COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA 18 (2) (1998) 201–226

RANDOM MATCHINGS IN REGULAR GRAPHS

JEFF KAHN* and JEONG HAN KIM

Received April 12, 1996

For a simple d-regular graph G, let M be chosen uniformly at random from the set of all matchings of G, and for $x \in V(G)$ let $p(\overline{x})$ be the probability that M does not cover x.

We show that for large d, the $p(\overline{x})$'s and the mean μ and variance σ^2 of $|M|$ are determined to within small tolerances just by d and (in the case of μ and σ^2) $|V(G)|$:

Theorem. *For any d-regular graph G,*
(a) $p(\overline{x}) \sim d^{-1/2}$ $\forall x \in V(G)$, *so that* $|V(G)| - 2\mu \sim |V(G)|/\sqrt{d}$,
(b) $\sigma^2 \sim |V(G)|/(4\sqrt{d})$,
where the rates of convergence depend only on d.

1. Introduction

Given a graph $G=(V,E)$, write $\mathcal{M}(G)$ for the set of matchings of G, and let M be chosen uniformly at random from $\mathcal{M}(G)$. (For graph theory background see e.g. [24]. We use "graph" to mean *simple* graph.) In this paper we are concerned with the behavior of M, and in particular of the random variable $\xi = \xi_G = |M|$, when G is regular of large degree.

Set $p_k = p_k(G) = \Pr(\xi = k)$. The distribution $\{p_k\}$ (for a general G) has been considered in many contexts, in physics and chemistry as well as mathematics. We will not try to give a thorough bibliography, but see e.g. [20], [13], [23], [7], [8], [9], [24, Chapter 8].

These distributions are in some ways very nice. For instance, as shown in [12], [13], [23], for any G the probability generating function

$$f(G;\lambda) = \sum_k p_k \lambda^k \tag{1}$$

has real roots. This gives log-concavity of the sequence $\{p_k\}$ (*c.f.* "Newton's inequalities," e.g. [10, p.51]), and implies that the distribution is approximately normal provided the variance $\sigma^2 = \sigma_\xi^2 =: \sigma^2(G)$ is large. (The latter is essentially

Mathematics Subject Classification (1991): 05C65, 05C70, 05C80, 60C05, 60F05, 60G42

* Supported by NSF

due to L. Harper [11]. See the two paragraphs preceding Theorem 1.2 for some discussion and references concerning the question of *when* σ^2 is large.)

Here we show that for regular G the behavior of $\{p_k\}$ is nice in another sense: the mean ($\mu=\mu_\xi=:\mu(G)$) and variance of ξ are remarkably well determined just by the degree and number of vertices of G.

Before stating this we need a finer parameter than μ. For $x\in V$, write $x\prec M$ if x is covered by (i.e. is contained in some edge of) the matching M, and set

$$p(\overline{x}) = p_G(\overline{x}) = \Pr(x \not\prec M).$$

Thus $\mu=(n-\sum_{x\in V} p(\overline{x}))/2$, where, here and throughout the paper, we set $|V|=n$.

Theorem 1.1. *For any d-regular graph G,*
(a) $p(\overline{x})\sim d^{-1/2}\ \ \forall x\in V(G)$, *so that* $n-2\mu(G)\sim n/\sqrt{d}$,
(b) $\sigma^2(G)\sim n/(4\sqrt{d})$.

Here the limits are taken as $d\to\infty$; so for example $p(\overline{x})\sim d^{-1/2}$ means

$$(1-o(1))d^{-1/2} < p(\overline{x}) < (1+o(1))d^{-1/2},$$

where *o(1) depends only on d, and not on G or x.* Let us stress that what's interesting here is the *existence* of the limiting values ($d^{-1/2}$, $n/(4\sqrt{d})$), rather than the values themselves.

(The values themselves are easily seen to be a natural expression of the idea that the events $\{x\prec M\}$ are roughly independent. To see this, we observe the easy identity (see (5))

$$p(\overline{x}) = \left(1+\sum_{y\sim x} p(\overline{y}|\overline{x})\right)^{-1}$$

(where the conditional probability $p(\overline{y}|\overline{x})$ has the obvious meaning). Using this, if we pretend the events $\{x\prec M\}$ are mutually independent with $p(\overline{x})=p$ for all x, then

$$p = (2d)^{-1}(-1+\sqrt{1+4d}) = d^{-1/2}+(2d)^{-1}+O(d^{-3/2}) \tag{2}$$

gives (a) (see also Conjecture 1.3); while (b) derives from the fact that ξ is half the random variable $|\{x\in V : x\prec M\}|$, which has the binomial distribution $B(n,1-p)$, so variance $np(1-p)\sim nd^{-1/2}$.)

Let us also mention that it is not even easy to show that a large regular G has large $\sigma^2(G)$; precisely: if G_α is d_α-regular ($d_\alpha\neq 0$) with $n_\alpha:=|V(G_\alpha)|\to\infty$, then $\sigma^2(G_\alpha)\to\infty$ as $\alpha\to\infty$. This was shown in [7] provided $d_\alpha/n_\alpha\to 0$, but in full generality only in [18] (with a proof quite different from the arguments used here). So it is, again, rather surprising that one can say something as precise as Theorem 1.1.

(As more or less observed following (1), the condition $\sigma^2(G_\alpha)\to\infty$ is equivalent to asymptotic normality of $\{p_k(G_\alpha)\}_{k\geq 0}$. This was the reason for most of the earlier work on $\sigma^2(G)$—see [28], [24, Ch. 8] in addition to [7], [18]—though not our principal motivation here.)

The actual bounds we establish are given in

Theorem 1.2. *For any d-regular graph G and* $\varepsilon>0$

(a) $p(\overline{x})=d^{-1/2}+O(d^{-3/4+\varepsilon}) \quad \forall x\in V(G)$,

(b) $\sigma^2(G)=\left(1+O(d^{-1/4+\varepsilon})\right)\dfrac{n}{4\sqrt{d}}$.

Again note the error terms depend only on d. (This is slightly abusive, in a standard way. For example the error term in (a) does depend on G, x, but is bounded by some $O(d^{-3/4+\varepsilon})$ which depends only on d. So we should really write $|p(\overline{x})-d^{-1/2}|<O(d^{-3/4+\varepsilon})$.)

Theorem 1.2 is proved beginning in Section 2. Before closing the present section, we just mention a few related questions.

First, surprisingly accurate though they are, it seems possible that the bounds in Theorem 1.2 can be strengthened considerably (compare (2)):

Conjecture 1.3. *For any d-regular graph G and* $x\in V(G)$,

(a) $p(\overline{x})=d^{-1/2}-(2d)^{-1}+O(d^{-3/2})$,

(b) $\sigma^2(G)=\frac{1}{4}nd^{-1/2}+O(nd^{-1})$.

This may be wishful thinking. It does, admittedly, seem too good to be true, but the same might have been (and was) said of Theorem 1.1 when it was not yet a theorem. Certainly the conjecture, if true, would be fairly remarkable.

Second, it would be of considerable interest if something like Theorem 1.1 were true for hypergraphs of fixed edge size. We recall a few definitions. (For further background see e.g. [6] or [16].) A *hypergraph* $\mathcal{H}$ on a *vertex* set V is simply a collection of subsets of V, and is *k-uniform* if each of its members (called *edges*) is of size k. (So a 2-uniform hypergraph is a graph.) A hypergraph is *d-regular* if each of its vertices is contained in exactly d edges, and *simple* if no two of its vertices are contained in two distinct edges.

A *matching* in a hypergraph is again a collection of pairwise disjoint edges, and, as for graphs, we write $\nu(\mathcal{H})$ for the size of a largest matching of $\mathcal{H}$. We extend our earlier notation ($\mathcal{M}$, ξ, $p(\overline{x})\ldots$) in the natural ways.

Conjecture 1.4. *Fix k. If* $\mathcal{H}$ *is a simple, k-uniform, d-regular hypergraph on a vertex set V of size n, then*

(a) $p(\overline{x})\sim d^{-1/k}\ \ \forall x\in V$, *so that* $n-k\mu(\mathcal{H})\sim nd^{-1/k}$, *and in particular* $\mu(\mathcal{H})\sim n/k$,

(b) $\sigma^2(\mathcal{H})\sim n/(k^2d^{1/k})$

(where again limits are taken as $d\to\infty$*).*

This would be extremely interesting, not only for its own sake, but also because of its relation to work done over the last fifteen or so years on the asymptotic behavior of hypergraphs of bounded edge size. A central result in this area, proved by N. Pippenger following ideas of Ajtai, Komlós and Szemerédi [1], Rödl [27] and Frankl and Rödl [5], says in part:

Theorem 1.5. (unpublished; see [29], [6]) *Fix* k. *If* $\mathcal{H}$ *is as in Conjecture* 1.4, *then*

$$\nu(\mathcal{H}) > (1 - o(1))n/k, \tag{3}$$

where $o(1)$ *depends only on* d.

(See also e.g. [26], [6], [16], [15], [21], [19], [14], [30], [17], [22] for exposition and related work. For Pippenger's Theorem in full we should relax "simple" to a (uniform) bound $o(d)$ on the pairwise degrees $d(x,y)$; but we are in deep enough waters with the present hypotheses and will not explore this extra generality.)

As was pointed out to us by Anders Johansson, it is not hard to show (using Pippenger's Theorem) that (3) is still true with μ in place of ν. Of course Conjecture 1.4 (a) implies a much stronger result. (In fact the error term $(1+o(1))d^{-1/k}n/k$ (in the approximation $\mu \approx n/k$) implied by the conjecture is far better than what was known, even for ν, when the present paper was written. The stronger bounds

$$n - k\nu = \begin{cases} O(nd^{-1/2}\log^{3/2} d) & \text{if } k = 3 \\ O(nd^{-1/(k-1)}) & \text{if } k > 3 \end{cases}$$

were subsequently established in [2].)

In contrast, for a *graph* G as in Theorem 1.1: (i) Vizing's Theorem ([31] or e.g. [24, Theorem 7.4.1]) implies $\nu(G) \geq (1 - 1/(d+1))n/2$; and (ii) $\mu > (1 - O(d^{-1/2}))n/2$—a less precise version of Theorem 1.1 (a)—is not too hard to prove using the approach of Section 2 (see (8)).

2. Path-trees and indication of proof

We first recall Godsil's [8] notion of the *path-tree* $T(G,v)$ associated with a graph G and $v \in V(G)$. (This is called a *tree of walks* in [8]. The present name is from [24]. Were it not for its length, we would prefer "tree of self-avoiding walks," since we will eventually view the vertices of $T(G,v)$ as outcomes of a random self-avoiding walk in G.)

The vertices of $T = T(G,v)$ are the paths of G which begin at v. (For our purposes a *path* is a sequence $(y_0, y_1 \ldots y_l)$ of distinct vertices with $y_i \sim y_{i-1}$.) Two vertices of T are adjacent if one is a maximal proper subpath of the other.

We will usually use $\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \ldots$ for vertices of T, and in particular write V for the singleton path (v), which we regard as the root of T. Later we will be interested in a random path $(v = y_0, y_1 \ldots y_k)$ in G, and will write Y_l for the vertex $(y_0 \ldots y_l)$ (where $l \leq k$). For $\mathrm{W} \in V(T)$ we write $|\mathrm{W}|$ for the length of the path W, in other words the depth of W in T, and set $T_l=\{\mathrm{W}\in V(T):|\mathrm{W}|=l\}$. We use $S(\mathrm{W})$ for the set of children of W, $s(\mathrm{W})$ for $|S(\mathrm{W})|$ and $T(\mathrm{W})$ for the subtree rooted at W.

Path-trees $T(G,v)$ turn out to capture considerable information about matchings in G, and to be in some respects easier to work with than the graph itself. (Again see [8] or the exposition in [24].) For present purposes the relevant connection is given by

Lemma 2.1. *With notation as above,* $p_G(\overline{v})=p_{T(G,v)}(\overline{\mathrm{V}})$.

That is, the probability that a random matching of G misses v is the same as the probability that a random matching of T misses V.

Proof. This is an immediate consequence of the main result of [8], which we repeat here for the reader's convenience.

The *matching generating polynomial* of G is

$$g(G;\lambda)=|\mathcal{M}(G)|f(G;\lambda)=\sum_k m_k\lambda^k$$

where $m_k=m_k(G)$ is the number of matchings of size k in G. The main result of [8] is (equivalent to)

$$g(G-v;\lambda)/g(G;\lambda)=g(T-\mathrm{V};\lambda)/g(T;\lambda).$$

Evaluation at $\lambda=1$ gives the lemma. ∎

An advantage of working with $T(G,v)$ is that it allows us to compute probabilities $p_G(\overline{x})$ recursively. Let us extend our earlier notation, writing $p(\overline{y}|\overline{x})$ for $p_{G-x}(\overline{y})$ and $p(\overline{x},\overline{y})$ for $\Pr(x\not\sim M,y\not\sim M)$. Since $p(\overline{x},\overline{y})=p(\{x,y\}\in M)$ when $x\sim y$, we have

$$p(\overline{x})+\sum_{y\sim x}p(\overline{x},\overline{y})=1, \tag{4}$$

which, when divided by $p(\overline{x})$, gives the basic identity

$$p(\overline{x})=\left(1+\sum_{y\sim x}p(\overline{y}|\overline{x})\right)^{-1}. \tag{5}$$

For trees this takes the form

$$p_{T(\mathrm{Y})}(\overline{\mathrm{Y}})=\left(1+\sum_{\mathrm{Z}\in S(\mathrm{Y})}p_{T(\mathrm{Z})}(\overline{\mathrm{Z}})\right)^{-1} \tag{6}$$

where we write $T(\mathrm{Y})$ for the subtree rooted at Y. Thus in principle we may compute the probabilities $p_{T(\mathrm{Y})}(\overline{\mathrm{Y}})$ recursively, beginning at the leaves and working up to the root, V, for which $p_{T(\mathrm{V})}(\overline{\mathrm{V}})=p_T(\overline{\mathrm{V}})$.

For example, if $T=T(G,v)$ with G d-regular, then it's not hard to use this recursion together with the obvious

$$d-l \le s(\mathrm{W}) \le d \quad \forall \mathrm{W} \in T_l \tag{7}$$

to show

$$c_1 d^{-1/2} < p_G(\overline{v}) < c_2 d^{-1/2} \quad \forall v \in V(G) \tag{8}$$

for some positive constants c_1, c_2. (This gives the bound $\mu(G) > (1-O(d^{-1/2}))n/2$ mentioned at the end of Section 1.)

For Theorem 1.2 the inequalities (7) are not enough—e.g. the reader could try evaluating the extreme case

$$s(\mathrm{W}) = \begin{cases} d & \text{if } |\mathrm{W}| \text{ is even} \\ d-|\mathrm{W}| & \text{if } |\mathrm{W}| \text{ is odd} \end{cases} \tag{9}$$

—and we must show that degree fluctuations in $T(G,v)$ are, in some usable sense, much more moderate than those in (9).

This is accomplished by comparing the degree $s(\mathrm{W})$ of a vertex W with the average of the degrees of its children,

$$\overline{s}(\mathrm{W}) = \frac{1}{s(\mathrm{W})} \sum_{\mathrm{U} \in S(\mathrm{W})} s(\mathrm{U}).$$

We show that, in contrast to (9), $s(\mathrm{W})$ and $\overline{s}(\mathrm{W})$ are close for most $\mathrm{W} \in V(T)$. For the precise technical statement, set

$$\Gamma(l,\varepsilon) := \{\mathrm{W} \in T_l : |s(\mathrm{W}) - \overline{s}(\mathrm{W})| > d^{1/4+\varepsilon}\} \quad \text{and} \quad \gamma(l,\varepsilon) := |\Gamma(l,\varepsilon)|,$$

and let $t=4\lfloor\sqrt{d}\log d\rfloor$. (To prove (8) it's enough to consider something like the first $\sqrt{d}\log d$ levels of T, and this will again be true for the proof of Theorem 1.2.)

Lemma 2.2. *For any fixed $\varepsilon>0$, if d is sufficiently large and $l \le t$, then*

$$\gamma(l,\varepsilon) < t^{-1}(d-t)^l e^{-d^{\varepsilon}}. \tag{10}$$

This is proved in Section 3, and the derivation of Theorem 1.2 is completed in Sections 4 and 5. The bound (10) is given in a form convenient for later calculations, and is slightly weaker than what's produced in Section 3.

3. Proof of Lemma 2.2

For $w \in V(G)$, we write $N(w)$ for the set of neighbors of w.

For $\mathrm{W} = (v = w_0 \dots w_l) \in V(T)$ let

$$\delta(\mathrm{W}) = d - s(\mathrm{W}) = |\{w_0 \dots w_{l-1}\} \cap N(w_l)|.$$

Set

$$\overline{\delta}(\mathrm{W}) = \frac{1}{s(\mathrm{W})} \sum_{\mathrm{U} \in S(\mathrm{W})} \delta(\mathrm{U}) = d - \overline{s}(\mathrm{W}).$$

Our proof of Lemma 2.2 is more naturally expressed in terms of these parameters, that is, with $\gamma(l,\varepsilon)$ rewritten as

$$\gamma(l,\varepsilon) = |\{\mathrm{W} \in T_l : |\delta(\mathrm{W}) - \overline{\delta}(\mathrm{W})| > d^{1/4+\varepsilon}\}|.$$

Let $(v = y_0, y_1 \dots y_t)$ be the natural random self-avoiding walk given by $y_0 = v$ and

$$\Pr(y_i = w | y_0 \dots y_{i-1}) = s(y_{i-1})^{-1} 1_{\{w \in S(y_{i-1})\}}, \tag{11}$$

where, in agreement with our notation for T,

$$S(y_{i-1}) = N(y_{i-1}) \setminus \{y_0 \dots y_{i-2}\}$$

and $s(y_{i-1}) = |S(y_{i-1})|$; that is, the walk chooses y_i uniformly from the as yet unvisited neighbors of y_{i-1}.

As earlier, we write Y_l for $(y_0 \dots y_l)$, thought of as a random vertex of T_l. We will show that for Y_l chosen according to this (not quite uniform) distribution on T_l, $|\delta(\mathrm{Y}_l) - \overline{\delta}(\mathrm{Y}_l)|$ is very unlikely to be large; precisely, for any $\alpha > 0$,

$$\Pr(|\delta(\mathrm{Y}_l) - \overline{\delta}(\mathrm{Y}_l)| > \alpha + 32 \log^2 d) < 2d^2 t \exp\left(-\frac{\alpha^2}{2t}\right). \tag{12}$$

To see that this implies Lemma 2.2, note that for any $\mathrm{W} \in T_l$,

$$\Pr(\mathrm{Y}_l = \mathrm{W}) \geq d^{-1}(d-1)^{-(l-1)} > d^{-l}$$

whence, setting

$$W_\alpha = \{\mathrm{W} \in T_l : |\delta(\mathrm{W}) - \overline{\delta}(\mathrm{W})| > \alpha + 32 \log^2 d\},$$

we have

$$|W_\alpha| \leq \Pr(\mathrm{Y}_l \in W_\alpha) \left(\min_{\mathrm{W} \in T_l} \Pr(\mathrm{Y}_l = \mathrm{W}) \right)^{-1} < 2d^2 t \exp\left(-\frac{\alpha^2}{2t}\right) d^l.$$

Taking $\alpha = d^{1/4+\varepsilon} - 32\log^2 d$ then gives Lemma 2.2 (and a bit more). ∎

The key observation for the proof of (12) is that while $\delta(\mathrm{Y}_l)$ is the number of visits to $N(y_l)$ by $(y_0 \dots y_{l-1})$, $\overline{\delta}(\mathrm{Y}_l)$ is roughly the "expected" number of such visits, where "expected" is used in the dynamic sense given by the function f below. A little martingale analysis then shows that these actual and expected numbers are likely to be close.

For fixed $l \in [t]$ and $w \in V(G)$, define

$$f(w) = \sum_{i=1}^{l} \Pr(y_i \in N(w) | y_0 \dots y_{i-1}), \tag{13}$$

$$g(w) = |N(w) \cap \{y_1 \dots y_l\}|. \tag{14}$$

Lemma 3.1. *For any $\alpha > 0$*

$$\Pr(\exists w \text{ with } |f(w) - g(w)| > \alpha) < 2d^2 t \exp\left(-\frac{\alpha^2}{2t}\right). \tag{15}$$

Remark. The reader may observe below that we only use the fact that $|f(w) - g(w)|$ is usually small when $w = y_l$; but the proof gives the stated inequality, and in fact we don't see how to establish what we need for y_l without proving something like (15).

Before proving Lemma 3.1, let us see why it implies (12). Notice that

$$g(y_l) = \delta(\mathrm{Y}_l). \tag{16}$$

On the other hand, we show that $f(y_l)$ is a good approximation of $\overline{\delta}(\mathrm{Y}_l)$. We have

$$\overline{\delta}(\mathrm{Y}_l) = \frac{1}{s(y_l)} \sum \{|N(u) \cap \{y_0 \dots y_l\}| : u \in N(y_l) \setminus \{y_0 \dots y_{l-1}\}\}, \tag{17}$$

while a similar expression for $f(w)$ is

$$\sum_{i=1}^{l} \frac{1}{s(y_{i-1})} |(N(y_{i-1}) \cap N(w)) \setminus \{y_0 \dots y_{i-2}\}|. \tag{18}$$

Now when $w = y_l$, the sum of the set cardinalities appearing in (18) is not much different than the sum in (17): the former—that is,

$$\sum_{i=1}^{l} |(N(y_{i-1}) \cap N(y_l)) \setminus \{y_0 \dots y_{i-2}\}| \tag{19}$$

—counts ordered pairs (u, y_{i-1}) with $1 \le i \le l$, $y_l \sim u \sim y_{i-1}$, and $u \notin \{y_0 \dots y_{i-2}\}$; whereas the latter counts all such pairs for which $u \notin \{y_0 \dots y_{l-1}\}$, together with the pairs (u, y_l) with $u \in N(y_l) \setminus \{y_0 \dots y_{l-1}\}$.

The difference between these sums is thus bounded by

$$\max\left\{|\{(j,i) : i \le j \le l-1, y_l \sim y_j \sim y_{i-1}\}|,\ d\right\} \le \binom{t}{2},$$

and we have (using (7))

$$\begin{aligned} |f(y_l) - \overline{\delta}(\mathrm{Y}_l)| &\le \frac{1}{s(y_l)}\binom{t}{2} + \sum_{i=1}^{l}\left|\frac{1}{s(y_{i-1})} - \frac{1}{s(y_l)}\right| |N(y_{i-1}) \cap N(y_l)| \\ &\le \frac{1}{d-t}\binom{t}{2} + \left(\frac{1}{d-l} - \frac{1}{d}\right) ld \\ &< 2t^2 d^{-1} \quad \le \quad 32 \log^2 d. \end{aligned}$$

Of course this together with (16) shows that Lemma 3.1 implies (12). ∎

Proof of Lemma 3.1. Let us for the moment fix $w \in V$ and write

$$f(w) - g(w) = \sum_{i=1}^{l} X_i$$

where

$$X_i = X_i(w) = \Pr(y_i \in N(w) | y_0 \dots y_{i-1}) - 1_{\{y_i \in N(w)\}}.$$

Now $\{X_i\}_{i=1}^{l}$ is a martingale difference sequence (that is, $E[X_i | X_1 \dots X_{i-1}] = 0$), with

$$|X_i| \le 1. \tag{20}$$

So according to "Azuma's inequality" (see, e.g., [4], [25], [3]), for any $\alpha > 0$,

$$\Pr(|f(w) - g(w)| > \alpha) < 2\exp\left(-\frac{\alpha^2}{2t}\right). \tag{21}$$

Thus we have a bound like (15) for any fixed w.

For (15) we must somehow control the number of w's under consideration. *A priori* this number could be something like the number of vertices within distance t of v (which swamps the bound in (21)); but we can reduce it by only beginning to keep track of $f(w) - g(w)$ when (and if) our random walk gets to within distance 2 of w.

To do this, let us fix, solely for bookkeeping purposes, some linear ordering "$\prec$" of V. For each $w \in V$ define the random variable $j(w)$ by

$$j(w) = \begin{cases} 0 & \text{if } d(v,w) \leq 2 \\ \infty & \text{if } d(y_i,w) > 2 \quad 0 \leq i \leq t-1 \\ \min\{i : d(y_i,w) = 2\} & \text{otherwise,} \end{cases}$$

and then let v_s be the s^{th} vertex in the (lexicographic) ordering in which w precedes w' if either $j(w) < j(w')$ or $j(w) = j(w')$ and $w \prec w'$. (Note this is a random ordering determined by $(y_0 \ldots y_t)$.)

Now for $1 \leq s \leq d^2 t$ and $1 \leq i \leq t$, set

$$X_i^s = X_i(v_s).$$

(Note $X_i^s = 0$ if $i \leq j(v_s)$. We could omit the restriction $s \leq d^2 t$, but this adds nothing since for larger s we have $j(v_s) = \infty$ and so $X_i^s = 0$ for all i.)

Now for each fixed s, $f(v_s) - g(v_s) = \sum_{i=1}^{l} X_i^s$, and $\{X_i^s\}_{i=1}^{l}$ is again a martingale difference sequence satisfying (20). Thus

$$\Pr(|f(v_s) - g(v_s)| > \alpha) < 2\exp\left(-\frac{\alpha^2}{2t}\right)$$

for each s, and

$$\Pr(\exists s \in [d^2 t], \quad |f(v_s) - g(v_s)| > \alpha) < 2d^2 t \exp\left(-\frac{\alpha^2}{2t}\right).$$

But this gives (15), since (trivially) $f(v_s) = g(v_s) = 0$ if $s > d^2 t$. ∎

4. Proof of Theorem 1.2(a)

As mentioned in Section 2, we use $T := T(G,v)$ and $p_T(\overline{\mathrm{V}})$ to estimate $p(\overline{v})$. In a sense, Lemma 2.2 says that degree fluctuations in T are much more moderate than the extreme case (9); namely,

$$|s(\mathrm{W}) - \overline{s}(\mathrm{W})| \leq d^{1/4+\varepsilon} \tag{22}$$

for "almost all" $\mathrm{W} \in V(T)$. If stronger conditions

$$d - |\mathrm{W}| \leq s(\mathrm{W}) \leq d - |\mathrm{W}| + d^{1/4+\varepsilon} \quad \text{for all } \mathrm{W} \in V(T) \tag{23}$$

or

$$d - d^{1/4+\varepsilon} \leq s(\mathrm{W}) \leq d \quad \text{for all } \mathrm{W} \in V(T)$$

were true, Theorem 1.2(a) would be straightforward. (For example, a complete graph satisfies (23); cf. (7).) This is because the worst cases would be essentially

$$s(\mathrm{W}) = \begin{cases} d - |\mathrm{W}| & \text{if } |\mathrm{W}| \text{ is even} \\ d - |\mathrm{W}| + d^{1/4+\varepsilon} & \text{if } |\mathrm{W}| \text{ is odd.} \end{cases} \tag{24}$$

It turns out that (22) yields nothing worse than (24). If $\mathrm{W} \in V(T)$ satisfies (22) and all of its grandchildren X satisfy $p_{T(\mathrm{x})}(\overline{\mathrm{X}}) \le p$ for some constant p, then the identity (6) gives

$$\begin{aligned} p_{T(\mathrm{w})}(\overline{\mathrm{W}}) &= \left(1 + \sum_{\mathrm{U} \in S(\mathrm{W})} p_{T(\mathrm{u})}(\overline{\mathrm{U}})\right)^{-1} \\ &= \left(1 + \sum_{\mathrm{U} \in S(\mathrm{W})} \left(1 + \sum_{\mathrm{X} \in S(\mathrm{U})} p_{T(\mathrm{x})}(\overline{\mathrm{X}})\right)^{-1}\right)^{-1} \\ &\le \left(1 + \sum_{\mathrm{U} \in S(\mathrm{W})} \frac{1}{1 + s(\mathrm{U})p}\right)^{-1} \end{aligned}$$

and Jensen's inequality, (22) and $s(\mathrm{W}) \ge d - |\mathrm{W}|$ imply that

$$\begin{aligned} p_{T(\mathrm{w})}(\overline{\mathrm{W}}) &\le \left(1 + \frac{s(\mathrm{W})}{1 + \overline{s}(\mathrm{W})p}\right)^{-1} \\ &\le \left(1 + \frac{s(\mathrm{W})}{1 + (s(\mathrm{W}) + d^{1/4+\varepsilon})p}\right)^{-1} \\ &\le \left(1 + \frac{d - |\mathrm{W}|}{1 + (d - |\mathrm{W}| + d^{1/4+\varepsilon})p}\right)^{-1}. \end{aligned} \tag{25}$$

A similar lower bound can be also found. Notice that the equalities hold when $p_{T(\mathrm{x})}(\overline{\mathrm{X}}) = p$ for all X and T is the tree described in (24).

Set $p_t = p_{t-1} = 1$ and for $i = 0, 1 \ldots t-2$,

$$p_i = \left(1 + \frac{d - i}{1 + (d - i + d^{1/4+\varepsilon})p_{i+2}}\right)^{-1}.$$

We first show that p_i is close to $1/\sqrt{d}$ for $i \le t/2 + 1$, and then that the effect of vertices violating (22) is negligible, so that $p_{T(\mathrm{w})}(\overline{\mathrm{W}})$ is close to $p_{|\mathrm{w}|}$ for most W.

Claim.

$$p_i = \frac{1}{\sqrt{d}} + O(d^{-3/4+\varepsilon}) \quad \textit{for } 0 \le i \le t/2 + 1\,. \tag{26}$$

Proof. We prove this for i even; odd i is handled similarly. For $i=0,2,\cdots,t$, define

$$f_i(x) := \left(1+\frac{d-i}{1+(d-i+d^{1/4+\varepsilon})x}\right)^{-1}, \qquad x>0$$

(so $f_i(p_{i+2})=p_i$) and denote by a_i the unique positive solution of $f_i(a_i)=a_i$; that is,

$$a_i := \frac{d^{1/4+\varepsilon}-1+\sqrt{(d^{1/4+\varepsilon}-1)^2+4(d-i+d^{1/4+\varepsilon})}}{2(d-i+d^{1/4+\varepsilon})}.$$

It is easy to check that

$$a_i = \frac{1}{\sqrt{d}}+O(d^{-3/4+\varepsilon}) \quad \text{and} \quad 0<a_{i+2}-a_i\le d^{-3/2}.$$

We will prove the following inequalities using induction in reverse order:

$$0\le p_i-a_i\le \exp\left(-\frac{3(t-i)}{4\sqrt{d}}\right)+\frac{(t-i)d^{-3/2}}{2}, \tag{27}$$

for $i=t,t-2,\cdots,0$. (Recall that $t=4\lfloor\sqrt{d}\log d\rfloor$.)

The base case $i=t$ is trivial. Suppose (27) is true for $i+2\le t$. Since f_i is increasing, the lower bound of the induction hypothesis gives

$$a_i=f_i(a_i)\le f_i(a_{i+2})\le f_i(p_{i+2})=p_i.$$

On the other hand, the Mean Value Theorem implies that there exists x with $a_i\le x\le p_{i+2}$ and such that

$$p_i-a_i=f_i(p_{i+2})-f_i(a_i)=f_i'(x)(p_{i+2}-a_i).$$

Since $f_i'(z)\le e^{-3/(2\sqrt{d})}$ for $z\ge a_i=1/\sqrt{d}+O(d^{-3/4+\varepsilon})$, we have

$$\begin{aligned} p_i-a_i &\le e^{-3/(2\sqrt{d})}\left(\exp\left(-\frac{3(t-i-2)}{4\sqrt{d}}\right)+\frac{(t-i-2)d^{-3/2}}{2}+d^{-3/2}\right)\\ &\le \exp\left(-\frac{3(t-i)}{4\sqrt{d}}\right)+\frac{(t-i)d^{-3/2}}{2}. \end{aligned}$$

∎

If a vertex does not have many descendants which violate (22), then we only need a small modification of the upper bound in (25) (see (31)). If it has many, it will be called a “bad” vertex. To be formal, let $0<\varepsilon<0.1$ (fixed) and $\Gamma_i:=\Gamma(i,\varepsilon)$. Then Lemma 2.2 says that

$$|\Gamma_i|<t^{-1}(d-t)^i e^{-d^{\varepsilon}}. \tag{28}$$

Each vertex of T_i has at least $(d-t)^{j-i}$ descendants in T_j. Roughly speaking, (28) says that at most a $t^{-1}e^{-d^\varepsilon}$ proportion of the descendants of an average vertex violate (22). If a vertex W has more than $t^{-1}e^{-d^\varepsilon/2} \times (d-t)^{j-|\mathrm{W}|}$ descendants violating (22) for some $j \geq |\mathrm{W}|$, it is called bad. More precisely, let, for $0 \leq i \leq j \leq t$,

$$(29)\quad B_{i,j} := \{\mathrm{W} \in T_i : |T(\mathrm{W}) \cap \Gamma_j| \geq t^{-1}(d-t)^{j-i}e^{-d^\varepsilon/2}\} \quad \text{and} \quad B := \bigcup_{i=0}^{t} \bigcup_{j=i}^{t} B_{i,j}$$

(B for bad). Clearly, (28) implies that neither v nor any of its children is in B. Moreover, it is easy to see that

$$(30)\qquad |B \cap T_i| \leq \sum_{j=i}^{t} |B_{i,j}| \leq \sum_{j=i}^{t} \frac{|\Gamma_j|}{t^{-1}(d-t)^{j-i}e^{-d^\varepsilon/2}} \leq t(d-t)^i e^{-d^\varepsilon/2}.$$

(We will not use this inequality until the last part of the next section.)

The following lemma and its corollary show that the effect of bad vertices is negligible, which in particular implies Theorem 1.2(a). They will also be used in the proof of Theorem 1.2(b).

Lemma 4.1. *Let $0 \leq i \leq t-1$ and $\mathrm{X} \in T_i \setminus B$. Then*

$$(31)\qquad p_{T(\mathrm{X})}(\overline{\mathrm{X}}) \leq p_i + e^{-d^\varepsilon/2},$$

and

$$(32)\qquad p_{T(\mathrm{X})}(\overline{\mathrm{W}}) \leq p_{i+1} + (d-t)e^{-d^\varepsilon/2} \quad \textit{for all } \mathrm{W} \in S(\mathrm{X}).$$

Note that (6) and (32) yield

$$p_{T(\mathrm{X})}(\overline{\mathrm{X}}) \geq \Big(1 + d(p_{i+1} + (d-t)e^{-d^\varepsilon/2})\Big)^{-1} \quad \text{for } \mathrm{X} \in T_i \setminus B\,.$$

Thus the following corollary follows from Lemma 4.1 and (26).

Corollary 4.2. *Let $0 \leq i \leq t/2$ and $\mathrm{X} \in T_i \setminus B$. Then*

$$\left| p_{T(\mathrm{X})}(\overline{\mathrm{X}}) - \frac{1}{\sqrt{d}} \right| = O(d^{-3/4+\varepsilon}).$$

In particular, we have Theorem 1.2(a).

We prove (31) for $|\mathrm{X}|$ even (recall that t is even) and (32) for $|\mathrm{X}|$ odd. The proof when these parities are reversed is identical, except that one truncates at odd rather than even levels in the following definition of T'.

If a descendant U of X violates (22) (i.e. $\mathrm{U}\in\Gamma_i$ for $i=|\mathrm{U}|$) and $|\mathrm{U}|$ is even, the trivial upper bound $p_{T(\mathrm{u})}(\overline{\mathrm{U}})\le 1$ will be used. We also neglect vertices U of level $|\mathrm{U}|>t$ and use the trivial bound for vertices at level t. To do this, it is convenient to introduce an auxiliary subtree $T'(\mathrm{X})$ of $T(\mathrm{X})$ obtained by removing all descendants of even vertices violating (22) and all vertices at levels greater than t. In this subtree, leaves are vertices violating (22) or at level t. Notice that the trees $T'(\mathrm{U})$ generated by leaves U do not have edges and so $p_{T'(\mathrm{u})}(\overline{\mathrm{U}})=1$, which was the bound we wanted. Clearly (6) implies that for every W of even level

$$p_{T(\mathrm{w})}(\overline{\mathrm{W}})\le p_{T'(\mathrm{w})}(\overline{\mathrm{W}}),$$

In particular,

$$p_{T(\mathrm{x})}(\overline{\mathrm{X}})\le p_{T'(\mathrm{x})}(\overline{\mathrm{X}}). \tag{33}$$

Inductive applications of the arguments used to obtain (25) will yield the following lemma.

Lemma 4.3. *For all* $\mathrm{W}\in T'(\mathrm{X})$ *with* $|\mathrm{W}|$ *even, we have*

$$p_{T'(\mathrm{w})}(\overline{\mathrm{W}})\le p_{|\mathrm{w}|}+\sum_{\mathrm{U}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{U}|+|\mathrm{W}|}(1-p_{|\mathrm{U}|}), \tag{34}$$

where $L'(\mathrm{W})$ *is the set of leaves of* $T'(\mathrm{W})$. *(The vertex of a singleton tree is regarded as a leaf.)*

Proof. Let $q'(\overline{\mathrm{W}})=p_{T'(\mathrm{w})}(\overline{\mathrm{W}})$. If W is a leaf of T', then the right side of (34) is $p_{T'(\mathrm{w})}(\overline{\mathrm{W}})+1-p_{T'(\mathrm{w})}(\overline{\mathrm{W}})=1$, so the result follows. Suppose W is not a leaf, $|\mathrm{W}|$ is even and (34) is true for all descendants of W with even levels. Then (6) and Jensen's inequality give

$$\begin{aligned}
q(\overline{\mathrm{W}}) &= \left(1+\sum_{\mathrm{U}\in S'(\mathrm{W})} q(\overline{\mathrm{U}})\right)^{-1}\\
&= \left(1+\sum_{\mathrm{U}\in S'(\mathrm{W})}\frac{1}{1+\sum_{\mathrm{X}\in S'(\mathrm{U})} q(\overline{\mathrm{X}})}\right)^{-1}\\
&\le \left(1+\frac{s'(\mathrm{W})}{1+s'(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S'(\mathrm{W})}\sum_{\mathrm{X}\in S'(\mathrm{U})} q(\overline{\mathrm{X}})}\right)^{-1}.
\end{aligned}$$

Clearly, for a non-leaf W and its children U, $S'(\mathrm{W})=S(\mathrm{W})$, $S'(\mathrm{U})=S(\mathrm{U})$ and hence $s'(\mathrm{W})=s(\mathrm{W})$, $s'(\mathrm{U})=s(\mathrm{U})$. The induction hypothesis yields

$$s(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S(\mathrm{W})}\sum_{\mathrm{X}\in S(\mathrm{U})} q(\overline{\mathrm{X}}) \le$$

$$s(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S(\mathrm{W})}\sum_{\mathrm{X}\in S(\mathrm{U})} p_{|\mathrm{X}|}+s(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S(\mathrm{W})}\sum_{\mathrm{X}\in S(\mathrm{U})}\sum_{\mathrm{Z}\in L'(\mathrm{X})}(d-t)^{-|\mathrm{Z}|+|\mathrm{X}|}(1-p_{|\mathrm{Z}|}).$$

The first term of this bound appeared when (25) was derived. That is,

$$s(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S(\mathrm{W})}\sum_{\mathrm{X}\in S(\mathrm{U})} p_{|\mathrm{X}|} =$$

$$s(\mathrm{W})^{-1}\sum_{\mathrm{U}\in S(\mathrm{W})} s(\mathrm{U})p_{|\mathrm{X}|} = \overline{s}(\mathrm{W})p_{|\mathrm{W}|+2} \le (s(\mathrm{W}) + d^{1/4+\varepsilon})p_{|\mathrm{W}|+2},$$

where the inequality uses the fact that W is not a leaf. The second term is nothing but

$$s(\mathrm{W})^{-1}\sum_{\mathrm{Z}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{Z}|+|\mathrm{W}|+2}(1-p_{|\mathrm{Z}|}) \le \sum_{\mathrm{Z}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{Z}|+|\mathrm{W}|+1}(1-p_{|\mathrm{Z}|}).$$

We now use the easy inequality

$$\left(1+\frac{\alpha}{\beta+x}\right)^{-1} \le \left(1+\frac{\alpha}{\beta}\right)^{-1} + \alpha^{-1}x \quad \text{for all } \alpha,\beta,x>0,$$

to obtain

$$\begin{aligned} q(\overline{\mathrm{W}}) &\le \left(1+\frac{s(\mathrm{W})}{1+(s(\mathrm{W})+d^{1/4+\varepsilon})p_{|\mathrm{W}|+2}}\right)^{-1} \\ &\quad + s(\mathrm{W})^{-1}\sum_{\mathrm{Z}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{Z}|+|\mathrm{W}|+1}(1-p_{|\mathrm{Z}|}) \\ &\le \left(1+\frac{d-|\mathrm{W}|}{1+(d-|\mathrm{W}|+d^{1/4+\varepsilon})p_{|\mathrm{W}|+2}}\right)^{-1} \\ &\quad + \sum_{\mathrm{Z}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{Z}|+|\mathrm{W}|}(1-p_{|\mathrm{Z}|}) \\ &= p_{|\mathrm{W}|} + \sum_{\mathrm{Z}\in L'(\mathrm{W})}(d-t)^{-|\mathrm{Z}|+|\mathrm{W}|}(1-p_{|\mathrm{Z}|}). \end{aligned}$$

∎

Proof of (31). Since Lemma 4.3 and (33) give

$$q(\overline{\mathrm{X}}) \le p_{|\mathrm{X}|} + \sum_{\mathrm{U}\in L'(\mathrm{X})} (d-t)^{-|\mathrm{U}|+|\mathrm{X}|}(1-p_{|\mathrm{U}|}),$$

it is enough to show that

$$\sum_{\mathrm{U}\in L'(\mathrm{X})} (d-t)^{-|\mathrm{U}|+|\mathrm{X}|}(1-p_{|\mathrm{U}|}) \le e^{-d^{\varepsilon}/2}.$$

But $\mathrm{X}\notin B$ and $p_t=1$ imply that

$$\sum_{\mathrm{U}\in L'(\mathrm{X})} (d-t)^{-|\mathrm{U}|+|\mathrm{X}|}(1-p_{|\mathrm{U}|}) \le \sum_{\substack{j:\text{ even}\\ |\mathrm{X}|\le j\le t-2}} \sum_{\mathrm{U}\in T(\mathrm{X})\cap\Gamma_j} (d-t)^{-j+|\mathrm{X}|}$$

$$\le \sum_{\substack{j:\text{ even}\\ |\mathrm{X}|\le j\le t-2}} t^{-1}(d-t)^{j-|\mathrm{X}|}e^{-d^{\varepsilon}/2}(d-t)^{-j+|\mathrm{X}|} \le e^{-d^{\varepsilon}/2}. \qquad \blacksquare$$

The proof of (32) is the same as that of (31), except we use

$$|T(\mathrm{W})\cap\Gamma_j| \le t^{-1}(d-t)^{j-|\mathrm{W}|+1}e^{-d^{\varepsilon}/2}$$

(which follows from $\mathrm{W}\in S(\mathrm{X})$ and $\mathrm{X}\notin B$).

5. Proof of Theorem 1.2(b)

First we note that

$$\sigma^2(G) = \mathrm{Var}\left[\frac{1}{2}\left(n - \sum_{v\in V(G)} 1_{\{v\not\sim M\}}\right)\right] = \frac{1}{4}\mathrm{Var}\left[\sum_{v\in V(G)} 1_{\{v\not\sim M\}}\right]$$

and clearly

$$\mathrm{Var}\left[\sum_{v\in V(G)} 1_{\{v\not\sim M\}}\right] = \sum_{v,w\in V(G)} (p_G(\overline{v},\overline{w}) - p_G(\overline{v})p_G(\overline{w}))$$

$$\le \sum_{v\in V(G)} p_G(\overline{v}) + \sum_{v\in V(G)} p_G(\overline{v}) \sum_{w\in V(G)\setminus\{v\}} (p_G(\overline{w}|\overline{v}) - p_G(\overline{w}))\,.$$

Set

$$I(G,v) := \sum_{w\in V(G)\setminus\{v\}} (p_G(\overline{w}|\overline{v}) - p_G(\overline{w}))\,.$$

Then Theorem 1.2(a) gives

$$\sigma^2(G) \le \left(1 + O(d^{-1/4+\varepsilon})\right) \frac{n}{\sqrt{d}} + \sum_{v \in V(G)} p(\overline{v}) I(G, v).$$

Theorem 1.2(b) will follow from the fact that

$$|p(\overline{v}) I(G, v)| \le 2d^{-3/4+4\varepsilon} \qquad \text{for all } v \in V(G). \tag{35}$$

Of course (35) is a concrete expression of the idea that the indicators $1_{\{v \not< M\}}$ are close to independent (compare the discussion in the vicinity of (2) of the limiting values in Theorem 1.2).

Central to our argument are the quantities $r_T(\overline{\mathrm{W}})$ which are the product of $p_{T(\mathrm{u})}(\overline{\mathrm{U}})$ over all ancestors U of W, including W itself:

$$r_T(\overline{\mathrm{W}}) := p_{T(\mathrm{w})}(\mathrm{W}) \cdot \prod_{\substack{\mathrm{U}:\text{ancestor} \\ \text{of } \mathrm{W}}} p_{T(\mathrm{u})}(\overline{\mathrm{U}}).$$

We know by (6) that

$$1 = p_T(\overline{\mathrm{V}}) \left(1 + \sum_{\mathrm{W} \in S(\mathrm{V})} p_{T(\mathrm{w})}(\overline{\mathrm{W}})\right)$$

and multiplying both sides by $p_T(\overline{\mathrm{V}})$ yields

$$p_T(\overline{\mathrm{V}}) = p_T^2(\overline{\mathrm{V}}) + p_T^2(\overline{\mathrm{V}}) \sum_{\mathrm{W} \in S(\mathrm{V})} p_{T(\mathrm{w})}(\overline{\mathrm{W}}). \tag{36}$$

The next lemma follows by inductive application of this argument.

Lemma 5.1.

$$p_T(\overline{\mathrm{V}}) = \sum_{\mathrm{W} \in V(T)} r_T^2(\overline{\mathrm{W}}).$$

Proof. We just use (36) and induction:

$$\begin{aligned} p_T(\overline{\mathrm{V}}) &= p_T^2(\overline{\mathrm{V}}) + p_T^2(\overline{\mathrm{V}}) \sum_{\mathrm{W} \in S(\mathrm{V})} p_{T(\mathrm{w})}(\overline{\mathrm{W}}) \\ &= p_T^2(\overline{\mathrm{V}}) + p_T^2(\overline{\mathrm{V}}) \sum_{\mathrm{W} \in S(\mathrm{V})} \sum_{\mathrm{U} \in T(\mathrm{W})} r_{T(\mathrm{w})}^2(\overline{\mathrm{U}}). \end{aligned}$$

Since $p_T^2(\overline{\mathrm{V}}) = r_T^2(\overline{\mathrm{V}})$ and $p_T^2(\overline{\mathrm{V}}) r_{T(\mathrm{w})}^2(\overline{\mathrm{U}}) = r_T(\overline{\mathrm{U}})$ for $\mathrm{W} \in S(\mathrm{V})$, the lemma follows. ∎

Since $p_T(\overline{\mathrm{V}}) = p_G(\overline{v})$, this lemma implies that

$$p_G(\overline{v}) = \sum_{\mathrm{W} \in V(T)} r^2(\overline{\mathrm{W}}). \tag{37}$$

The quantity $p_G(\overline{v}) I(G, v)$ (see (35)) turns out to be an alternating sum of the $r_T^2(\overline{\mathrm{W}})$'s:

Lemma 5.2.

$$p_G(\overline{v}) I(G, v) = \sum_{\mathrm{W} \in V(T) \setminus \{\mathrm{V}\}} (-1)^{|\mathrm{W}|-1} r^2(\overline{\mathrm{W}}).$$

Proof. The proof will be based on the recursive relations

$$I(G, v) = - \sum_{y \in N_G(v)} p_T(\overline{v}, \overline{y}) I(G \setminus v, y) \; + \sum_{y \in N_G(v)} p_G(\overline{v}, \overline{y}) p_G(\overline{y} | \overline{v}), \tag{38}$$

where $G \setminus v$ is the subgraph of G induced by $V(G) \setminus \{v\}$.

To prove (38) notice that if the event $\{v \prec M\}$ occurs, then there must be a unique $y \in N_G(v)$ such that the event $\{\{v, y\} \in M\}$ occurs. Hence

$$p_G(\overline{v}) + \sum_{y \in N(G)} p_G(\{y, v\} \in M) = 1, \tag{39}$$

and

$$p_G(\overline{w}) = p_G(\overline{v}) p_G(\overline{w} | \overline{v}) + \sum_{y \in N_G(v)} p_G(\{v, y\} \in M) p_G(\overline{w} | \{v, y\} \in M).$$

Using

$$p_G(\overline{w} | \overline{v}) = p_G(\overline{w} | \overline{v}) \left(p_G(\overline{v}) + \sum_{y \in N_G(v)} p_G(\{v, y\} \in M) \right)$$

we have

$$p_G(\overline{w} | \overline{v}) - p_G(\overline{w}) = \sum_{y \in N_G(v)} p_G(\{v, y\} \in M) \left(p_G(\overline{w} | \overline{v}) - p_G(\overline{w} | \{v, y\} \in M) \right).$$

Furthermore, since there is a bijection between the set all matchings containing an edge $\{v, y\}$ and the set of all matchings containing no edge incident to v or y, we have

$$p_G(\{v, y\} \in M) = p_G(\overline{v}, \overline{y}),$$

and

$$p_G(\overline{w} | \{v, y\} \in M) = \begin{cases} p_G(\overline{w} | \overline{y}, \overline{v}) & \text{if } w \notin \{v, y\} \\ 0 & \text{if } w \in \{v, y\}. \end{cases}$$

Thus

$$p_G(\overline{w}|\overline{v}) - p_G(\overline{w}) = \sum_{y\in N_G(v)} p_G(\overline{v},\overline{y})\left(p_G(\overline{w}|\overline{v}) - p_G(\overline{w}|\{v,y\}\in M)\right)$$

and

$$\begin{aligned} I(G,v) &= \sum_{w\in V(G)\setminus\{v\}}\ \sum_{y\in N_G(v)} p_G(\overline{v},\overline{y})\left(p_G(\overline{w}|\overline{v}) - p_G(\overline{w}|\{v,y\}\in M)\right) \\ &= \sum_{y\in N_G(v)} p_G(\overline{v},\overline{y}) \sum_{w\in V(G)\setminus\{v\}} \left(p_G(\overline{w}|\overline{v}) - p_G(\overline{w}|\{v,y\}\in M)\right) \\ &= \sum_{y\in N_G(v)} p_G(\overline{v},\overline{y}) \sum_{w\in V(G)\setminus\{y,v\}} \left(p_G(\overline{w}|\overline{v}) - p_G(\overline{w}|\overline{y},\overline{v})\right) \\ &\quad + \sum_{y\in N_G(v)} p_G(\overline{v},\overline{y}) p_G(\overline{y}|\overline{v}). \end{aligned}$$

Since $p_G(\overline{w}|\overline{v}) = p_{G\setminus v}(\overline{w})$ and $p_G(\overline{w}|\overline{v},\overline{y}) = p_{G\setminus v}(\overline{w}|\overline{y})$,

$$\sum_{w\in V(G)\setminus\{y,v\}} \left(p_G(\overline{w}|\overline{v}) - p_G(\overline{w}|\overline{y},\overline{v})\right) = -I(G\setminus v, y),$$

and (38) follows.

Suppose now that the lemma is true for $G\setminus y$, $y\in N_G(v)$. Then $p_G(\overline{y}|\overline{v}) = p_{G\setminus v}(\overline{y})$, $y\in N_G(v)$ and the induction hypothesis imply that

$$\begin{aligned} &p_G(\overline{v},\overline{y}) I(G\setminus v, y) = p_G(\overline{v}) p_{G\setminus v}(\overline{y}) I(G\setminus v, y) \\ &= p_G(\overline{v}) \sum_{\mathrm{W}\in V(T(\mathrm{Y}))\setminus\{\mathrm{Y}\}} (-1)^{|\mathrm{W}|_{\mathrm{y}}-1} r^2_{T(\mathrm{y})}(\overline{\mathrm{W}}), \end{aligned}$$

where $|\mathrm{W}|_{\mathrm{y}}$ is the level of W in $T(\mathrm{Y})$, which is $|\mathrm{W}|-1$. Thus (using $p_G(\overline{v},\overline{y})p_G(\overline{y}|\overline{v}) = p_G(\overline{v})p^2_G(\overline{y}|\overline{v})$)

(40) $$p_G(\overline{v}) I(G,v)$$
$$= -p^2_G(\overline{v}) \sum_{y\in N_G(v)}\ \sum_{\mathrm{W}\in V(T(\mathrm{Y}))\setminus\{\mathrm{Y}\}} (-1)^{|\mathrm{W}|-2} r^2_{T(\mathrm{y})}(\overline{\mathrm{W}}) + \sum_{y\in N_G(v)} p^2_G(\overline{v}) p^2_G(\overline{y}|\overline{v}).$$

For $y\in N_G(v)$, clearly

$$p^2_G(\overline{v}) r^2_{T(\mathrm{y})}(\overline{\mathrm{W}}) = p^2_T(\overline{\mathrm{V}}) r^2_{T(\mathrm{y})}(\overline{\mathrm{W}}) = r^2_T(\overline{\mathrm{W}}).$$

Using $p_G(\overline{y}|\overline{v}) = p_{T(\mathrm{y})}(\overline{\mathrm{Y}})$ we also know that

$$p^2_G(\overline{v}) p^2_G(\overline{y}|\overline{v}) = p^2_T(\overline{\mathrm{V}}) p^2_{T(\mathrm{y})}(\overline{\mathrm{Y}}) = r^2_T(\overline{\mathrm{Y}}).$$

So altogether we have

$$p_G(\overline{v})I(G,v) = \sum_{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}} (-1)^{|\mathrm{W}|-1} r^2(\overline{\mathrm{W}}).$$

∎

To complete the proof of (35), and therefore Theorem 1.2(b), we show tight lower bounds on the even and odd sums of the $r_T^2(\overline{\mathrm{W}})$'s:

$$\sum_{\substack{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}\\ |\mathrm{W}|\ \text{even}}} r_T^2(\overline{\mathrm{W}}) \geq \frac{p_G(\overline{v})(1-d^{-1/4+4\varepsilon})}{2} \tag{41}$$

and

$$\sum_{\substack{\mathrm{W}\in V(T)\\ |\mathrm{W}|\ \text{odd}}} r_T^2(\overline{\mathrm{W}}) \geq \frac{p_G(\overline{v})(1-d^{-1/4+4\varepsilon})}{2}. \tag{42}$$

(We wind up with the (4ε)'s because we often use extra factors d^{ε} to subsume smaller but clumsier error terms.) These inequalities together with Theorem 1.2(a) and (37) imply that

$$\begin{aligned}\sum_{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}} (-1)^{|\mathrm{W}|-1} r_T^2(\overline{\mathrm{W}}) &= \sum_{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}} r_T^2(\overline{\mathrm{W}}) - 2\sum_{\substack{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}\\ |\mathrm{W}|\ \text{even}}} r_T^2(\overline{\mathrm{W}}) \\ &\leq p_G(\overline{v}) - 2\cdot\frac{p_G(\overline{v})(1-d^{-1/4+4\varepsilon})}{2} \leq 2d^{-3/4+4\varepsilon}\end{aligned}$$

and similarly

$$\sum_{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}} (-1)^{|\mathrm{W}|-1} r_T^2(\overline{\mathrm{W}}) \geq -2d^{-3/4+4\varepsilon}.$$

So (35) follows.

The proofs of (41) and (42) are essentially identical, so we prove only (41). Recall $s(\mathrm{W})$ is the number of chidren of W in T. Let $\mathrm{W}\in V(T)$ be of even level. Denote by $a(\mathrm{W})$ the product of $s(\mathrm{U})^{-1}$ over all even ancestors of W; that is,

$$a(\mathrm{W}) = \left(\prod_{\substack{\mathrm{U}:\ \text{ancestors of W}\\ |\mathrm{U}|\ \text{even}}} s(\mathrm{U})\right)^{-1}$$

($a(\mathrm{V}) := 1$). Of course, we should really use $a_T(\mathrm{W})$. The subscript T is ommitted for simplicity. We also write $r(\mathrm{W})$ for $r_T(\mathrm{W})$.

Now consider $\left(\sum_{\mathrm{W}\in T_l} a(\mathrm{W})r(\overline{\mathrm{W}})\right)^2$ and apply the Cauchy–Schwarz inequality to obtain

$$\left(\sum_{\mathrm{W}\in T_l} a(\mathrm{W})r(\overline{\mathrm{W}})\right)^2 \leq \sum_{\mathrm{W}\in T_l} a^2(\mathrm{W}) \sum_{\mathrm{W}\in T_l} r^2(\overline{\mathrm{W}}). \tag{43}$$

The next two claims give tight lower and upper bounds (respectively) on the left hand side of (43) and the first term on the right hand side for even l, which will yield the desired lower bound on $\sum_{\mathrm{W}\in T_l} r^2(\overline{\mathrm{W}})$ for even l.

Claim 1.

$$\sum_{\mathrm{W}\in T_l} a(\mathrm{W})r(\overline{\mathrm{W}}) \geq p_G(\overline{v})\left(1 - d^{-1/2} - 2d^{-3/4+2\varepsilon}\right)^{l/2} \qquad \text{for even } l \leq t/2 - 2.$$

Proof. We show by induction that

$$\sum_{\mathrm{W}\in T_l} a(\mathrm{W})r(\overline{\mathrm{W}}) \geq p_G(\overline{v})\left(1 - d^{-1/2} - d^{-3/4+2\varepsilon}\right)^{l/2} - \frac{l \cdot e^{-d^{\varepsilon}/5}}{2}.$$

The base case $l=0$ is trivial. For $l>0$ note that for $\mathrm{W}\in T_{l+2}$ and its parent U and grandparent X,

$$a(\mathrm{W}) = a(\mathrm{X})s^{-1}(\mathrm{X}) \quad \text{and} \quad r(\overline{\mathrm{W}}) = r(\overline{\mathrm{X}})p_{T(\mathrm{u})}(\overline{\mathrm{U}})p_{T(\mathrm{w})}(\overline{\mathrm{W}}).$$

Hence

$$\begin{aligned}\sum_{\mathrm{W}\in T_{l+2}} a(\mathrm{W})r(\overline{\mathrm{W}}) &= \sum_{\mathrm{X}\in T_l}\sum_{\mathrm{U}\in S(\mathrm{X})}\sum_{\mathrm{W}\in S(\mathrm{U})} a(\mathrm{W})r(\overline{\mathrm{W}}) \\ &= \sum_{\mathrm{X}\in T_l} a(\mathrm{X})r(\overline{\mathrm{X}})s^{-1}(\mathrm{X}) \sum_{\mathrm{U}\in S(\mathrm{X})} p_{T(\mathrm{u})}(\overline{\mathrm{U}}) \sum_{\mathrm{W}\in S(\mathrm{U})} p_{T(\mathrm{w})}(\overline{\mathrm{W}}).\end{aligned}$$

Applying (6) we have

$$p_{T(\mathrm{u})}(\overline{\mathrm{U}}) \sum_{\mathrm{W}\in S(\mathrm{U})} p_{T(\mathrm{w})}(\overline{\mathrm{W}}) = 1 - p_{T(\mathrm{u})}(\overline{\mathrm{U}})$$

and

$$\sum_{\mathrm{W}\in T_{l+2}} a(\mathrm{W})r(\overline{\mathrm{W}}) = \sum_{\mathrm{X}\in T_l} a(\mathrm{X})r(\overline{\mathrm{X}})s^{-1}(\mathrm{X}) \sum_{\mathrm{U}\in S(\mathrm{X})} (1 - p_{T(\mathrm{u})}(\overline{\mathrm{U}})).$$

For $\mathrm{X} \in T_l \setminus B$ and $\mathrm{U} \in S(\mathrm{X})$, we know by (26) and (32) that

$$p_{T(\mathrm{u})}(\overline{\mathrm{U}}) \le d^{-1/2} + d^{-3/4+2\varepsilon}.$$

Therefore

$$\begin{aligned}\sum_{\mathrm{W} \in T_{l+2}} a(\mathrm{W}) r(\overline{\mathrm{W}}) &\ge \sum_{\mathrm{X} \in T_l \setminus B} a(\mathrm{X}) r(\overline{\mathrm{X}}) s^{-1}(\mathrm{X}) \sum_{\mathrm{U} \in S(\mathrm{X})} (1 - p_{T(\mathrm{u})}(\overline{\mathrm{U}})) \\ &\ge \left(1 - d^{-1/2} - d^{-3/4+2\varepsilon}\right) \sum_{\mathrm{X} \in T_l \setminus B} a(\mathrm{X}) r(\overline{\mathrm{X}}) \\ &\ge \left(1 - d^{-1/2} - d^{-3/4+2\varepsilon}\right) \sum_{\mathrm{X} \in T_l} a(\mathrm{X}) r(\overline{\mathrm{X}}) - \sum_{\mathrm{X} \in B \cap \Gamma_l} a(\mathrm{X}) r(\overline{\mathrm{X}}).\end{aligned}$$

The induction hypothesis implies that

$$\begin{aligned}&\sum_{\mathrm{W} \in T_{l+2}} a(\mathrm{W}) r(\overline{\mathrm{W}}) \\ &\ge p_G(\overline{v}) \left(1 - d^{-1/2} - d^{-3/4+2\varepsilon}\right)^{(l+2)/2} - \frac{l \cdot e^{-d^\varepsilon/5}}{2} - \sum_{\mathrm{X} \in B \cap \Gamma_l} a(\mathrm{X}) r(\overline{\mathrm{X}}).\end{aligned}$$

Hence it is enough to show

$$\sum_{\mathrm{X} \in B \cap \Gamma_L} a(\mathrm{X}) r(\overline{\mathrm{X}}) \le e^{-d^\varepsilon/5},$$

which follows easily from Cauchy–Schwarz, (30) and (37):

$$\begin{aligned}\sum_{\mathrm{X} \in B \cap \Gamma_l} a(\mathrm{X}) r(\overline{\mathrm{X}}) &\le \left(\sum_{\mathrm{X} \in B \cap \Gamma_l} a^2(\mathrm{X})\right)^{1/2} \left(\sum_{\mathrm{X} \in B \cap \Gamma_l} r^2(\overline{\mathrm{X}})\right)^{1/2} \\ &\le \left(t(d-t)^l e^{-d^\varepsilon/2} (d-t)^{-l}\right)^{1/2} (p_G(\overline{v}))^{1/2} \le e^{-d^\varepsilon/5}.\end{aligned}$$

∎

Claim 2.

$$\sum_{\mathrm{W} \in T_l} a^2(\mathrm{W}) \le 1 + d^{-1/4+2\varepsilon} \quad \text{for all even } l \le t.$$

Proof. We show by induction that

$$\sum_{\mathrm{W} \in T_l} a^2(\mathrm{W}) \le \left(1 + 2d^{-3/4+\varepsilon}\right)^{l/2} \left(1 + \frac{l}{2t} e^{-d^\varepsilon}\right). \tag{44}$$

Since $l \le t = o(d^{1/2+\varepsilon})$, this is sufficient. As usual, the base case is trivial. For general cases, notice that

$$\begin{aligned}\sum_{\mathrm{W}\in T_{l+2}} a^2(\mathrm{W}) &= \sum_{\mathrm{X}\in T_l}\sum_{\mathrm{U}\in S(\mathrm{X})}\sum_{\mathrm{W}\in S(\mathrm{U})} a^2(\mathrm{W})\\ &= \sum_{\mathrm{X}\in T_l} a^2(\mathrm{X})s^{-2}(\mathrm{X}) \sum_{\mathrm{U}\in S(\mathrm{X})}\sum_{\mathrm{W}\in S(\mathrm{U})} 1\\ &= \sum_{\mathrm{X}\in T_l} a^2(\mathrm{X})s^{-1}(\mathrm{X})\overline{s}(\mathrm{X}).\end{aligned}$$

For $\mathrm{X}\in T_l\setminus\Gamma_l$, we have

$$s^{-1}(\mathrm{X})\overline{s}(\mathrm{X}) \le (1+2d^{-3/4+\varepsilon})$$

while for $\mathrm{X}\in\Gamma_l$, we use the trivial bounds

$$s(\mathrm{X}) \ge d-t \quad \text{and} \quad \overline{s}(\mathrm{X}) \le d.$$

These yield

$$\begin{aligned}\sum_{\mathrm{W}\in T_{l+2}} a^2(\mathrm{W}) &\le \left(1+2d^{-3/4+\varepsilon}\right)\sum_{\mathrm{X}\in T_l\setminus\Gamma_l} a^2(\mathrm{X}) + \frac{d}{d-t}\sum_{\mathrm{X}\in\Gamma_l} a^2(\mathrm{X})\\ &\le \left(1+2d^{-3/4+\varepsilon}\right)\sum_{\mathrm{X}\in T_l} a^2(\mathrm{X}) + \sum_{\mathrm{X}\in\Gamma_l} a^2(\mathrm{X}).\end{aligned}$$

On the other hand, (28) and $a(\mathrm{X})\le(d-t)^{-l}$ give

$$\sum_{\mathrm{X}\in\Gamma_l} a^2(\mathrm{X}) \le t^{-1}(d-t)^l e^{-d^\varepsilon}(d-t)^{-l} = t^{-1}e^{-d^\varepsilon},$$

so the induction hypothesis implies (44). ∎

Proof of (41)**.** Claims 1 and 2 with (43) imply that, for all even $l \le t/2-2 = o(d^{1/2+\varepsilon})$,

$$\begin{aligned}\sum_{\mathrm{W}\in T_l} r^2(\overline{\mathrm{W}}) &\ge p_G^2(\overline{v})\left(1-d^{-1/2}-2d^{-3/4+2\varepsilon}\right)^l\left(1+d^{-1/4+2\varepsilon}\right)^{-1}\\ &\ge p_G^2(\overline{v})\left(1-d^{-1/4+3\varepsilon}\right)\left(1-d^{-1/2}\right)^l.\end{aligned}$$

Thus

$$\sum_{\substack{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}\\ |\mathrm{W}|\ \text{even}}} r^2(\overline{\mathrm{W}}) \geq \sum_{\substack{l=2\\ l\ \text{even}}}^{t/2-2} \sum_{\mathrm{W}\in T_l} r^2(\overline{\mathrm{W}})$$

$$\geq p_G^2(\overline{v})\left(1-d^{-1/4+3\varepsilon}\right)\sum_{\substack{l=2\\ l:\ \text{even}}}^{t/2-2}\left(1-d^{-1/2}\right)^l.$$

Since

$$\sum_{\substack{l=2\\ l:\ \text{even}}}^{t/2-2}\left(1-d^{-1/2}\right)^l \geq (1-d^{-1/4+\varepsilon})d^{1/2}/2$$

and $p_G(\overline{v})=d^{-1/2}+O(d^{-3/4+\varepsilon})$, we have, finally,

$$\sum_{\substack{\mathrm{W}\in V(T)\setminus\{\mathrm{V}\}\\ |\mathrm{W}|\ \text{even}}} r^2(\overline{\mathrm{W}}) \geq \frac{p_G^2(\overline{v})(1-2d^{-1/4+3\varepsilon})d^{1/2}}{2}$$

$$\geq \frac{p_G(\overline{v})(1-d^{-1/4+4\varepsilon})}{2}.$$

∎

References

[1] M. Ajtai, J. Komlós and E. Szemerédi: A dense infinite Sidon sequence, *Europ. J. Combinatorics* **2** (1981), 1–11.

[2] N. Alon, J. H. Kim and J. Spencer: Nearly perfect matchings in regular simple hypergraphs, *Israel J. Math* **100** (1997), 171–187.

[3] N. Alon and J. H. Spencer: *The Probabilistic Method*, Wiley, New York, 1992.

[4] B. Bollobás: Martingales, isoperimetric inequalities and random graphs, in *Combinatorics,* A. Hajnal, L. Lovász and V. T. Sós Eds., Colloq. Math. Soc. János Bolyai **52**, North Holland, 1988.

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

[6] Z. Füredi: Matchings and covers in hypergraphs, *Graphs and Combinatorics* **4** (1988), 115–206.

[7] C. D. Godsil: Matching behavior is asymptotically normal, *Combinatorica* **1** (1981), 369–376.

[8] C. D. Godsil: Matchings and walks in graphs, *J. Graph Th.* **5** (1981), 285–297.

[9] C. D. Godsil and I. Gutman: On the matching polynomial of a graph, pp. 241–249 in *Algebraic Methods in Graph Theory, I* (L. Lovász and V. T. Sós, eds.), Colloq. Math. Soc. János Bolyai **25**, North-Holland, Amsterdam, 1981.

[10] G. H. Hardy, J. E. Littlewood and G. Pólya: *Inequalities* (second edition), Cambridge University Press, Cambridge, 1952.

[11] L. H. Harper: Stirling behavior is asymptotically normal, *Ann. Math. Stat.* **38** (1967), 410–414.

[12] O. J. Heilmann and E. H. Lieb: Monomers and dimers, *Phys. Rev. Letters* **24** (1970), 1412–1414.

[13] O. J. Heilmann and E. H. Lieb: Theory of monomer-dimer systems, *Comm. Math. Physics* **25** (1972), 190–232.

[14] A. Johansson: An improved upper bound on the choice number for triangle free graphs, manuscript, 1994.

[15] J. Kahn: Asymptotically good list-colorings, *J. Combinatorial Th. (A)* **73** (1996), 1–59.

[16] J. Kahn: Recent results on some not-so-recent hypergraph matching and covering problems, pp. 305–353 in *Extremal Problems for Finite Sets, Visegrád, 1991*, Bolyai Soc. Math. Studies **3**, 1994.

[17] J. Kahn: A linear programming perspective on the Frankl–Rödl–Pippenger Theorem, *Random Structures and Algorithms,* **8** (1996), 149–157.

[18] J. Kahn: A normal law for matchings, submitted.

[19] J. Kahn and P. M. Kayll: Fractional v. integer covers in hypergraphs of bounded edge size, *J. Combinatorial Th. (A)* **78** (1997), 199–235.

[20] I. Kaplansky and J. Riordan: The problem of the rooks and its applications, *Duke Math. J.* **13** (1946) 259–268.

[21] J. H. Kim: On Brooks' Theorem For Sparse Graphs, *Combinatorics, Probability & Computing* **4** (1995), 97–132.

[22] J. H. Kim: The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures and Algorithms* **7** (1995), 173–207.

[23] H. Kunz: Location of the zeros of the partition function for some classical lattice systems, *Phys. Lett. (A)* (1970), 311–312.

[24] L. Lovász and M. D. Plummer: *Matching Theory*, North Holland, Amsterdam, 1986.

[25] C. J. H. McDiarmid: On the method of bounded differences, pp. 148–188 in *Surveys in Combinatorics 1989, Invited Papers at the 12th British Combinatorial Conference,* J. Siemons Ed., Cambridge Univ. Pr., 1989.

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

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

[28] A. Ruciński: The behaviour of $\binom{n}{k,\ldots,k,n-ik}c^i/i!$ is asymptotically normal, *Discrete Math.* **49** (1984), 287–290.

[29] J. Spencer: Lecture notes, M.I.T., 1987.

[30] J. Spencer: Asymptotic packing via a branching process, *Random Structures and Algorithms* **7** (1995), 167–172.

[31] V. G. Vizing: On an estimate of the chromatic class of a p-graph (in Russian), *Diskret. Analiz* **3** (1964), 25–30.

Jeff Kahn

Department of Mathematics and RUTCOR
Rutgers University
New Brunswick, NJ 08903, USA
jkahn@math.rutgers.edu

Jeong Han Kim

AT&T Bell Laboratories
Murray Hill, NJ 07974, USA

Current address:

Microsoft Research
One Microsoft Way
Redmond WA 98052, USA