.17) we study what happens at v a bit more closely. Assume that $\Gamma = [1, M]$ and define a filtration

$$\mathcal{M}_{v-1} = \mathcal{N}_0 \subset \cdots \subset \mathcal{N}_\gamma \subset \cdots \subset \mathcal{N}_M = \mathcal{M}_v$$

where

$$\mathcal{N}_\gamma := \langle \epsilon'(v,\tilde{\gamma}), \eta'(v,\tilde{\gamma}) : \tilde{\gamma} \le \gamma \rangle \vee \mathcal{M}_{v-1}.$$

Thus we look at one excitation at a time. Set $\mathrm{E}_\gamma[\cdot] := \mathrm{E}[\cdot \mid \mathcal{N}_\gamma]$ and

$$\mathrm{Z}_\gamma[X] := \mathrm{E}[X \mid \mathcal{N}_\gamma] - \mathrm{E}[X \mid \mathcal{N}_{\gamma-1}].$$

To prove (6.16), it is by Lemma 12 enough to show

(6.18)
$$|\mathrm{Z}_\gamma[T_v]| \le \theta \sum_{\substack{u \sim v \\ u>v}} c_u p_c(u,\gamma)$$

and

(6.19)
$$\mathrm{E}_{\gamma-1}[|\mathrm{Z}_\gamma[T_v]|] \le \theta^2 \sum_{\substack{u \sim v \\ u>v}} c_u p_c(u,\gamma) p_c(v,\gamma).$$

Note that, by the definition of p_c, the r.h.s. of (6.18) is of order $O(c \cdot K \cdot \theta) = o(c)$ whenever $p_c(v,\gamma) > 0$, which is the only case we have to consider, since otherwise, clearly, both bounds are zero.

By definition, for $u \in G$,

$$Y_u = c_u \cdot \prod_{\gamma} (1 - \theta \eta'(u,\gamma)\epsilon'(u,\gamma)\zeta'(u,\gamma))$$

where

$$\zeta'(u,\gamma) := \prod_{\substack{w \sim u \\ w < u}} (1 - \epsilon'(w,\gamma)).$$

Therefore, by independence, for $v < u$,

$$\mathrm{E}_\gamma[Y_u] = c_u \prod_{\tilde{\gamma}} (1 - \frac{\theta}{2} p_c(u,\tilde{\gamma}) \mathrm{E}_\gamma[\zeta'(u,\tilde{\gamma})])$$

where

(6.20)

$$\mathrm{E}_\gamma[\zeta'(u,\tilde{\gamma})] = \left(\prod_{\substack{w \sim u \\ v < w < u}} (1 - \theta p_c(w,\tilde{\gamma})) \right) \cdot \left(\prod_{\substack{w \sim u \\ w \le v}} (1 - \epsilon'(w,\tilde{\gamma})) \right).$$

Hence

$$\mathrm{Z}_\gamma[Y_u] = -c_u \cdot P \cdot \frac{\theta}{2} p_c(u,\gamma) \mathrm{Z}_\gamma[\zeta'(u,\gamma)]. \tag{6.21}$$

where

$$P = \prod_{\tilde{\gamma} \neq \gamma} (1 - \frac{\theta}{2} p_c(u,\tilde{\gamma}) \mathrm{E}_\gamma[\zeta'(u,\tilde{\gamma})]) \le 1.$$

From (6.20) it is clear that

$$|\mathrm{Z}_\gamma[\zeta'(u,\gamma)]| \le 1 \quad \text{and} \quad \mathrm{E}_{\gamma-1}[|\mathrm{Z}_\gamma[\zeta'(u,\gamma)]|] \le 2\theta p_c(v,\gamma),$$

which, by (6.21), implies that

(6.22)

$$|\mathrm{Z}_\gamma[Y_u]| \le c_u \frac{\theta}{2} p_c(u,\gamma) \quad \text{and} \quad \mathrm{E}_{\gamma-1}[|\mathrm{Z}_\gamma[Y_u]|] \le c_u \theta^2 p_c(u,\gamma) p_c(v,\gamma).$$

By summing over $u > v$ we get (6.18) and (6.19). □

Acknowledgements

References

[AEKS81] M. Ajtai, P. Erdös, J. Komlós, and E. Szemerédi, *On Turán's theorem for sparse graphs*, Combinatorica 1 (1981), 313–317.

[AS92] N. Alon and J. H. Spencer, *The probabilistic method*, vol. 1, Wiley, New York, 1992.

[Kah90] Jeff Kahn, *Recent results on some not-so-recent hypergraph matching and covering problems*, 1990, Preprint from Proc. Visegrad Conference of Finite Sets.

[Kah93] Jeff Kahn, *Asymptotically good list-colourings*, 1993, Preprint.

[She95] J. Shearer, *A note on the independence number of k_r-free graphs*, Random Structures and Algorithms 7 (1995), no. 3, 269–271.