A normal law for matchings

Jeff Kahn
Department of Mathematics and RUTCOR
Rutgers University
New Brunswick NJ 08903

Dedicated to the memory of Paul Erdös, both for his pioneering discovery of normality in unexpected places, and for his questions, some of which led (eventually) to the present work.

ABSTRACT

For a simple graph G, let ξ_G be the size of a matching drawn uniformly at random from the set of all matchings of G. Motivated by work of C. Godsil [11], we give, for a sequence $\{G_n\}$ and $\xi_n = \xi_{G_n}$, several necessary and sufficient conditions for asymptotic normality of the distribution of ξ_n, for instance

$$\{\Pr(\xi_n = k)\}_{k \geq 0} \text{ is asymptotically normal iff } \nu_n - \mu_n \to \infty$$

(where $\mu_n = \mathbf{E}\xi_n$ and ν_n is the size of a largest matching in G_n). In particular this gives asymptotic normality for any sequence of regular graphs (of positive degree) or graphs with perfect matchings.

The material presented here suggests numerous related questions, some of which are discussed in the last section of the paper.

1 Introduction

Given a graph $G = (V, E)$, we write $\mathcal{M}(G)$ for the set of matchings of G. In this paper we are concerned with the behavior of the random variable $\xi_G = |M|$, where M is drawn uniformly from $\mathcal{M}(G)$. We set $p_k(G) = \Pr(\xi_G = k)$,

---

Supported in part by NSF.
1991 Mathematics Subject Classification. Primary 05C70, 60F05; Secondary 05B35; 05C65.

and let $\mu(G)$ and $\sigma(G)$ denote the mean and standard deviation of ξ_G. (Our notation here mainly follows [34, p.341].)

In what follows we often deal with sequences $\{G_n\}$ of (*simple*) graphs. Given such a sequence, we abbreviate ξ_{G_n}, $\mu(G_n)$ and $\sigma(G_n)$ to ξ_n, μ_n and σ_n, and in addition set $\Delta(G_n) = \Delta_n$, $\delta(G_n) = \delta_n$ (where Δ, δ denote maximum and minimum degrees), $|V(G_n)| = v_n$, $|E(G_n)| = e_n$, and $\nu(G_n) = \nu_n$, $\tau(G_n) = \tau_n$, $\alpha(G_n) = \alpha_n$, where, as usual, ν, τ and α are matching, vertex cover and independence numbers. (For graph theory background see e.g. [3] or [34].) To avoid trivialities we always assume

$$v_n \to \infty \;\; (n \to \infty) \quad \text{and} \quad \delta_n \geq 1. \tag{1}$$

The sequence of distributions $\{p_k(G_n)\}_{k\geq 0}$ is said to be *asymptotically normal* if for each $x \in \mathbf{R}$

$$\Pr\left(\frac{\xi_n - \mu_n}{\sigma_n} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \quad (n \to \infty).$$

In 1981, following earlier work of Harper [16] (see below), C. Godsil [11] gave necessary and sufficient conditions for asymptotic normality of the distributions $\{p_k(G_n)\}$:

Theorem 1.1 *The distribution $\{p_k(G_n)\}_{k\geq 0}$ is asymptotically normal if and only if*

$$\sigma_n \to \infty \quad (n \to \infty). \tag{2}$$

Necessity of the condition is obvious; Godsil's proof of sufficiency is sketched below.

Given Theorem 1.1, one is left with the problem of understanding *when* the variance of ξ is large. This turns out to be surprisingly difficult and is the primary concern of the present paper. Our main results—Theorems 1.6 and 1.10—characterize asymptotic normality in terms of several graph statistics which, as indicated in Proposition 1.8, are more manageable than the variance. In fact the content of the paper is wholly combinatorial: the connection with normality being established by Theorem 1.1, the task here is to develop a sort of rough statistical minimax theory, saying that each of our statistics is large if and only if all are large.

Before stating these results, we slightly elaborate on [11] and earlier work. For a more thorough discussion see [34, Chapter 8].

Let $p(G,x)$ denote the probability generating function of the sequence $\{p_k(G)\}$,

$$p(G,x) = \sum_{k=0}^{\nu(G)} p_k(G)x^k.$$

A fundamental discovery of Heilmann and Lieb ([17], [18]; see also [29]) says that for every G,

$$p(G,x) \text{ has real roots.} \tag{3}$$

(Note that the roots are then necessarily negative. When G is bipartite, $p(G,x)$ is, up to scaling by $|\mathcal{M}(G)|$, the "rook polynomial" associated with G [28], [40]. For this case (3) was conjectured (some years after [17], [18], [29]) in [13], and a proof was provided in [37]. For more in related directions, see [48] and the references cited there.)

The proof of Theorem 1.1 is modeled on Harper's proof [16] of asymptotic normality of the sequence $\{S(n,k)/B_n\}_{k\geq 1}$, with $S(n,k)$ the Stirling number of the second kind and B_n the Bell number. Harper's basic observation is that, if (3) holds, then (2) is necessary and sufficient for asymptotic normality of $\{p_k(G_n)\}_{k\geq 0}$. Indeed, if the roots of $p(G,x)$ are $-\lambda_i$, then we must have $p(G,x) = \prod \frac{x+\lambda_i}{1+\lambda_i}$, so that

$$\xi = \sum_{i=1}^{\nu} \xi_i,$$

where the ξ_i are independent Bernoulli random variables with

$$\Pr(\xi_i = 0) = \frac{\lambda_i}{1+\lambda_i}, \qquad \Pr(\xi_i = 1) = \frac{1}{1+\lambda_i}. \tag{4}$$

(This observation, rediscovered by Harper, is due to P. Lévy [31], [32].) The sufficiency of (2) for asymptotic normality of $\{p_k(G_n)\}$ is then given by, for example, either Ljapunov's or Lindeberg's version of the Central Limit Theorem (e.g. [8]).

Actually, as Godsil shows, one has a bit more than asymptotic normality; namely, for fixed $x \in \mathbf{R}$, if $(k_n - \mu_n)/\sigma_n \to x$, then

$$\sigma_n p_k(G_n) \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (n \to \infty).$$

This is called *asymptotic local normality* in [34]. See also [38] for a survey of probabilistic consequences of probability generating functions with real roots.

Godsil used Theorem 1.1 to give one fairly general sufficient condition for asymptotic normality:

Theorem 1.2 *If $e_n/\Delta_n^2 \to \infty$ then $\sigma_n \to \infty$ (and so $\{p_k(G_n)\}_{k\geq 0}$ is asymptotically normal).*

Ruciński [41] then showed that the same conclusion holds under a weaker hypothesis:

Theorem 1.3 *If $\nu_n/\Delta_n \to \infty$ then $\sigma_n \to \infty$.*

(This implies Theorem 1.2 because $e_n \leq 2\nu_n\Delta_n$.)

As Godsil observed, Theorem 1.2 gives asymptotic normality for sequences of regular graphs, provided the degrees are not too large:

Corollary 1.4 *If G_n is regular of degree d_n with $v_n/e_n \to \infty$, then $\{p_k(G_n)\}_{k\geq 0}$ is asymptotically normal.*

The proofs of (2) under the hypotheses of Theorems 1.2 and 1.3 are based on the fact, shown in [18], that

$$\text{if } \ p(G, x_0) = 0, \text{ then } -x_0 \geq 1/(4(\Delta(G) - 1)).$$

Godsil also proved asymptotic normality for $\{p_k(K_n)\}_{k\geq 0}$ and went on to conjecture two further sufficient conditions, *viz.*

(a) $\quad \nu_n = \Omega(v_n)$, and

(b) $\quad e_n = \Omega(v_n^2)$.

As it turns out, both these conjectures are false:

Example 1.5 *Fix $\varepsilon > 0$ and take $G_n = K_{n,\lfloor(1+\varepsilon)n\rfloor}$, with the parts of the bipartition denoted X and Y (say with $|X| = n$) for future reference. Then* (a) *and* (b) *hold, but $\sigma_n = O(1)$.*

Nonetheless, it seems clear that something more general than Godsil's (or Ruciński's) results should be true, and his conjectures appear to have been made in this spirit. (He describes the first conjecture as "offered more in the hope of provoking an answer than from a confident belief in its truth.")

For example, neither of Theorems 1.2, 1.3 applies to sequences of regular graphs in which degree grows in proportion to the number of vertices (though Godsil's conjectures would have covered, and were perhaps motivated by, this case). That one does have asymptotic normality for such sequences is one consequence of our first result, the "normal law" of the title:

Theorem 1.6 *The distribution* $\{p_k(G_n)\}_{k\geq 0}$ *is asymptotically normal if and only if*

$$\nu_n - \mu_n \to \infty \quad (n \to \infty). \tag{5}$$

Equivalently,

$$\sigma_n \to \infty \quad \text{iff} \quad \nu_n - \mu_n \to \infty. \tag{6}$$

Notice that, with notation as in (4),

$$\sigma^2(G) \;=\; \sum_{i=1}^{\nu} \frac{\lambda_i}{(1+\lambda_i)^2} \;\leq\; \min\{\mu(G), \nu(G) - \mu(G)\}, \tag{7}$$

which in particular gives the necessity of (5), though this, like the necessity of (2), is clear on more general grounds. Of course $\mu_n \to \infty$ is also necessary, but we do not need to assume it here, since μ is in fact always of the same order of magnitude as ν. This was discovered by L. Babai (see [11]), who showed $\mu(G) \geq \nu(G)/3$ for all G. Here we give an exact result:

Proposition 1.7 *For every multigraph* G, $\mu(G) \geq \nu(G)/2$.

This is sharp, for example, when G is a matching or a path of length 3, and also in various other cases. While we don't really need Proposition 1.7 in what follows (we do use it in the proof of Theorem 1.3 given in Section 5), we include it as it seems natural and not quite obvious. The easy proof is given at the end of Section 2.

Let us again stress that Theorem 1.6 is by no means a matter of replacing one unverifiable condition by another. This is illustrated by the next Proposition, which is proved in Section 3. (Recall we assume (1).)

Proposition 1.8 *Each of the following implies* (5):

(a) $\nu_n > (1 - o(1))v_n/2$;

(b) $\nu_n/\Delta_n \to \infty$;

(c) $\delta_n = \omega(\Delta_n^2/v_n)$;

(d) $\delta_n = (1 - o(1))\Delta_n$.

In particular we recover (via (b)) Theorems 1.2 and 1.3. (Of course Theorem 1.3 also implies that (b) gives (5).) Note that since $\nu_n > e_n/(2\Delta_n) \geq v_n\delta_n/(4\Delta_n)$, condition (c) implies condition (b), so implies (5) provided (b) does. We include (c) because it together with (d) characterizes those sequences $\{(v_n, \delta_n, \Delta_n)\}$ which guarantee (5). (That is, we have (5) for all $\{G_n\}$ with $\{(v_n, \delta_n, \Delta_n)\}$ iff at least one of (c), (d) is satisfied. For instance, to see necessity when $(\delta_n+\Delta_n)|v_n$, take G_n to consist of $v_n/(\delta_n+\Delta_n)$ copies of K_{δ_n,Δ_n}.) Of course (d) shows in particular that we have asymptotic normality for sequences of regular graphs.

Note also that (a) gives asymptotic normality whenever the G_n have perfect matchings. This includes, for example, Godsil's asymptotic normality of $\{p_k(K_n)\}$ (as does (d)). It also includes Harper's asymptotic normality of $\{S(n,k)/B_n\}$ (see [11] or [40, p.213] for the connection). Interestingly, (a) is best possible: for positive constant ε, the assumption $\nu_n > (1-\varepsilon)v_n/2$ does not imply (5), as shown by Example 1.5 above.

As indicated above, Theorem 1.6 provides the first proof of (2) for sequences of regular graphs. In fact, as shown even more recently in [27], the values of μ and σ^2 for a regular graph are remarkably well determined just by degree and number of vertices. For $x \in V(G)$, we write $p(\overline{x})$ for the probability that x is not covered by the random matching M.

Theorem 1.9 ([27]) *For any d-regular simple graph G,*

(a) $p(\overline{x}) \sim d^{-1/2}\ \ \forall x \in V(G)$, *so that* $v(G) - 2\mu(G) \sim v(G)/\sqrt{d}$,

(b) $\sigma^2(G) \sim v(G)/(4\sqrt{d})$

(with limits taken as $d \to \infty$, so e.g. the convergence in (a) of $p(\overline{x})d^{1/2}$ to 1 is uniform in G, x).

Theorems 1.6 and 1.9 and Proposition 1.8 were announced in [21], [22].

The next result expands Theorem 1.6 to include a few more equivalent conditions. We need two additional parameters, which may be of independent interest. The first, which we may call the *cover defect* (name suggested by L. Babai), is

$$\begin{aligned}\kappa(G) &= \min\{\sum_{y\in Y} p(\overline{y}) : Y \text{ a vertex cover of } G\} \\ &= \min\{\mathbf{E}[|Y \setminus \cup_{A\in M} A|] : Y \text{ a vertex cover of } G\}.\end{aligned}$$

Write F_M for the set of edges not meeting any edges of M. Our second new parameter—call it the *residual matching number*—is

$$\lambda(G) = \mathbf{E}[\nu(F_M)].$$

As in other cases, we set $\kappa(G_n) = \kappa_n$, $\lambda(G_n) = \lambda_n$. We state our conditions now in the negative, the form in which we shall usually treat them in what follows. (As usual, τ denotes vertex cover number.)

Theorem 1.10 *The following are equivalent:*

(a) $\sigma_n = O(1)$;

(b) $\nu_n - \mu_n = O(1)$;

(c) $\tau_n - \mu_n = O(1)$;

(d) $\kappa_n = O(1)$;

(e) $\lambda_n = O(1)$.

The equivalence of (a) and (b) is just Theorem 1.6. As noted above, (b) $\Rightarrow$ (a) follows from (7). That (c) implies (b) is trivial. Note this also shows that $\tau_n - \nu_n \to \infty$ is sufficient for asymptotic normality; it is not necessary, however (take $G_n = K_{n,n}$). Other easy implications are (b) $\Rightarrow$ (e) and (d) $\Rightarrow$ (e). Condition (d) is central in that the proof that (a) $\Rightarrow$ (b) (that is, Theorem 1.6) goes via (a) $\Rightarrow$ (d) $\Rightarrow$ ((c) $\Rightarrow$) (b). The hardest part of this is (a) $\Rightarrow$ (d), which is shown in Section 8. The other, easier implications are given in Section 7.

Of course, Theorem 1.10 simply says that each of the quantities $\sigma(G)$, $\nu(G) - \mu(G)$ etc. appearing in (a)-(e) can be bounded in terms of the others. Though the bounds in the implications we prove directly are all quite tight, most of the implied bounds should be improvable. For example, we show, provided $\sigma(G) = \Omega(1)$, that $\nu(G) - \mu(G) = O(\kappa^2(G))$ and $\kappa(G) = O(\sigma^4(G))$, whence $\nu(G) - \mu(G) = O(\sigma^8(G))$. But while the first two bounds are sharp (except for the constants), the third probably is not: we conjecture that

$$\text{if } \sigma(G) = \Omega(1), \text{ then } \nu(G) - \mu(G) = O(\sigma^6(G)). \quad (8)$$

This would be best possible; see Example 7.2. (The rather strange fact that $\nu - \mu$ can be so much larger than σ^2 suggests that it may not be easy to find a very simple proof of Theorem 1.6. When $\sigma(G)$ is small then we do get the right bound here, namely $\nu(G) - \mu(G) = O(\sigma^2(G))$. See Section 7 for more on what we know and would like to know in this vein.)

We record one more sufficient condition, also proved in Section 7.

Proposition 1.11 *If*

$$\tau_n > (1 - o(1))v_n/2, \tag{9}$$

then $\{p_k(G_n)\}$ *is asymptotically normal.*

Note (9) is weaker than (a) of Proposition 1.8, and is equivalent to

$$\alpha_n < (1 + o(1))v_n/2$$

(with α as usual denoting independence number). It was initially thought that an argument similar to those needed for Proposition 1.8 would suffice for Proposition 1.11, but we do not see this at the moment. (Of course, (9) is also not so easy to check as the conditions of Proposition 1.8.)

In connection with an earlier version of this paper, Anant Godbole [10] asked me whether one could prove a Poisson limit for ξ_n in case σ_n or $\nu_n - \mu_n$ tends to a constant rather than infinity. A partial answer is given by the next theorem, in which we write $p(A)$ for the probability that $A \in E(G_n)$ belongs to M drawn uniformly from $\mathcal{M}(G_n)$.

Theorem 1.12 *Let c be a positive constant, and suppose $\nu_n - \mu_n \to c$. Then the distribution of $\nu_n - \xi_n$ converges to* $\mathrm{Po}(c)$ *(that is,* $\Pr(\nu_n - \xi_n = k) \to e^{-c}c^k/k!$ *for each k) provided* $\max\{p(A) : A \in E(G_n)\} \to 0$.

In fact this may be a full answer: it seems likely that the condition $\max\{p(A) : A \in E(G_n)\} \to 0$ is also necessary for the Poisson limit, though showing this seems less straightforward than one might expect.

Before closing this section, we just sketch a few points from the proof of Theorem 1.10.

For $M \in \mathcal{M}$ we again write F_M for the set of edges of G not met by any edges of M, and define the random variable ψ on $\mathcal{M}$ by

$$\psi(M) = |F_M|.$$

Notice—see (11) below—that $\mathbf{E}[\psi] = \mu$. More generally, for any $F \subseteq E$, we define

$$\psi_F(M) = |F_M \cap F|,$$

the number of edges of F not met by any edges of M.

The central idea of the proof is that there is a tradeoff between concentration of ξ and concentration of ψ: if the variance of ψ is small relative to μ^2, then $\sigma^2 = \sigma_\xi^2$ must be large (see Lemma 6.1). The original proof of asymptotic normality for regular graphs (which was the first step in the direction of the present work) is a direct application of this idea, and is given following Lemma 6.1.

In general, however, ξ and ψ can both fail to be concentrated. For instance, we may take G_n to be the union of two graphs, one with large σ and the other with large σ_ψ. A more interesting example is obtained from Example 1.5 by adding $\lfloor \varepsilon n \rfloor$ new vertices to X, and joining them by a matching N to some $\lfloor \varepsilon n \rfloor$ vertices of Y.

Such examples suggest that σ may be mostly attributable to some $F \subseteq E$, and that we should try to identify such an F and derive a lower bound on σ from concentration of ψ_F. (In the second example, most of σ comes from $|M \cap N|$ (the variance of $|M \setminus N|$ is $O(1)$), and ψ_N *is* concentrated. (The distribution of $|M \cap N|$ is close to the binomial distribution $B(\lfloor \gamma n \rfloor, 1/2)$, with $\gamma = \varepsilon^2/(1+\varepsilon)$.))

As it turns out, we *can* show (Lemma 6.5) that $\sigma \to \infty$ if some ψ_F has small variance, but have not so far been able to complete a proof based entirely on this fact. What we can do instead is to extract enough structural information from the lack of concentration of appropriate ψ_F's to finish in a simpler way, based on the observation (Lemma 5.1) that if σ is small, then no matching N can have large $\mathbf{E}[|M \cap N|]$. (Section 5 also includes a surprisingly simple proof of Theorem 1.3 based on this observation.)

Crucial to many of our arguments is the fact that there is considerable approximate independence in the probability spaces in question. For example, conditioning on some particular $x \in V$ being covered by M has little effect on the expected size of M, or, for most other vertices y, on the probability that y is also covered. Results in this vein, based on a simple but surprisingly powerful observation from [25], are given in Section 4.

The rest of the paper is organized as follows. Section 2 contains notation, a few easy identities, and the proof of Proposition 1.7. Section 3 contains the proof of Proposition 1.8.

The main part of the paper consists of Sections 4-8. The first three of these are devoted to the various Lemmas mentioned above. In Section 7, we prove Proposition 1.11 and the easier parts of Theorem 1.10, and in Section 8 we complete the proofs of Theorems 1.6 and 1.10 by showing that (a) of Theorem 1.10 implies (d).

Section 9 supplies the proof of Theorem 1.12, which is mainly an application of some of the earlier arguments.

Finally, Section 10 discusses some of the many questions suggested by the present work.

2 Preliminaries

Throughout the paper we take "graph" to mean simple graph. In what follows, we deal with a fixed graph $G = (V, E)$, and write $\mathcal{M} = \mathcal{M}(G)$, $\xi = \xi_G$ and so on. For $M \in \mathcal{M}$ and $x \in V$, we write $x \prec M$ if x is contained in some edge of M. As already mentioned, we write F_M for the set of edges not met by any edges of M, i.e.

$$F_M = \{A \in E \setminus M : M \cup \{A\} \in \mathcal{M}\},$$

and set $\psi = \psi(M) = |F_M|$, and $\psi_F(M) = |F_M \cap F|$ if $F \subseteq E$.

A natural generalization of uniform distribution on $\mathcal{M}$ is the distribution $p = p_\alpha$ derived from some $\alpha : E \to \mathbf{R}^+$ according to

$$w(M) = \prod_{A \in M} \alpha_A,$$

$$p(M) = w(M) / \sum_{M' \in \mathcal{M}} w(M').$$

We will call p_α the *hard-core distribution* associated with α. The name is from statistical physics (e.g. [2]), where the weights α_A are sometimes called *activities*. ("Monomer-dimer system" and "exclusion model" are also used; see e.g. [17], [18], [29].) Other recent, rather diverse contexts in which hard-core distributions have proved important include [33], [39], [30] (see also [7] for an exposition of [30]). They are called "normal populations" in [39], and in [30], are not thought of as probability distributions at all, but as "canonical" convex representations of points in $\mathbf{R}^n$. See also [26], [23], [24] for applications of hard-core distributions to combinatorial problems. (The last two references are more recent than the present work, but were written a little more quickly.)

Of course $\alpha \equiv 1$ gives uniform distribution, and, more generally, if $\alpha : E \to \mathbf{N}$, then p_α corresponds to uniform distribution on $\mathcal{M}(G(\alpha))$, where the multigraph $G(\alpha)$ is obtained from G by taking α_A copies of A for each $A \in E$.

Though we deal with hard-core distributions only in passing, it would be very interesting to understand how our results might generalize to this setting; see Section 10. (We introduce the distributions here rather than in Section 10 because they appear briefly in Section 4.)

For p a probability distribution on $\mathcal{M}$, $M \in \mathcal{M}$ chosen according to p, $x_i \in V$ and $A = \{x, y\} \in E$, set

$$p(x_1, \ldots, x_t) = \Pr(x_1, \ldots, x_t \prec M),$$

$$p(\overline{x_1}, \ldots, \overline{x_t}) = \Pr(x_1, \ldots, x_t \not\prec M),$$

and

$$p(A) = p(xy) = \Pr(A \in M).$$

(Note that $p(xy)$ is not the same as $p(x, y)$; but we will make no use of the latter, so hopefully this will not cause any confusion.) We also extend this notation to conditional probabilities—e.g. $p(\overline{x}|\overline{y})$—in the obvious way.

From now on, except where stated otherwise, we confine ourselves to uniform distribution on $\mathcal{M}(G)$ with G simple.

For $A = \{x, y\} \in E$ we have the trivial but important identity

$$p(A) = p(\overline{x}, \overline{y}), \tag{10}$$

which implies in particular

$$E[\psi] \quad = \quad \mu. \tag{11}$$

Another quantity which plays an important role in what follows is

$$\varphi(x) := \sum_{y \sim x} p(\overline{y}). \tag{12}$$

As will appear below (see Section 4), $\varphi(x)$ is usually about the same as $p(\overline{x})^{-1}$. This is made at least plausible by the basic identity

$$p(\overline{x}) = [1 + \sum_{y \sim x} p(\overline{y}|\overline{x})]^{-1} \tag{13}$$

Proof. By (10),

$$1 - p(\overline{x}) \ = \ p(x) \ = \ \sum_{x \in A \in E} p(A) \ = \ \sum_{y \sim x} p(\overline{x}, \overline{y}).$$

Dividing by $p(\overline{x})$ and rearranging gives (13).

□

Let us mention another identity in the same vein. We do not need this for our main argument, but will use it later (see Section 6) to give a relatively simple proof of asymptotic normality for regular graphs. (We use $d(x)$ for the degree of x.)

$$p(\overline{x}) \ = \ \frac{1}{d(x)+1}[1 + \sum \{p(\overline{x}, \overline{y}, \overline{z}) : x \sim y \sim z \neq x\}] \tag{14}$$

Proof. Multiplying by $(d(x)+1)/p(\overline{x})$ and using (13), we find that (14) is equivalent to

$$1 + \sum_{y \sim x} [p(\overline{y}|\overline{x}) + \sum \{p(\overline{y}, \overline{z}|\overline{x}) : x \neq z \sim y\}] \ = \ d(x) + 1.$$

But the term in square brackets is easily seen (use (10) in $G - x$) to be equal to 1 for each y.

□

We close this Section with the promised

Proof of Proposition 1.7. (Recall p here is uniform distribution on $\mathcal{M}(G)$ with $G = (V, E)$ a *multigraph*.) We use induction on $|E|$. If G is a matching then $\mu = \nu/2$. If not, choose $A \in E$ with ends x, y such that some maximum matching of G does not contain A. Set $p(A) = p$. Let $G' = G - x - y$, $G'' = G - A$, and let M', M'' be random matchings of G', G'' respectively. Then

$$\nu(G') \geq \nu - 2$$

and (by choice of A)

$$\nu(G'') = \nu.$$

So by induction we have

$$\begin{aligned}\mu &= p\mathbf{E}[|M||A \in M] + (1-p)\mathbf{E}[|M||A \notin M] \\ &= p(\mathbf{E}[|M'|] + 1) + (1-p)\mathbf{E}[|M''|] \\ &\geq p(\nu(G')/2 + 1) + (1-p)\nu(G'')/2 \\ &\geq \nu/2.\end{aligned}$$

□

3 Sufficient conditions

In this section we prove Proposition 1.8. The proof is based on some consideration of the random variable $\psi(M)$. As mentioned earlier, ψ and the related random variables ψ_F will again play a central role in the proofs of our main results.

We have already noted that (c) ⇒ (b). For the implication (d) ⇒ (5), we mainly appeal to work of Bollobás and Eldridge [4] which says in particular that if $\Delta_n = \omega(1)$ and (d) holds then (a) holds as well. (Bollobás and Eldridge actually give the minimum possible ν for any specified values of δ, Δ and v, but of course nothing so precise is needed here.) If $\Delta_n = O(1)$ then (d) does not imply (a), but instead (trivially, by (1)) implies (b). (Of course it is enough to prove (d) ⇒ (5) under each of the assumptions $\Delta_n = \omega(1)$, $\Delta_n = O(1)$.)

Thus we really only need to derive (5) from (a) and (b). We now dispense with the sequence notation of Section 1 and work with a fixed G (taking v, ν etc. to be parameters of G), our task then being to show that $\nu - \mu$ is bounded below by appropriate functions of the quantities in (a) and (b) (see (19), (21)).

We use the fact that if the (arbitrary) probability distribution $\{p_i\}_{i=0}^{\nu}$ satisfies, for some k,

$$p_k \geq p_{k+1} \geq \cdots \geq p_\nu,$$

then for its mean μ we have $\mu \leq (\nu + k)/2$, and so

$$\nu - \mu \geq (\nu - k)/2. \tag{15}$$

Let $\mathcal{M}_i = \mathcal{M}_i(G)$ denote the set of i-matchings (matchings of size i) of G, and set $m_i = |\mathcal{M}_i|$ and $m = |\mathcal{M}| = \sum m_i$. Thus $p_i := p_i(G) = m_i/m$.

Let k be the largest index for which

$$p_k < p_{k+1}. \tag{16}$$

For $i = 1, \ldots, \nu$ form the natural bipartite graph B_i on $\mathcal{M}_{i-1} \cup \mathcal{M}_i$; namely, for $M \in \mathcal{M}_{i-1}$, $M' \in \mathcal{M}_i$,

$$M \sim_{B_i} M' \Longleftrightarrow M \subseteq M'.$$

Thus for degrees in B_i we have

$$d_{_{B_i}}(M) = \begin{cases} \psi(M) & \text{if } M \in \mathcal{M}_{i-1} \\ |M| \;\; (= i) & \text{if } M \in \mathcal{M}_i. \end{cases}$$

Let $\overline{\psi}_i$ denote the average of ψ on $\mathcal{M}_i$:

$$\overline{\psi}_i = \frac{1}{m_i} \sum_{M \in \mathcal{M}_i} \psi(M).$$

Then

$$m_k \overline{\psi}_k = |E(B_{k+1})| = m_{k+1}(k+1),$$

so (16) is the same as

$$\overline{\psi}_k > k+1. \tag{17}$$

Note that for any $M \in \mathcal{M}_k$ the $\psi(M)$ edges of F_M have matching number at most $\nu - k$, so can be covered by at most $2(\nu - k)$ vertices; that is,

$$\tau(F_M) \leq 2(\nu - k). \tag{18}$$

For (a) $\Rightarrow$ (5) we show

$$\nu - \mu = \Omega\left(\frac{\nu}{\max\{v - 2\nu, \sqrt{v}\}}\right). \tag{19}$$

(Note this is sharp for any $G = K_{\nu, v-\nu}$.) To see this, note that by (18) we have

$$\psi(M) < 2(\nu - k)(v - 2k)$$

for any $M \in \mathcal{M}_k$, so also, using (17),

$$k \;<\; \overline{\psi}_k \;<\; 2(\nu - k)(v - 2k). \tag{20}$$

Using (15) and the equivalent

$$k \;\geq\; 2\mu - \nu \;=\; \frac{1}{2}[v - (v - 2\nu) - 4(\nu - \mu)],$$

and setting $x = 4(\nu - \mu)$, $y = v - 2\nu$, we have from (20), $v - (y + x) \leq 2x(x + y)$, which implies

$$\begin{aligned} x \;&\geq\; \frac{1}{4}[-(2y+1) + \sqrt{(2y+1)^2 - 8(y - v)}\;] \\ &=\; \frac{4\nu}{2y + 1 + \sqrt{(2y-1)^2 + 8v}}\;. \end{aligned}$$

This gives (19).

□

For the derivation of (5) from Ruciński's condition (b), we again use (18) and (15), which imply

$$\nu - 2(\nu - \mu) \;\leq\; k \;<\; \overline{\psi}_k \;\leq\; 2(\nu - k)\Delta \;\leq\; 4(\nu - \mu)\Delta$$

and

$$\nu - \mu \geq \nu/(4\Delta + 2). \tag{21}$$

Note this is again about right; for example, $\nu - \mu = \nu/(\Delta + 1)$ if G is a disjoint union of Δ-stars.

4 Approximate independence

A guiding idea in the present work is that, while we usually cannot expect exact independence of events of interest, we do have considerable *approximate* independence. The following simple lemma, essentially from [25], is the basis for the results we need in this direction. Though we will mostly be concerned with uniform distribution in what follows, the lemma is true for all hard-core distributions; thus p should be understood as any p_α.

For $W \subseteq V$ we set

$$\mu(W) \ = \ \mu_G(W) \ = \sum_{w \in W} p(w),$$

and extend this notation to conditional situations—e.g. $\mu(W|w)$, $\mu(W|\overline{w})$—in the obvious way. (So for example $\mu(W|\overline{w}) = \mu_{G-w}(W \setminus \{w\})$.)

Lemma 4.1

(a) *If $w \in W$, then*

$$0 \leq \mu(W) - \mu(W|\overline{w}) \leq 2p(w).$$

(b) *If $w \in V \setminus W$ then*

$$-p(w) \leq \mu(W) - \mu(W|\overline{w}) \leq p(w).$$

Remarks

1. In particular this says that deleting a vertex cannot increase the expected size of our random matching. The same is *not* true for edge deletions (exercise). Also, somewhat curiously in view of the limited effect of vertex deletions on any $\mu(W)$, there is no constant bound on $|\mathbf{E}[|M \cap F|] - \mathbf{E}[|M \cap F||\overline{v}]|$ for general $F \subseteq E$. For example, take G to be a tree of depth D rooted at v, with vertices at depth $i < D$ having degree, say, $(D-i)^2$, and let F consist of edges at odd depth (where those containing v are at depth 0). Then it's not hard to see that $\mathbf{E}[|M \cap F||\overline{v}] - \mathbf{E}[|M \cap F|] = \Omega(D)$. (For multigraphs, a path of length D with initial vertex v and ith edge having multiplicity $(D-i+1)^2$ achieves the same effect.)

2. The cases of equality in Lemma 4.1 are easily extracted from the proof: equality holds in the lower bounds iff w is an isolated vertex, and in the upper bounds iff the component of G containing w is a star (possibly singleton) with center w and all other vertices contained in W. Relaxing the bounds in the Lemma by replacing $p(w)$ by (say) 1 would not make any serious difference in the applications below.

Proof of Lemma 4.1. We proceed by induction on $|V(G)|$; base cases are easily verified. Notice that (a) with $W = X$ is the same as (b) with $W = X \setminus \{w\}$, so it's enough to prove (a).

Write $N(w)$ for the neighborhood of (set of vertices adjacent to) w, and set

$$N(w) \cap W = \{w_1, \ldots, w_a\}, \qquad N(w) \setminus W = \{v_1, \ldots, v_b\}.$$

Then

$$\begin{aligned}
\mu(W) - \mu(W|\overline{w}) &= p(\overline{w})\mu(W|\overline{w}) + \sum_i p(ww_i)[2 + \mu(W|\overline{w}, \overline{w}_i)] \\
&\quad + \sum_j p(wv_j)[1 + \mu(W|\overline{w}, \overline{v}_j)] - \mu(W|\overline{w}) \\
&= \sum_i p(ww_i)[2 + \mu(W|\overline{w}, \overline{w}_i) - \mu(W|\overline{w})] \\
&\quad + \sum_j p(wv_j)[1 + \mu(W|\overline{w}, \overline{v}_j) - \mu(W|\overline{w})] \quad (22)
\end{aligned}$$

Each of the terms in square brackets in (22) is between 0 and 2, because, by induction,

$$-2 \leq \mu(W|\overline{w}, \overline{w}_i) - \mu(W|\overline{w}) \leq 0,$$
$$-1 \leq \mu(W|\overline{w}, \overline{v}_j) - \mu(W|\overline{w}) \leq 1.$$

The bounds of (a) follow, since $\sum_{x \sim w} p(wx) = p(w)$.

□

We record a number of consequences, now confining ourselves to uniform distribution. (Proofs will follow.) Neither Lemma 4.1 nor the corollaries below are the last word on this subject, but they are more than sufficient for our purposes.

Corollary 4.2 *Fix $W \subseteq V$ and define the random variable $\zeta = \zeta_W$ by*

$$\zeta(M) = |\{w \in W : w \not\sim M\}|.$$

Then (denoting the mean and variance of ζ by μ_ζ, σ_ζ^2)

$$\sigma_\zeta^2 \leq 2 \sum_{x \in W} p(\overline{x})p(x) \leq 2\mu_\zeta.$$

Corollary 4.3 *For all distinct $x, y \in V$,*

$$\frac{1}{1 + p(\overline{x})} \;\leq\; \frac{p(\overline{x}, \overline{y})}{p(\overline{x})p(\overline{y})} \;\leq\; \min\left\{\frac{1}{p(x)}, 2\right\}.$$

Corollary 4.4 *For all distinct $x, y \in V$,*

$$|p(\overline{x}) - p(\overline{x}|\overline{y})| \le 2p(\overline{x})^2.$$

Corollary 4.5 *For any $x \in V$*

$$p(x)\max\{p(x), 1/2\} \le p(\overline{x})\varphi(x) \le 1.$$

(See (12) for φ.)

For $X, Y \subseteq V$, set $E(X) = \{\{x,y\} \in E : x, y \in X\}$ and $E(X,Y) = \{\{x,y\} \in E : x \in X, y \in Y\}$.

Corollary 4.6 *Let $W \subseteq V$, $x \in V$ and suppose $F \subseteq E(W) \cup E(W, V \setminus W)$ contains no edges containing x. Then*

$$|\sum_{A \in F}(p(A|\overline{x}) - p(A))| \le 4\sum_{w \in W} p(\overline{w}). \tag{23}$$

Corollary 4.7 *Suppose $W \subseteq V$ satisfies*

$$p(\overline{x}) \ge \alpha \qquad \forall x \in W$$

and

$$\mathbf{E}[|M \cap N|] \le C \qquad \textit{for every matching } N \subseteq E(W).$$

Then there exists a cover X of $E(W)$ satisfying

$$\sum_{x \in X} p(\overline{x}) \le 4C\alpha^{-1}.$$

We proceed to the proofs.

Proof of Corollary 4.2. (Notice that only the first inequality requires proof.) For $x \in W$ let ζ_x be the indicator of the event $\{x \not\sim M\}$. Then $\zeta = \sum_{x \in W} \zeta_x$, and (using Lemma 4.1 for the inequality)

$$\begin{aligned}
\sigma_\zeta^2 &= \sum_{x \in W}\sum_{y \in W}(p(\overline{x},\overline{y}) - p(\overline{x})p(\overline{y})) \\
&= \sum_{x \in W} p(\overline{x}) \sum_{y \in W}(p(\overline{y}|\overline{x}) - p(\overline{y})) \\
&\le 2\sum_{x \in W} p(\overline{x})p(x).
\end{aligned}$$

□

(When $W = V$ the bound $\sigma_\zeta^2 \leq 2\mu_\zeta$ follows from (7), since in this case $\mu_\zeta = v - 2\mu \geq 2(\nu - \mu)$, implying

$$\sigma_\zeta^2 \quad = \quad 4\sigma_\xi^2 \quad \leq \quad 4(\nu - \mu) \quad \leq \quad 2\mu_\zeta.)$$

For the proofs of Corollaries 4.3-4.5, let

$$A = \sum_{z \sim x} p(\overline{z}|\overline{x}), \qquad B = \sum \{p(\overline{z}|\overline{x}, \overline{y}) : y \neq z \sim x\},$$

so that

$$p(\overline{x}) = (1 + A)^{-1}, \qquad p(\overline{x}|\overline{y}) = (1 + B)^{-1}$$

and

$$\frac{p(\overline{x}, \overline{y})}{p(\overline{x})p(\overline{y})} = \frac{1 + A}{1 + B}. \tag{24}$$

We assert that

$$-p(y|\overline{x}) \leq A - B \leq 1. \tag{25}$$

If $y \not\sim x$ this is immediate from Lemma 4.1 (applied to $G - x$; here we could replace the upper bound in (25) by $p(y|\overline{x})$).

If $y \sim x$ then

$$A - B = \sum \{p(\overline{z}|\overline{x}) - p(\overline{z}|\overline{x}, \overline{y}) : y \neq z \sim x\} + p(\overline{y}|\overline{x}).$$

The sum again has absolute value at most $p(y|\overline{x})$, so we again have (25) (and could increase the lower bound to $-p(y|\overline{x}) + p(\overline{y}|\overline{x})$).

Proof of Corollary 4.3. In view of (24), the first inequality of (25) gives

$$\frac{p(\overline{x}, \overline{y})}{p(\overline{x})p(\overline{y})} \;\geq\; \frac{1 + A}{1 + A + p(y|\overline{x})} \;=\; \frac{1}{1 + p(\overline{x}, y)} \;\geq\; \frac{1}{1 + p(\overline{x})},$$

while the second gives (noting $B \geq 0$)

$$\frac{p(\overline{x}, \overline{y})}{p(\overline{x})p(\overline{y})} \;\leq\; \min\left\{\frac{1 + A}{A}, \frac{2 + B}{1 + B}\right\} \;\leq\; \min\left\{\frac{1}{p(x)}, 2\right\}.$$

□

Both the upper bounds of Corollary 4.3 are sharp when $\{x,y\}$ is an isolated edge of G. The lower bound can be improved, as indicated above, to $[1+p(\overline{x},y)]^{-1}$ when $x \not\sim y$ (this is sharp if G is a path of length 2 with ends x, y), and to $[1+p(\overline{x})-2p(\overline{x},\overline{y})]^{-1}$ when $x \sim y$ (this is sharp whenever x is the unique neighbor of y).

Proof of Corollary 4.4.

$$\begin{aligned} |p(\overline{x})-p(\overline{x}|\overline{y})| &= |(1+A)^{-1}-(1+B)^{-1}| \\ &= |B-A|p(\overline{x})p(\overline{x}|\overline{y}) \\ &\leq p(\overline{x})p(\overline{x}|\overline{y}) \\ &\leq 2p(\overline{x})^2 \end{aligned}$$

(the first inequality by (25), and the second by Corollary 4.3).

□

Proof of Corollary 4.5. The upper bound follows from $\varphi(x) \leq A+1$, a consequence of Lemma 4.1. For the lower bound we use Corollary 4.3 to obtain

$$\varphi(x) \;=\; \sum_{z\sim x} p(\overline{z}) \;\geq\; \max\{p(x),1/2\}A,$$

and

$$p(\overline{x})\varphi(x) \;\geq\; p(\overline{x})\max\{p(x),1/2\}[p(\overline{x})^{-1}-1] \;=\; p(x)\max\{p(x),1/2\}.$$

□

Proof of Corollary 4.6. We may assume $x \notin W$. For each $A \in F$, choose an end $w(A) \in W$ of A. In what follows, w ranges over W, and, given w, z ranges over vertices joined to w by edges A of F with $w(A)=w$. The left hand side of (23) is then

$$\begin{aligned} |\sum_w\sum_z(p(\overline{w},\overline{z}|\overline{x})-p(\overline{w},\overline{z}))| &= |\sum_w\sum_z\{p(\overline{w}|\overline{x})(p(\overline{z}|\overline{x},\overline{w})-p(\overline{z}|\overline{w})) \\ &\qquad + p(\overline{z}|\overline{w})(p(\overline{w}|\overline{x})-p(\overline{w}))\}| \\ &\leq \sum_w\{p(\overline{w}|\overline{x})|\sum_z(p(\overline{z}|\overline{x},\overline{w})-p(\overline{z}|\overline{w}))| \\ &\qquad + 2p(\overline{w})^2\sum_z p(\overline{z}|\overline{w})\} \\ &\leq 2\sum_w p(\overline{w})(1+p(w)) \;\leq\; 4\sum p(\overline{w}) \end{aligned}$$

(using Corollary 4.4 for the first inequality and Lemma 4.1, Corollary 4.3 and (13) for the second).

□

The bound here, while sufficient for our purposes, seems far from the truth, and ought to be improvable (though, as noted in Remark 1 following the statement of Lemma 4.1, it cannot be improved to $O(1)$).

Proof of Corollary 4.7. Let $N = \{\{x_i, y_i\} : i \in I\}$ be a maximal matching of $E(W)$. Then by Corollary 4.3 and the assumptions on W,

$$\begin{aligned} C &\geq \mathbf{E}[|M \cap N|] \\ &= \sum_{i \in I} p(\overline{x}_i, \overline{y}_i) \\ &\geq (1/2) \sum_{i \in I} p(\overline{x}_i) p(\overline{y}_i) \\ &\geq (\alpha/4) \sum_{i \in I} (p(\overline{x}_i) + p(\overline{y}_i)). \end{aligned}$$

Since $\cup_{i \in I} \{x_i, y_i\}$ covers $E(W)$, this gives Corollary 4.7.

□

5 Few edges per matching

Lemma 5.1 *For any matching $N \subseteq E$,*

$$\mathbf{E}[|M \cap N|] \leq 2\sigma^2. \tag{26}$$

Proof. Let $Y = M \setminus N$, and define $\beta(Y)$ to be the number of edges of N not meeting any edge of Y. Then $Var[\xi | Y] = \beta(Y)/4$, so

$$\sigma^2 = Var[\xi] \geq \mathbf{E}\{Var[\xi | Y]\} = \frac{1}{4} \mathbf{E}[\beta(Y)]. \tag{27}$$

On the other hand, $\mathbf{E}[|M \cap N| | Y] = \beta(Y)/2$, so

$$\mathbf{E}[|M \cap N|] = \mathbf{E}\{\mathbf{E}[|M \cap N| | Y]\} = \frac{1}{2} \mathbf{E}[\beta(Y)],$$

and this with (27) gives (26).

□

Corollary 5.2 $\sigma^2 \geq \frac{1}{2}\sum_{A\in E} p(A)^2$

Proof. Since

$$\mathbf{E}[\sum_{A\in M} p(A)] = \sum_{A\in E} p(A)^2, \tag{28}$$

there exists a matching N (namely one for which $\sum_{A\in N} p(A)$ is at least the left hand side of (28)) with

$$\mathbf{E}[|M\cap N|] \ = \ \sum_{A\in N} p(A) \ \geq \ \sum_{A\in E} p(A)^2.$$

The Corollary now follows from Lemma 5.1.

□

We can use Lemma 5.1 to give a very simple proof of Theorem 1.3 (so also Theorem 1.2), as follows.

Given G (simple) of maximum degree Δ, let $N_1\cup\cdots\cup N_{\Delta+1}$ be a partition of $E(G)$ into matchings (this exists by Vizing's Theorem [47]). Then by Lemma 5.1 (for the first inequality) and Proposition 1.7 (for the third), we have

$$\sigma^2 \ \geq \ \frac{1}{2}\max_i \mathbf{E}[|M\cap N_i|] \ \geq \ \frac{\mu}{2(\Delta+1)} \ \geq \ \frac{\nu}{4(\Delta+1)}.$$

□

(Note this gives $\sigma^2 \geq |E|/(4(\Delta+1)^2)$, since $\nu \geq |E|/(\Delta+1)$, e.g. by Vizing's Theorem. The original Godsil and Ruciński bounds for Theorems 1.2, 1.3 are $\sigma^2 > |E|/(16\Delta^2)$ and $\sigma^2 > \nu/(36\Delta)$, the latter improvable to $\sigma^2 > \nu/(16\Delta)$ by substitution of Proposition 1.7 for Babai's $\mu \geq \nu/3$. The preceding argument does a little better than these, even if—to make "very simple" unexceptionable—we replace Vizing's Theorem by the trivial $\chi'(G) \leq 2\Delta - 1$.)

6 ξ vs. ψ

As mentioned in the Introduction, the next lemma, a sort of uncertainty principle for ξ and its "transform" ψ, is at the heart of our proofs of Theorems 1.6 and 1.10. That something of this nature should be true is not surprising, but it's interesting that, at least at present, such a circuitous approach to the variance of ξ seems the most effective.

Lemma 6.1 *There are positive constants A, B such that either*

$$\sigma^2 \geq \mu/A$$

or

$$\sigma_\psi^2 \sigma^2 \geq \mu^2/B.$$

Lemma 6.1 is valid as a purely numerical statement, as follows. Suppose X is a random variable taking values in $\{0, \ldots, \nu\}$ (ν just some natural number), with

$$p_k := \Pr(X = k) > 0 \qquad 0 \leq k \leq \nu.$$

Set

$$\overline{\psi}_k = \begin{cases} (k+1)p_{k+1}/p_k & \text{if } 0 \leq k \leq \nu - 1 \\ 0 & \text{if } k = \nu, \end{cases}$$

and define the random variable $\overline{\psi}$ by $\Pr(\overline{\psi} = \overline{\psi}_k) = p_k$. Notice that if $X = \xi$ as in our situation, then $\overline{\psi}$ has the same distribution as the conditional expectation $\mathbf{E}[\psi|\xi]$, so $\sigma_\psi^2 \geq \sigma_{\overline{\psi}}^2$.

(Of course this applies to any finite collection of sets $\mathcal{M}$ closed under taking subsets, with

$$\mathcal{M}_k := \{M \in \mathcal{M} : |M| = k\} \neq \emptyset \quad \text{iff} \quad 0 \leq k \leq \nu,$$

$$\psi(M) = |\{M' \in \mathcal{M}_{k+1} : M' \supset M\}| \;\; \text{for } M \in \mathcal{M}_k,$$

and X the size of M chosen uniformly from $\mathcal{M}$. Combinatorially this seems the natural level of generality for the present discussion.)

Thus, with $\mu := \mathbf{E}[X] = \mathbf{E}[\overline{\psi}]$, Lemma 6.1 is contained in

Lemma 6.2 *With notation as above, there are positive constants A, B such that either*

$$\sigma_X^2 \geq \mu/A \tag{29}$$

or

$$\sigma_{\overline{\psi}}^2 \sigma_X^2 \geq \mu^2/B. \tag{30}$$

This is, of course, more generality than we need, and in particular the special case of small μ is not relevant for us; but for completeness' sake we say the few extra words needed to prove the Lemma as stated.

It should be noted that the option (29) cannot be excluded. For example, if X has the binomial distribution $B(\nu, p)$, then, with $q = 1 - p$, we have $\mu = \nu p$, $\sigma_X^2 = \nu pq$ and $\sigma_{\overline{\psi}}^2 = (p/q)^2 \nu pq$, so (30) fails for small p. We don't know whether (29) is still needed when $X = \xi(G)$, but suspect it is not. (See also Conjecture 6.3 below.)

Proof of Lemma 6.2. We prove the Lemma with $A = 1000$ and $B = 4000$. These values are by no means optimal, but for present purposes there seems little point in trying to improve them.

Write σ for σ_X. We first show that we may assume σ is not too small. Suppose $\sigma \leq 1/4$. We may also suppose

$$\mu \geq 1/2,$$

since it's easy to see that otherwise $\sigma^2 \geq \mu/2$.

Now since

$$1/16 \geq \sigma^2 \geq \sum \{p_k/4 : |k - \mu| \geq 1/2\}, \tag{31}$$

there is a (unique) integer k_0 with $|k_0 - \mu| < 1/2$, and, clearly,

$$k_0 > 2\mu/3. \tag{32}$$

Let $k = k_0 - 1$. By (31),

$$p_{k_0} \geq 1 - 4\sigma^2 \geq 3/4 \qquad \text{and} \qquad p_k \leq 4\sigma^2 \leq 1/4. \tag{33}$$

Combining (32), (33), we have $\overline{\psi}_k = k_0 p_{k_0}/p_k \geq 2\mu$ and

$$\begin{aligned} \sigma_{\overline{\psi}}^2 &\geq p_k \left(\frac{k_0 p_{k_0}}{p_k} - \mu \right)^2 \geq p_k \left(\frac{1}{2} \frac{k_0 p_{k_0}}{p_k} \right)^2 \\ &= \frac{(k_0 p_{k_0})^2}{4 p_k} > \frac{(\mu/2)^2}{4 \cdot 4\sigma^2} = \frac{\mu^2}{64\sigma^2} \ . \end{aligned}$$

(So we have (30).)

We may thus suppose that

$$\sigma > 1/4. \tag{34}$$

Assume also that (29) and (30) fail; that is, that

$$\sigma^2 < \mu/A \qquad \text{and} \qquad \sigma_{\psi}^2\sigma^2 < \mu^2/B, \tag{35}$$

with A and B as above.

Notice first of all that there is some

$$b \geq \mu - 2\sigma \tag{36}$$

with

$$p_b \geq \frac{1}{8\sigma+1}, \tag{37}$$

since otherwise (using (34)),

$$\sigma^2 \quad \geq \quad (2\sigma)^2 \Pr(|X-\mu| \geq 2\sigma) \quad \geq \quad (2\sigma)^2(1-\frac{4\sigma+1}{8\sigma+1}) \quad > \quad \sigma^2.$$

Let us choose such a b with p_b as large as possible.

Suppose

$$p_k \geq p_b/2 \qquad \text{for} \quad a \leq k \leq b,$$

$$p_{a-1} < p_b/2.$$

Then, noting that the variance of uniform distribution on $\{a,\ldots,b\}$ is $(b-a+2)(b-a)/12$, and that we can partition our probability space as $Q \cup R$ with $\Pr(Q) = (b-a+1)p_b/2$ and $X|Q$ uniform on $\{a,\ldots,b\}$, we have

$$\sigma^2 \geq (b-a+2)(b-a+1)(b-a)p_b/24,$$

whence

$$(b-a)^3 < 24(8\sigma+1)\sigma^2 < 24 \cdot 12\sigma^3$$

(using (37) for the first inequality and (34) for the second). Thus

$$b - a < C'\sigma \tag{38}$$

and (by (36))

$$\mu - a < C\sigma, \tag{39}$$

where $C' = \sqrt[3]{288}$ and $C = C' + 2 < 9$.

Set $\gamma = C\sigma/\mu$, and note that (by (34) and (35))

$$\gamma < 36/A. \tag{40}$$

Define δ_k by

$$p_{k+1} = (1 + \delta_k)p_k.$$

(We think of δ_k as positive, though it will not always be so.)

Using (39), we have, for $k \geq a - 1$,

$$\begin{aligned} \overline{\psi}_k &= (k+1)p_{k+1}/p_k \;>\; (1-\gamma)\mu(1+\delta_k) \\ &= (1 + (\delta_k - \gamma - \gamma\delta_k))\mu \;=:\; (1+\zeta_k)\mu. \end{aligned} \tag{41}$$

Suppose first that

$$\delta_{a-1} \geq 1/2.$$

Then

$$\zeta_{a-1} = (1-\gamma)p_a/p_{a-1} - 1 \geq (\frac{1}{3} - \gamma)\frac{p_a}{p_{a-1}}$$

and

$$\begin{aligned} \frac{\mu^2}{B\sigma^2} \;>\; \sigma^2_{\overline{\psi}} \;&\geq\; p_{a-1}\zeta^2_{a-1}\mu^2 \;\geq\; (\frac{1}{3} - \gamma)^2 \frac{p_a^2\mu^2}{p_{a-1}} \\ &\geq\; (\frac{1}{3} - \gamma)^2 \frac{3}{2} p_a \mu^2 \;\geq\; (\frac{1}{3} - \gamma)^2 \frac{3\mu^2}{4(8\sigma + 1)} \;, \end{aligned}$$

contradicting (34) or (40).

We may therefore suppose that

$$\delta_{a-1} < 1/2.$$

Then by our choice of a, b,

$$2 < \prod_{k=a-1}^{b-1} (1 + \delta_k) < \frac{3}{2} \prod_{k=a}^{b-1} (1 + \delta_k),$$

whence

$$\sum_{k=a}^{b-1} \delta_k > \ln(4/3)$$

and (using (38), (40) and (35))

$$\begin{aligned}
\sum_{k=a}^{b-1} \zeta_k^+ &\geq (1-\gamma)\sum_{k=a}^{b-1} \delta_k - (b-a)\gamma \\
&> \ln(4/3) - \gamma(\ln(4/3) + C'\sigma) \\
&> \ln(4/3) - (36\ln(4/3) + C'C)/A \quad =: \quad X.
\end{aligned}$$

But then, finally (again using (41), (38) and (34)),

$$\begin{aligned}
\sigma_{\overline{\psi}}^2 &= \sum p_k(\overline{\psi}_k - \mu)^2 \\
&\geq \sum_{k=a}^{b-1} p_k(\zeta_k^+)^2\mu^2 \\
&\geq \frac{\mu^2}{2(8\sigma+1)} \sum_{k=a}^{b-1} (\zeta_k^+)^2 \\
&\geq \frac{\mu^2}{2(8\sigma+1)} \frac{1}{b-a} (\sum_{k=a}^{b-1} \zeta_k^+)^2 \\
&\geq \frac{\mu^2 X^2}{2(8\sigma+1)C'\sigma} \\
&\geq \mu^2/(B\sigma^2).
\end{aligned}$$

□

Though Lemma 6.2 is adequate for our purposes, it would be nice to have a more definitive version—for example under the additional assumption that the roots of $\sum p_k x^k$ are real—since the question the Lemma addresses seems a natural one. We mention one possibility in this direction, which would be of some interest for the case of regular graphs, as discussed below.

Conjecture 6.3 *Let X, $\overline{\psi}$ be as in Lemma 6.2, and suppose that the roots of $\sum_{k=0}^{\nu} p_k x^k$ are real. Then $\sigma_{\overline{\psi}}^2 \sigma_X^2 \geq \mu^4 \nu^{-2}$.*

This is sharp when X has the binomial distribution $B(\nu, p)$. (Note ξ_G is binomial if G is a disjoint union of triangles or of stars of a fixed size.) As far as we know, Conjecture 6.3 might even be true under the weaker assumption that $\{p_k\binom{\nu}{k}^{-1}\}$ is log concave, though we are inclined to doubt this.

As mentioned earlier, we will also need to consider more general ψ_F's. Before doing so, let us show that Lemma 6.1 is already enough to establish asymptotic normality, and even the correct order of magnitude of σ^2, for sequences of regular graphs. Namely, we show that there are positive constants a_1, a_2 such that for d-regular G,

$$a_1 v d^{-1/2} < \sigma^2 < a_2 v d^{-1/2} \ . \tag{42}$$

We will use (the upper bound from) the following rough version of Theorem 1.9(a).

Proposition 6.4 *There are positive constants c_1, c_2 such that for any d-regular G and $x \in V(G)$,*

$$c_1 d^{-1/2} < p(\overline{x}) < c_2 d^{-1/2}. \tag{43}$$

Of course we could just use Theorem 1.9(a) here, but want to emphasize that (42) can be proved relatively simply. (The arguments of [27] are as of now not very simple.) We are also inclined to record the proof of (42) in its original form.

Proof of Proposition 6.4 (sketch). With the pair G, v we associate the tree $T = T(G, v)$ whose vertices are the self-avoiding walks beginning at v, with the natural adjacencies (see [12]). In particular we may regard v itself as a vertex of T, and take it to be the root. It's then easy to see that $p_T(\overline{v}) = p_G(\overline{v})$ $(= p(\overline{v}))$.

Writing $T(x)$ for the subtree rooted at x, we compute $p(\overline{x})$'s recursively (beginning at the leaves) via (13), which in the present context becomes

$$p_{T(x)}(\overline{x}) = [1 + \sum\{p_{T(y)}(\overline{y}) : y \text{ a child of } x\}]^{-1} \ .$$

The proof of (43) is then an easy calculation which requires of T only the trivial inequalities

$$d - i \le d_{T(x)}(x) \le d - 1$$

for any x at depth $i \ge 1$.

□

The proofs of (both parts of) Theorem 1.9 are again based on T, but require more precise information on degrees in T—this is established using martingales—as well as considerably more delicate calculations.

Proof of (42). Of course the upper bound (with $a_2 = c_2/2$) is immediate from (43) and (7).

Let $E = \{\{x_i, y_i\} : i \in I\}$, set $X_i = 1_{\{x_i, y_i \not\sim M\}}$, and, for $i, j \in I$, write

$$i \not\sim j \quad \text{if } |\{x_i, y_i, x_j, y_j\}| = 4,$$

$$i \sim j \quad \text{if } |\{x_i, y_i, x_j, y_j\}| = 3.$$

Then

$$\begin{aligned} \mathbf{E}[\psi^2] &= \sum_i \sum_j \mathbf{E}[X_i X_j] \\ &= \sum_i \mathbf{E}[X_i] + \sum_{i \sim j} \mathbf{E}[X_i X_j] + \sum_{i \not\sim j} \mathbf{E}[X_i X_j]. \end{aligned} \tag{44}$$

Here the first term is $\mathbf{E}[\psi] = \mu$. The second term is the interesting one. We may bound it using (14) and Proposition 6.4:

$$\begin{aligned} \sum_{i \sim j} \mathbf{E}[X_i X_j] &= \sum_x \sum_{y \sim x} \sum_{y \sim z \neq x} p(\overline{x}, \overline{y}, \overline{z}) \\ &= \sum_x [(d+1)p(\overline{x}) - 1] \\ &< c_2 v \sqrt{d}, \end{aligned}$$

where x, y, z range over V.

For the third term in (44) we first observe that, for any fixed i, if we set $W = V \setminus \{x_i, y_i\}$, then Lemma 4.1 implies

$$\begin{aligned} \sum_{j \not\sim i} p(\overline{x}_j, \overline{y}_j | \overline{x}_i, \overline{y}_i) &= \sum_{j \not\sim i} p(x_j y_j | \overline{x}_i, \overline{y}_i) \\ &= \mu(W | \overline{x}_i, \overline{y}_i)/2 \\ &\leq (\mu(W) + 2)/2 \quad \leq \quad \mu + 1. \end{aligned}$$

This gives

$$\begin{aligned}\sum_i \sum_{j\not\sim i} \mathbf{E}[X_i X_j] &= \sum_i p(\overline{x}_i, \overline{y}_i) \sum_{j\not\sim i} p(\overline{x}_j, \overline{y}_j | \overline{x}_i, \overline{y}_i) \\ &\leq \sum_i p(\overline{x}_i, \overline{y}_i)(\mu+1) \quad = \quad \mu^2 + \mu.\end{aligned}$$

Combining the above bounds gives

$$\sigma_\psi^2 < c_2 v\sqrt{d} + 2\mu. \tag{45}$$

But then, noting that Proposition 1.7 implies $\mu = \Omega(v)$ for regular graphs, we may apply Lemma 6.1 to obtain

$$\sigma^2 = \Omega(vd^{-1/2}),$$

completing the proof of (42).

□

Remark If we replace Proposition 6.4 by Theorem 1.9(a), then (45) becomes $\sigma_\psi^2 \lesssim v\sqrt{d}$, which with Conjecture 6.3 would give $\sigma^2 \gtrsim v/(4\sqrt{d})$. This is the lower bound in Theorem 1.9(b). (Part (b) of Theorem 1.9 is quite a bit harder than part (a), so this derivation would be of some interest. Note that part (a) with (7) gives $\sigma^2 \lesssim v/(2\sqrt{d})$.)

We now extend Lemma 6.1 to ψ_F's. We write μ_F for $\mathbf{E}[|M \cap F|] = \mathbf{E}[\psi_F]$.

Lemma 6.5 *For A, B as in Lemma 6.1 and any $F \subseteq E$, either*

$$\sigma^2 \geq \mu_F/(8A)$$

or

$$\sigma_{\psi_F}^2 \geq \frac{1}{16}\mu_F^2 \min\{\frac{1}{B\sigma^2}, 1\}.$$

The second alternative here could almost certainly be replaced by

$$\sigma_{\psi_F}^2 \sigma^2 = \Omega(\mu_F^2). \tag{46}$$

In fact it's probably also true that the first alternative is unnecessary, and that (46) is *always* true. This possibility was already mentioned for Lemma 6.1

(the case $F = E$; see the remark following the statement of Lemma 6.2). We can at least show that such a strengthening of Lemma 6.1 implies (46) in general. Of course Lemma 6.2 gives the generalization of Lemma 6.1 obtained by replacing ξ by $\xi_F := |M \cap F|$ (and ψ by ψ_F, μ by μ_F and σ^2 by $\sigma_F^2 := Var[\xi_F]$); but note this doesn't imply Lemma 6.5, since σ_F can be much larger than σ.

It would be interesting to decide whether one can base a proof of Theorem 1.6 entirely on Lemma 6.5. In other words, is it true that if $\nu - \mu$ is large, then there must be *some* $F \subseteq E$ for which $\mu_F^2 \sigma_{\psi_F}^{-2}$ is also large?

Proof of Lemma 6.5. We condition on $N := M \setminus F$, and write $\mu_{F|N}$ for $\mathbf{E}[|M \cap F||N] = \mathbf{E}[\psi_F|N]$. Define events

$$\begin{aligned} Q &= \{\mu_{F|N} < \mu_F/2\}, \\ R &= \overline{Q} \wedge \{Var[\xi|N] \geq \mu_{F|N}/A\}, \\ S &= \overline{Q} \wedge \overline{R}. \end{aligned}$$

Then

$$\sigma_{\psi_F}^2 \geq \Pr(Q)\mu_F^2/4$$

and

$$\sigma^2 \geq \Pr(R)\mu_F/(2A).$$

On the other hand, if $N \in S$, then Lemma 6.1 (applied to the graph consisting of edges of F not met by edges of N) implies

$$Var[\psi_F|N]Var[\xi|N] \ \geq \ \mu_{F|N}^2/B \ \geq \ \mu_F^2/(4B),$$

whence (by Hölder's inequality)

$$Var[\psi_F|S]Var[\xi|S] \geq \mathbf{E}\{Var[\psi_F|N]|S\}\mathbf{E}\{Var[\xi|N]|S\} \geq \mu_F^2/(4B), \quad (47)$$

and

$$\sigma_{\psi_F}^2 \sigma^2 \geq (\Pr(S))^2 \mu_F^2/(4B).$$

So for example,

$$\begin{aligned} \sigma_{\psi_F}^2 &\geq \mu_F^2/16 \quad \text{if } \Pr(Q) \geq 1/4, \\ \sigma^2 &\geq \mu_F/(8A) \quad \text{if } \Pr(R) \geq 1/4, \\ \sigma_{\psi_F}^2 \sigma^2 &\geq \mu_F^2/(16B) \quad \text{if } \Pr(S) \geq 1/2. \end{aligned}$$

Since one of these must occur, we have Lemma 6.5.

□

7 Easier implications

(Here and in the next section, items (a)-(e) are those in Theorem 1.10.)

In this section we give the easier parts of the proof of Theorem 1.10, and also prove Proposition 1.11, by showing that (b) implies (e), (e) implies (d), and (d) implies both (c) and the "negation" of (9), namely,

(f) $\quad \tau_n < (1 - \Omega(1))\upsilon_n/2.$

Recall (c) trivially implies (b). In the next section we show that (a) implies (d), which will complete the proof of Theorem 1.10, since we already know that (b) implies (a).

As in Section 3, we may dispense with the sequence notation: to prove Theorem 1.10 it's enough to show that each of the quantities in (a)-(e) is bounded in terms of the others; while (d) $\Rightarrow$ (f) will follow from

$$\upsilon - 2\tau > \Omega((\kappa+1)^{-1})\upsilon - O(\kappa^2 + \kappa). \tag{48}$$

Thus we continue to work with a fixed G.

We first mention two examples which will help put our bounds in perspective.

Example 7.1 *Let* $V(G) = \{x_1, \ldots, x_n, y_1, \ldots, y_n\}$ *and*

$$E(G) = \{\{x_i, x_j\} : 1 \le i < j \le n\} \cup \{\{x_i, y_i\} : 1 \le i \le n\}.$$

Then $\nu - \mu = n/2$, $\sigma^2 = \sqrt{n}/2 - 1/2 + O(n^{-1/2})$ *and* $\lambda = \kappa = \sqrt{n} - 1 + O(n^{-1/2})$.

(Briefly because: For $\nu - \mu$ just observe that $p(\overline{x}_i) + p(\overline{y}_i) = 1$. For the other values set $p_n = p(\overline{x}_i)$. Then (13) gives $p_n = [(n-1)p_{n-1} + 2]^{-1}$ which inductively gives, e.g.,

$$(\sqrt{n+1} + 1)^{-1} < p_n \le (\sqrt{n} + 1)^{-1}$$

so in particular $p_n = n^{-1/2} - n^{-1} + O(n^{-3/2})$. This gives λ, κ, σ^2 as above: $\lambda = \kappa = np_n$ and (due to some cancellation of covariances) $\sigma^2 = np_n/2$.)

Example 7.2 *Let* $V(G) = X \cup Y \cup Z$ *with* X,Y,Z *pairwise disjoint and* $|X| = |Z| = n$, $|Y| = n + n^{2/3}$, *and let* $E(G)$ *consist of all edges joining* X

and Y, together with a perfect matching on Y and a perfect matching of X and Z. (We think of this as Example 1.5 with $\varepsilon = n^{-1/3}$, augmented by the two matchings.) Then σ^2 and λ are $\theta(n^{1/3})$, while $\kappa = \theta(n^{2/3})$ and $\nu - \mu = \theta(n)$. (If we delete Z, then $\nu - \mu$ drops to $\theta(n^{2/3})$, but the other values are not much affected.)

The bounds we establish for Theorem 1.10 are given in (52), (54), (56) and (85). They imply in particular

$$\left.\begin{array}{c}\sigma^2\\ \lambda\end{array}\right\} \;\leq\; \nu - \mu \;\leq\; \tau - \mu \;\leq\; \kappa^2/2 + O(\kappa) \tag{49}$$

and

$$\kappa \;\leq\; \left\{\begin{array}{l} O(\sigma^4 + \sigma^2) \\ O(\lambda^2 + \lambda). \end{array}\right. \tag{50}$$

These are all tight up to the implied constants. To see this, first note that for any disjoint union of stars, we actually have

$$\nu - \mu \;=\; \tau - \mu \;=\; \kappa \;=\; \lambda, \tag{51}$$

with $\sigma^2 = m(\nu - \mu)/(m+1)$ for m-stars. This shows that all bounds but the last in (49) are tight (except that σ^2 cannot be *equal* to $\nu - \mu$), and that *all* of the bounds in (49) and (50) are tight if any of the parameters in (51) is $O(1)$. Accuracy of the remaining inequalities is given by Example 7.1 for the last inequality in (49)—actually for the implied inequality $\nu - \mu \leq \kappa^2/2 + O(\kappa)$—and by Example 7.2 for the inequalities in (50).

As mentioned in Section 1, most of the remaining inequalities implied by (49) and (50) appear not to be tight, and it would be interesting to understand these relations more precisely. Some of these are probably not hard, though at the moment the only tight bound we know other than those above is the rather trivial $\kappa \leq \tau - \mu$. (For example, we don't even know, though we expect, that $\kappa = O(\nu - \mu)$.) Perhaps most interesting among these several questions are the conjecture (8) (note we have instead $\nu - \mu = O(\sigma^8)$), and

Question 7.3 *How closely related are σ^2 and λ? In particular, is it true that $\lambda = \Theta(\sigma^2)$ (that is, are there bounds on the ratios λ/σ^2 and σ^2/λ)?*

We now turn to the proofs. The implications given in this section are based on the results of Section 4. It is only in Section 8 that we require the lemmas of Sections 5 and 6.

Proof of (b) $\Rightarrow$ (e). Recalling that F_M is the set of edges not met by edges of M, and noticing that $\nu(F_M) \leq \nu - |M|$, we have

$$\lambda \leq \mathbf{E}[\nu - |M|] = \nu - \mu. \tag{52}$$

□

Proof of (d) $\Rightarrow$ (c). Suppose Y is a vertex cover of G with

$$\sum_{x \in Y} p(\overline{x}) = \kappa. \tag{53}$$

We show

$$\tau - \mu < (\kappa^2 + 3\kappa)/2. \tag{54}$$

Let

$$T_M = |\{x \in Y : x \not\sim M\}|,$$
$$Q_M = |M \cap E(Y)|.$$

(Recall $E(Y)$ is the set of edges contained in Y.) Then $|M| = |Y| - (T_M + Q_M)$ and

$$\tau - \mu \ \leq \ |Y| - \mathbf{E}[|M|] \ = \ \mathbf{E}[T_M] + \mathbf{E}[Q_M] \ = \ \kappa + \mathbf{E}[Q_M].$$

But

$$\begin{aligned} \mathbf{E}[Q_M] &= \frac{1}{2} \sum_{x \in Y} \sum_{x \sim y \in Y} p(\overline{x}, \overline{y}) \\ &= \frac{1}{2} [\sum_{x \in Y} p(\overline{x}) (\sum_{x \sim y \in Y} (p(\overline{y}|\overline{x}) - p(\overline{y})) + \sum_{x \sim y \in Y} p(\overline{y}))] \\ &< \frac{1}{2} [\sum_{x \in Y} p(\overline{x}) + (\sum_{x \in Y} p(\overline{x}))^2] \\ &= (\kappa + \kappa^2)/2, \end{aligned} \tag{55}$$

so we have (54). (We used Lemma 4.1 for the inequality. The slightly worse bound $\tau - \mu < \kappa^2 + \kappa$ could be gotten a little more easily using Corollary 4.3.)

□

Proof of (d) ⇒ (f). Again suppose Y is a cover with (53). Let $Z = V \setminus Y$. Then for each $z \in Z$, by Lemma 4.1,

$$p(\overline{z}) \;=\; [1 + \sum_{y \sim z} p(\overline{y}|\overline{z})]^{-1} \;\geq\; [2 + \sum_{y \sim z} p(\overline{y})]^{-1} \;\geq\; (\kappa + 2)^{-1}.$$

Combining this with (54) and (55), we have

$$\frac{\kappa+1}{\kappa+2}|Z| \;\geq\; \sum_{z \in Z} p(z) \;=\; \mu - \mathbf{E}[Q_M] \;>\; \tau - (\kappa^2 + 2\kappa)$$

and

$$\begin{aligned} v - 2\tau \;&\geq\; v - 2\mu - (\kappa^2 + 3\kappa) \\ &\geq\; \sum_{z \in Z} p(\overline{z}) - (\kappa^2 + 3\kappa) \\ &\geq\; (\kappa + 2)^{-1}|Z| - (\kappa^2 + 3\kappa) \\ &\geq\; \frac{1}{\kappa+1}(\tau - (\kappa^2 + 2\kappa)) - (\kappa^2 + 3\kappa), \end{aligned}$$

which implies (48).

□

When $\tau = \omega(\kappa^3)$ the bound here is essentially $v - 2\tau > \tau/\kappa$, which is about right, e.g. for the complete bipartite graphs of Example 1.5. We haven't worked out what the truth is when τ is smaller relative to κ.

Proof of (e) ⇒ (d). We will find a vertex cover Y with

$$\sum_{x \in Y} p(\overline{x}) = O(\lambda^2 + \lambda). \tag{56}$$

Set $\varepsilon = [\max\{8\lambda + 1, 129\}]^{-1}$ and let

$$W = \{x : p(\overline{x}) \geq \varepsilon\}.$$

Then for any matching $N \subseteq E(W)$,

$$\lambda \geq \mathbf{E}[\nu(F_M \cap E(W))] \geq \mathbf{E}[|M \cap N|].$$

So by Corollary 4.7, there is a cover Z of $E(W)$ with

$$\sum_{x \in Z} p(\overline{x}) \le 4\lambda/\varepsilon. \tag{57}$$

Set $X = V \setminus W = \{x : p(\overline{x}) < \varepsilon\}$, and suppose

$$\sum_{x \in X} p(\overline{x}) = T.$$

We will show that

$$T \le 16\lambda, \tag{58}$$

which with (57) gives (56) for $Y = Z \cup X$.

Proof of (58). Assume (58) is false, and set $\eta = (1-\varepsilon)/\varepsilon = \max\{8\lambda, 128\}$. For $x \in X$ define the random variable

$$\zeta_x = |\{y \sim x : y \not\sim M\}|,$$

and let

$$\eta_x = \mathbf{E}[\zeta_x | \overline{x}] = \sum_{y \sim x} p(\overline{y}|\overline{x}) = p(\overline{x})^{-1} - 1 > \eta$$

(see (13)). By Corollary 4.2, $Var[\zeta_x|\overline{x}] \le 2\eta_x$, so Chebyshev's inequality gives, for any $\delta > 0$,

$$\Pr(\zeta_x < (1-\delta)\eta|\overline{x}) \;\le\; \Pr(\zeta_x < (1-\delta)\eta_x|\overline{x}) \;<\; \frac{2}{\delta^2\eta_x} \;<\; \frac{2}{\delta^2\eta},$$

and

$$\Pr(x \not\sim M,\, \zeta_x \ge (1-\delta)\eta) \;>\; p(\overline{x})(1 - \frac{2}{\delta^2\eta}).$$

Thus, (taking $\delta = 1/2$ and) setting

$$R = R_M = |\{x \in X : x \not\sim M, \zeta_x \ge \eta/2\}|$$

and $\beta = 8/\eta$, we have

$$\mu_R \;>\; \sum_{x \in X} p(\overline{x})(1-\beta) \;=\; (1-\beta)T. \tag{59}$$

Notice that matching vertices of X greedily in F_M shows that

$$\nu(F_M) \geq \frac{1}{2}\min\{R, \eta/2\}. \tag{60}$$

But λ, the expected value of $\nu(F_M)$, is much less than $\eta/4$. So if R is reasonably concentrated then $\mu_R/2 \approx T/2$ should be close to a lower bound on λ.

To make this idea concrete, set $Q = |\{x \in X : x \not\sim M\}|$. Then $Q \geq R$, $\mathbf{E}[Q] = T$,

$$\mu_{Q-R} < \beta T \tag{61}$$

(by (59)), and, again using Corollary 4.2,

$$\sigma_Q^2 \leq 2T. \tag{62}$$

Suppose first that

$$T \geq \eta \tag{63}$$

and set $T - \eta/2 = Y$. By (60),

$$\lambda \geq \Pr(R \geq \eta/2)\eta/4. \tag{64}$$

On the other hand, using (62) and (61) (in conjunction with Chebyshev's and Markov's inequalities) and the fact that $Y \geq T/2$, we have

$$\begin{aligned}
\Pr(R < \eta/2) &= \Pr(T - R > Y) \\
&\leq \Pr(T - Q > Y/2) + \Pr(Q - R > Y/2) \\
&< \frac{2T}{(Y/2)^2} + \frac{2\beta T}{Y} \\
&\leq 32/T + 32/\eta \ \leq \ 64/\eta.
\end{aligned}$$

Combining this with (64) leads to the contradiction $\lambda > (1 - 64/\eta)\eta/4 \geq \lambda$.

Thus we cannot have (63) (note this already gives (58) when $\lambda \geq 16$) and may suppose instead that

$$T < \eta. \tag{65}$$

Set

$$\begin{aligned}
A &= \Pr(R \leq 2\eta)\mathbf{E}[R|R \leq 2\eta] \\
B &= \Pr(R > 2\eta)\mathbf{E}[R|R > 2\eta].
\end{aligned}$$

Now (60) implies $\nu(F_M) \geq R/8$ whenever $R \leq 2\eta$, so that

$$\lambda = \mathbf{E}[\nu(F_M)] \geq A/8. \tag{66}$$

On the other hand, by (59),

$$(1-\beta)T < \mu_R = A + B, \tag{67}$$

so an upper bound on B gives a lower bound on A. We may derive such an upper bound from (62): Notice that (65) gives $Q \leq (2/\eta)(Q-T)^2$ whenever $Q \geq 2\eta$. Thus (recalling that $Q \geq R$ and $\mu_Q = T$)

$$\begin{aligned} B &\leq \Pr(R > 2\eta)\mathbf{E}[Q|R > 2\eta] \\ &\leq \Pr(R > 2\eta)\frac{2}{\eta}\mathbf{E}[(Q-T)^2|R > 2\eta] \\ &\leq \frac{2}{\eta}\sigma_Q^2 \leq \frac{4T}{\eta}. \end{aligned}$$

Finally, combining this with (66) and (67) gives

$$\lambda \geq (1 - 8/\eta - 4/\eta)T/8 \geq T/16,$$

so we have (58).

□

8 (a) ⇒ (d)

In this section we complete the proof of Theorem 1.10 (and of Theorem 1.6) by showing that (a) implies (d). The proof produces a vertex cover Y with $\sum_{y \in Y} p(\overline{y}) = O(\sigma^4)$ unless σ is small (say $\sigma < 1$), in which case $\sum_{y \in Y} p(\overline{y}) = O(\sigma^2)$. As noted in Section 7, both these bounds are essentially best possible.

Set $\alpha_i = 2^i/v$ for $i \geq 0$, and for $i \geq 1$,

$$S_i = \{x \in V : \alpha_{i-1} \leq p(\overline{x}) < \alpha_i\}.$$

(Note $\cup S_i = V$, e.g. by (13).) Denote by E_i the set of edges meeting S_i, and set $\psi_i = \psi_{E_i}$, $\mu_i = \mu_{E_i} = \mathbf{E}[|M \cap E_i|] = \mathbf{E}[\psi_i]$, and $\sigma_i^2 = Var(\psi_i)$.

Notice that

$$\mu_i \le |S_i|, \tag{68}$$

and

$$\mu_i \ge \frac{1}{2}\sum_{x\in S_i} p(x) \ge (1-\alpha_i)|S_i|/2. \tag{69}$$

Let

$$\varepsilon = \min\left\{\frac{1}{16B\sigma^2}, \frac{1}{16}\right\}.$$

(We retain the constants A, B from Lemma 6.1.) We first show that if

$$\alpha_i < \varepsilon/100, \tag{70}$$

then

$$|S_i| \le 16\varepsilon^{-1}\alpha_i^{-1}, \tag{71}$$

and consequently,

$$\sum_{x\in S_i} p(\overline{x}) < 16\varepsilon^{-1}. \tag{72}$$

The proof of this is based on some analysis of σ_i^2. We first deduce from Lemma 6.5 that, except in easy cases, σ_i^2 is at least about $\varepsilon|S_i|^2$. We then proceed roughly as we did in the regular case (Section 6) to show that $\sigma_i^2 = O(|S_i|\alpha_i^{-1})$.

By Lemma 6.5 we know that either

$$\sigma^2 \ge \mu_i/(8A) \tag{73}$$

or

$$\sigma_i^2 \;\ge\; \frac{1}{16}\mu_i^2 \min\left\{\frac{1}{B\sigma^2}, 1\right\} \;=\; \varepsilon\mu_i^2. \tag{74}$$

Since (73) gives (71) immediately (because of (69)), we may assume we have (74).

Suppose

$$E_i = \{\{x_j, y_j\} : j \in J\},$$

let $X_j = 1_{\{x_j, y_j \not\sim M\}}$, and write

$$\begin{array}{ll} j \not\sim k & \text{if } |\{x_j, y_j, x_k, y_k\}| = 4 \\ j \sim k & \text{if } |\{x_j, y_j, x_k, y_k\}| = 3. \end{array}$$

Then

$$\mathbf{E}[\psi_i^2] = \sum_j \mathbf{E}[X_j] + \sum_{j\sim k} \mathbf{E}[X_jX_k] + \sum_{j\not\sim k} \mathbf{E}[X_jX_k].$$

The first term here is just μ_i, and we will show that the third term is not much more than μ_i^2. The quantity of interest is the second term: σ_i^2 cannot satisfy (74) unless the second term is at least about $\varepsilon\mu_i^2$. (The reader might find it helpful to check that in Example 1.5, the second term in the analogous expansion of $\mathbf{E}[\psi^2]$ is indeed on the order of μ^2.)

For the third term we use Corollary 4.6 (twice), as well as Lemma 4.1, both with $W = W_j := S_i \setminus \{x_j, y_j\}$ in the j^{th} summand:

$$\begin{aligned}
\sum_{j\not\sim k} \mathbf{E}[X_jX_k] - \mu_i^2 &\le \sum_j p(\overline{x}_j, \overline{y}_j) \sum_{k\not\sim j} (p(\overline{x}_k, \overline{y}_k | \overline{x}_j, \overline{y}_j) - p(\overline{x}_k, \overline{y}_k)) \\
&= \sum_j p(\overline{x}_j, \overline{y}_j) \sum_{k\not\sim j} [(p(\overline{x}_k, \overline{y}_k | \overline{x}_j, \overline{y}_j) - p(\overline{x}_k, \overline{y}_k | \overline{x}_j)) \\
&\qquad\qquad + (p(\overline{x}_k, \overline{y}_k | \overline{x}_j) - p(\overline{x}_k, \overline{y}_k))] \\
&\le \sum_j p(\overline{x}_j, \overline{y}_j) \cdot 4 \sum_{w\in W_j} (p(\overline{w}|\overline{x}_j) + p(\overline{w})) \\
&\le \sum_j p(\overline{x}_j, \overline{y}_j)(8 \sum_{w\in W_j} p(\overline{w}) + 4) \\
&\le \mu_i(8|S_i|\alpha_i + 4) \\
&\le 8|S_i|^2\alpha_i + 4|S_i|.
\end{aligned}$$

Thus the burden falls on the second term:

$$\sum_{j\sim k} \mathbf{E}[X_jX_k] \ge \varepsilon\mu_i^2 - \mu_i - 8|S_i|^2\alpha_i - 4|S_i|. \tag{75}$$

Now

$$\sum_{j\sim k} \mathbf{E}[X_jX_k] = \sum\{p(\overline{x}, \overline{y}, \overline{z}) : x, y, z \text{ distinct}, \ xy, yz \in E_i\}. \tag{76}$$

We show that most of this must come from terms with $y \in S_i$. To see this, note that the right hand side of (76) is at most $\Sigma_1 + \Sigma_2$, where

$$\Sigma_1 = \sum\{p(\overline{x}, \overline{y}, \overline{z}) : y \in S_i,\ x \sim y \sim z\},$$

$$\Sigma_2 = \sum\{p(\overline{x}, \overline{y}, \overline{z}) : x, z \in S_i,\ y \sim x\}.$$

(We could also require $y \sim z$ here.)

As mentioned above, Σ_2 is not very big:

$$\begin{aligned}\Sigma_2 &= \sum_{x\in S_i}\sum_{y\sim x} p(\overline{x},\overline{y}) \sum_{z\in S_i} p(\overline{z}|\overline{x},\overline{y}) \\ &= \sum_{x\in S_i}\sum_{y\sim x} p(\overline{x},\overline{y}) \sum_{z\in S_i} (p(\overline{z}) + (p(\overline{z}|\overline{x},\overline{y}) - p(\overline{z}))) \\ &\leq \sum_{x\in S_i} p(x)(\sum_{z\in S_i} p(\overline{z}) + 4) \\ &\leq |S_i|(|S_i|\alpha_i + 4). \end{aligned} \tag{77}$$

Here we used Lemma 4.1 for the first inequality.

Combining (77) with (75) we have

$$\Sigma_1 \geq \varepsilon\mu_i^2 - (\mu_i + 9|S_i|^2\alpha_i + 8|S_i|). \tag{78}$$

On the other hand (again using Lemma 4.1, together with (13)),

$$\begin{aligned}\Sigma_1 &= \sum_{y\in S_i}\sum_{x\sim y} p(\overline{x},\overline{y}) \sum_{z\sim y} (p(\overline{z}|\overline{y}) + (p(\overline{z}|\overline{x},\overline{y}) - p(\overline{z}|\overline{y}))) \\ &\leq \sum_{y\in S_i}\sum_{x\sim y} p(\overline{x},\overline{y})(\sum_{z\sim y} p(\overline{z}|\overline{y}) + 2) \\ &= \sum_{y\in S_i} p(y)(p(\overline{y})^{-1} + 1) \\ &< 2|S_i|\alpha_i^{-1} . \end{aligned} \tag{79}$$

Combining this with (78) and using (68), (69), (70) yields, after a little rearranging,

$$|S_i| \; < \; (\varepsilon(1-\alpha_i)^2/4 - 9\alpha_i)^{-1}(2\alpha_i^{-1} + 9) \; < \; 16\varepsilon^{-1}\alpha_i^{-1},$$

so we have (71).

□

Let t be the largest index i for which $\alpha_i < \varepsilon/100$, or $t = 0$ if $\alpha_0 \geq \varepsilon/100$, and set $Z = V \setminus \cup_{i=1}^t S_i$. In addition, set

$$i_0 = \min\{i : \sum_{j<i}\sum_{x\in S_j} p(\overline{x}) > 1/(3\alpha_i)\},$$

or $i_0 = \infty$ if the set in question is empty, and $s = \min\{i_0 - 1, t\}$.

The precise reason for these definitions will appear below. Roughly, the parameters are chosen so that for any $x \in S_1, \ldots, S_s$, a substantial part of $\varphi(x)$ will be contributed by vertices of Z. (Recall—see (12)—$\varphi(x) = \sum_{y\sim x} p(\overline{y})$.)

We show that for $i \le s$ and $x \in S_i$,

$$\sum\{p(\overline{z}) : x \sim z \in Z\} > 1/(3p(\overline{x})). \tag{80}$$

Proof. By Corollary 4.5,

$$\sum_{y\sim x} p(\overline{y}) = \varphi(x) \ge p(x)/p(\overline{x}) - 1.$$

On the other hand,

$$\begin{aligned}\sum_{x\sim y\notin Z} p(\overline{y}) &= \sum\{p(\overline{y}) : y \in S_1 \cup \cdots \cup S_{i-1}\} + \sum\{p(\overline{y}) : y \in S_i \cup \cdots \cup S_t\} \\ &\le 1/(3\alpha_i) + (t - i + 1)16\varepsilon^{-1},\end{aligned}$$

using $i < i_0$ for the first term and (72) for the second. Now $t - i + 1 = \log(2\alpha_t/\alpha_i) \le \log(\varepsilon/(50\alpha_i))$, and a little calculation, using $i \le t$, shows that

$$\begin{aligned}\sum_{x\sim z\in Z} p(\overline{z}) &\ge p(x)/p(\overline{x}) - 1 - 1/(3\alpha_i) - 16\varepsilon^{-1}\log(\varepsilon/(50\alpha_i)) \\ &> 1/(3p(\overline{x})),\end{aligned}$$

as promised.

□

Set $S^* = S_1 \cup \cdots \cup S_s$. For $x \in S^*$ and summations running over $z \in N_Z(x) := \{z \in Z : x \sim z\}$, it follows from (80) and Corollary 4.3 that

$$\begin{aligned}\sum p(xz)^2 &= \sum p(\overline{x}, \overline{z})^2 \\ &= p(\overline{x})^2 \sum p(\overline{z}|\overline{x})^2 \\ &\ge (1/4)p(\overline{x})^2 \sum p(\overline{z})^2 \\ &\ge (1/4)p(\overline{x})^2 (\sum p(\overline{z}))^2/|N_Z(x)| \\ &\ge 1/(36|N_Z(x)|).\end{aligned} \tag{81}$$

But since $z \in Z$ implies $p(\overline{z}) \geq \varepsilon/200$, we may bound $|N_Z(x)|$ (using Corollary 4.5) via

$$1 \;\geq\; p(\overline{x})\varphi(x) \;\geq\; p(\overline{x}) \sum_{z \in N_Z(x)} p(\overline{z}) \;\geq\; p(\overline{x})(\varepsilon/200)|N_Z(x)|.$$

Inserting this in (81) we have

$$\sum_{z \in N_Z(x)} p(xz)^2 \geq \varepsilon p(\overline{x})/(7200)$$

and

$$\sum_{x \in S^*} \sum_{z \in N_Z(x)} p(xz)^2 \geq \varepsilon/(7200) \sum_{x \in S^*} p(\overline{x}),$$

which by Corollary 5.2 implies

$$\sum_{x \in S^*} p(\overline{x}) \;\leq\; 7200\varepsilon^{-1} \cdot 2\sigma^2 \;=:\; D\sigma^2\varepsilon^{-1}. \tag{82}$$

This is our main inequality. It says in particular that we can afford to take our cover Y to contain *all* of S^*.

In fact when $s = i_0 - 1 < t$, we may take Y to contain $V \setminus Z \setminus S^* = S_{s+1} \cup \cdots \cup S_t$ as well. For in this case, (82) gives

$$\varepsilon/(100) \;>\; \alpha_{s+1} \;>\; (1/3)[\sum_{x \in S^*} p(\overline{x})]^{-1} \;\geq\; \varepsilon/(3D\sigma^2), \tag{83}$$

and then (again using (72))

$$\begin{aligned} \sum\{p(\overline{y}) : y \in \cup_{i=s+1}^{t} S_i\} &\;\leq\; 16\varepsilon^{-1}(t-s) \\ &\;=\; 16\varepsilon^{-1}\log(\alpha_t/\alpha_s) \\ &\;\leq\; 16\varepsilon^{-1}\log\left(\frac{\varepsilon}{100}\frac{6D\sigma^2}{\varepsilon}\right). \end{aligned} \tag{84}$$

This is $O(\sigma^2 \log \sigma)$ if $\sigma > 2$ (say), and $O(\sigma^2)$ otherwise (the latter since (83) implies $\sigma = \Omega(1)$).

It remains to cover the edges contained in Z, which turns out to be easy: since

$$p(\overline{x}) \geq \varepsilon/200 \text{ for all } x \in Z,$$

Lemma 5.1 allows us to apply Corollary 4.7 with $W = Z$, $\alpha = \varepsilon/200$ and $C = 2\sigma^2$ to conclude that there is a cover X of Z with

$$\sum_{x \in X} p(\overline{x}) \leq 1600\sigma^2\varepsilon^{-1}.$$

So, finally, setting $Y = (V \setminus Z) \cup X$, we have

$$\sum_{y \in Y} p(\overline{y}) \;=\; \begin{cases} O(\sigma^4) & \text{if } \sigma \geq 1 \\ O(\sigma^2) & \text{if } \sigma < 1, \end{cases} \tag{85}$$

and in particular (a) $\Rightarrow$ (d).

$\square$

9 Asymptotically Poisson

Here we prove Theorem 1.12. Set $\overline{\mu} = \nu - \mu$ and $\alpha = \max_{A \in E} p(A)$. According to formula (1.23) of [1],

$$\sum_k |p_k - e^{-\overline{\mu}}\, \overline{\mu}^k/k!| < (1 - e^{-\overline{\mu}})\frac{\overline{\mu} - \sigma^2}{\overline{\mu}}\ .$$

So for Theorem 1.12 it is enough to show, assuming $C^{-1} < \overline{\mu} < C$ for some positive constant C, that

$$\overline{\mu} - \sigma^2 \to 0 \;\text{ as }\; \alpha \to \emptyset.$$

Note first that, setting $\delta = \sqrt{2\alpha}$, Corollary 4.3 gives for any $A = \{x, y\} \in E$,

$$\alpha \geq p(A) = p(\overline{x}, \overline{y}) \geq \frac{1}{2}p(\overline{x})p(\overline{y}).$$

Thus $Y := \{x \in V : p(\overline{x}) \leq \delta\}$ is a cover. Moreover, we have shown in Section 8 that for small enough δ

$$\sum_{x \in Y} p(\overline{x}) = O(1). \tag{86}$$

(Namely, with the parameters of Section 8, (82) and (84) give, provided $\delta < \varepsilon/(200)$,

$$\sum_{x \in Y} p(\overline{x}) \le \sum \{p(\overline{x}) : p(\overline{x}) < \varepsilon/200\} < O(\sigma^4 + \sigma^2). \tag{87}$$

This gives (86), since $\sigma^2 \le \overline{\mu} = O(1)$ implies that $\varepsilon = \Omega(1)$ and that the right hand side of (87) is $O(1)$.)

For distinct $x, y \in V$, set

$$\zeta_x = 1_{\{x \prec M\}}, \qquad \zeta_{xy} = 1_{\{xy \in M\}}.$$

Then since Y is a cover we have

$$\xi = |Y| - (\sum_{x \in Y} \zeta_x + \sum_{xy \in E(Y)} \zeta_{xy}), \tag{88}$$

whence

$$\overline{\mu} \;\le\; |Y| - \mu \;=\; \sum_{x \in Y} p(\overline{x}) + \sum_{xy \in E(Y)} p(xy) \;=:\; K.$$

Thus, since $\sigma^2 \le \overline{\mu}$, it is enough to show

$$K - \sigma^2 \to 0 \;\text{ as }\; \alpha \text{ (or } \delta) \to 0. \tag{89}$$

Now by (88), $\sigma^2 = \Sigma_1 + \Sigma_2 + \Sigma_3$, where

$$\begin{aligned} \Sigma_1 &= \sum_{x \in Y} \sum_{y \in Y} (p(\overline{x}, \overline{y}) - p(\overline{x}) p(\overline{y})), \\ \Sigma_2 &= \sum_{x \in Y} \sum_{yz \in E(Y)} (p(\overline{x}, yz) - p(\overline{x}) p(yz)), \\ \Sigma_3 &= \sum_{xy \in E(Y)} \sum_{zw \in E(Y)} (p(xy, zw) - p(xy) p(zw)), \end{aligned}$$

and we should show that each of $\Sigma_1 - \sum_{x \in Y} p(\overline{x})$, Σ_2, and $\Sigma_3 - \sum\{p(xy) : xy \in E(Y)\}$ tends to 0 as $\alpha \to \infty$. We will actually show that the absolute value of each is $O(\delta)$.

These calculations should be fairly routine by now, and we will run through them fairly quickly. We make repeated use of (86) without further notice.

First,

$$\Sigma_1 = \sum_{x\in Y} p(\overline{x})[(1-p(\overline{x})) + \sum_{y\in Y\setminus\{x\}} (p(\overline{y}|\overline{x}) - p(\overline{y}))],$$

so that by Corollary 4.4 and the definition of Y,

$$\begin{aligned} |\Sigma_1 - \sum_{x\in Y} p(\overline{x})| &< \sum_{x\in Y} p(\overline{x})[\delta + 2\sum_{y\in Y} p^2(\overline{y})] \\ &< \delta \sum_{x\in Y} p(\overline{x})(1+2\sum_{y\in Y} p(\overline{y})) \;=\; O(\delta). \end{aligned}$$

Next,

$$\Sigma_2 = \sum_{x\in Y} p(\overline{x})[\sum_{yz\in E(Y\setminus\{x\})} (p(yz|\overline{x}) - p(yz)) - \sum_{y\in Y\setminus\{x\}} p(xy)].$$

The first of the inner sums is at most

$$\frac{1}{2}\sum_{y\in Y\setminus\{x\}}\sum_{z\in Y\setminus\{x,y\}} (p(\overline{y},\overline{z}|\overline{x}) - p(\overline{y},\overline{z})) =$$

$$\frac{1}{2}\sum_{y\in Y\setminus\{x\}}\sum_{z\in Y\setminus\{x,y\}} [(p(\overline{y}|\overline{x}) - p(\overline{y}))p(\overline{z}|\overline{x},\overline{y}) + p(\overline{y})(p(\overline{z}|\overline{x},\overline{y}) - p(\overline{z}|\overline{y}))],$$

which by Corollaries 4.3 and 4.4 is at most

$$\frac{1}{2}\sum_{y\in Y\setminus\{x\}}\sum_{z\in Y\setminus\{x,y\}} (2p^2(\overline{y})\cdot 4p(\overline{z}) + p(\overline{y})\cdot 8p^2(\overline{z})) \;\le\; 8\delta(\sum_{y\in Y} p(\overline{y}))^2 \;=\; O(\delta).$$

The second inner sum is (again using Corollary 4.3)

$$\sum_{y\in Y\setminus\{x\}} p(\overline{x},\overline{y}) \le 2p(\overline{x})\sum_{y\in Y} p(\overline{y}) = O(\delta).$$

So we have $\Sigma_2 = O(\delta)$.

For Σ_3 we set $E(Y) = \{\{x_j, y_j\} : j \in J\}$ and, as earlier, write

$$\begin{aligned} j \not\sim k &\quad \text{if } |\{x_j, y_j, x_k, y_k\}| = 4 \\ j \sim k &\quad \text{if } |\{x_j, y_j, x_k, y_k\}| = 3. \end{aligned}$$

Then

$$\Sigma_3 = \sum_j p(\overline{x}_j, \overline{y}_j)[1 - p(\overline{x}_j, \overline{y}_j) + \sum_{k \not\sim j} (p(\overline{x}_k, \overline{y}_k | \overline{x}_j, \overline{y}_j) - p(\overline{x}_k, \overline{y}_k)) - \sum_{k \sim j} p(\overline{x}_k, \overline{y}_k)].$$

Here we need to observe that (86) implies, via Corollary 4.3, that

$$\sum_j p(\overline{x}_j, \overline{y}_j) \leq 2(\sum_{x \in Y} p(\overline{x}))^2 = O(1).$$

The proof that $|\Sigma_3 - \sum_j p(\overline{x}_j, \overline{y}_j)| = O(\delta)$ then parallels the preceding cases and is left to the reader.

10 Concluding remarks

In this section we briefly explore a few of the many possibilities suggested by the present work. Our main theme—though we will manage some digressions—is that it should be possible to prove central limit theorems in some extremely general combinatorial situations (more or less) analogous to that of Theorems 1.6, 1.10. I should warn the reader that many of these questions have not been thought about very seriously, so there may be some silly ones. Almost all do seem quite interesting if one conditions on their being sensible.

Before beginning, let us just recall that there are a number of unresolved issues from the main part of the paper, notably relations between our various parameters (see (8) and Section 7; it would also be of interest to identify other parameters which are similarly tied to the present ones); some questions from Section 6, especially Conjecture 6.3 and the question preceding the proof of Lemma 6.5; and necessity of the condition of Theorem 1.12. It would also be nice to find a simpler proof of Theorem 1.6: I very much doubt that the present one is optimal, though, as mentioned earlier, the fact that $\nu - \mu$ can be much larger than σ^2 suggests some lower bound on the difficulty of the problem. Of course one may also ask about analogues of Theorem 1.12, but we will not go into this here.

It should be noted immediately that in most of the situations described here, we lose (3), and with it the easy derivation of asymptotic normality from large variance. So we must hope to establish asymptotic normality by other means, the most obvious candidates being Stein's method [45] and martingales (e.g. [15]; see the discussion preceding Problem 10.10).

In what follows, ξ will always be an $\mathbf{N}$-valued random variable. We write $r = r_\xi$ for $\max \xi$—the *rank* of ξ—and, as earlier, μ and σ^2 for mean and variance of ξ. We now use $p_k(\xi)$, in place of $p_k(G)$, for $\Pr(\xi = k)$. For a sequence $\{\xi_n\}$ we again use r_n etc. in the natural way.

The possibilities here seem nearly endless. For instance, one could consider almost any special case of the following broad generalization of Theorem 1.6, formulated with Vera Sós. Let $\mathcal{G}$ be a fixed finite set of connected graphs such that any connected subgraph of a graph in $\mathcal{G}$ is also in $\mathcal{G}$. For a graph G, let $\mathcal{M}_\mathcal{G}(G)$ be the set of subgraphs of G having all components in $\mathcal{G}$, and let $\xi_\mathcal{G}(G)$ be the size (meaning number of edges, though one could also consider, e.g., the number of components) of a uniformly chosen member of $\mathcal{M}_\mathcal{G}(G)$.

Conjecture 10.1 *For any $\mathcal{G}$, sequence $\{G_n\}$ of graphs and $\xi_n = \xi_\mathcal{G}(G_n)$, asymptotic normality of $\{p_k(\xi_n)\}$ is equivalent to each of the conditions $\sigma_n \to \infty$, $r_n - \mu_n \to \infty$.*

Hard-core distributions on matchings in graphs

Let $G = (V, E)$ be a weighted graph (that is, with a weight function $\alpha : E \to \mathbf{R}^+$), and let $\xi = |M|$, with M drawn from the associated hard-core distribution (h.c.d.) on $\mathcal{M}(G)$ (see Section 2).

As (3) remains valid for h.c.d.'s (it is given in this generality in [18], [29]), we again have asymptotic normality of ξ as $\sigma \to \infty$. On the other hand, it's easy to see that none of the conditions $\nu - \mu \to \infty$, $\kappa \to \infty$, $\lambda \to \infty$ could reasonably be expected to imply asymptotic normality (For instance, let G_n consist of n copies of a path of length 3 in which the middle edge has weight n and the others weight 1. Then $\nu_n - \mu_n = n$ and $\sigma_n^2 \approx 2$.) At this writing we have no plausible substitute for any of these parameters, though we hope that at least for bipartite graphs, $\nu - \mu$ may be replaced by

$$\gamma(G) := \min_{E' \subseteq E} \{\nu(E') - \mathbf{E}[|M \cap E'|] + \mathbf{E}[|M \setminus E'|]\}. \tag{90}$$

Conjecture 10.2 *With notation as above and G_n bipartite, $\{p_k(\xi_n)\}$ is asymptotically normal as $\gamma_n \to \infty$.*

(It is not hard to see that $\sigma_n \to \infty$ implies $\gamma_n \to \infty$ for general weighted graphs. For a (nonbipartite) counterexample to the converse, let G_n consist of n disjoint copies of the graph with vertices $x_1, x_2, x_3, y_1, y_2, y_3$, edges x_ix_j and x_iy_i, and weights $\alpha(x_ix_j) = n^3$, $\alpha(x_iy_i) = n$.)

Incidentally, Theorems 1.2 and 1.3 are valid for multigraphs, as are their proofs in [11], [41] and Section 5 (except that for the latter we should replace Vizing's Theorem by Shannon's $\chi' \leq 3\Delta/2$ [43] or, again, just $\chi' \leq 2\Delta - 1$).

Hypergraphs

Recall that a *hypergraph* is a collection, say $\mathcal{H}$, of subsets (called *edges*) of some finite (*vertex*) set V. It is *simple* if $|A \cap B| \leq 1$ for distinct $A, B \in \mathcal{H}$ and *k-uniform* (*k-bounded*) if each of its edges has size k (at most k). For further hypergraph background see e.g. [9].

Here we fix k (thought of as at least 3) and, unless otherwise stated, take all hypergraphs $\mathcal{H}$ to be k-uniform and simple, usually with large degrees. In the background here is the idea that, in contrast to the well-known intractibility of general hypergraph problems, such hypergraphs tend to behave quite nicely in asymptotic senses. (See e.g. [20] or, again, [9] for overviews of developments in this vein.)

We now take ξ to be the size of M drawn uniformly from the set $\mathcal{M} = \mathcal{M}(\mathcal{H})$ of matchings of $\mathcal{H}$. (For non-uniform $\mathcal{H}$ we get a slightly different set of questions if we instead let ξ be, say, the number of vertices missed by M; but this seems unlikely to make much difference, and anyway not worth worrying about for the present.)

Our best guess at an analogue of Theorem 1.6 is (recall δ is minimum degree)

Conjecture 10.3 *For a sequence $\mathcal{H}_n$ of simple, k-bounded hypergraphs with $\delta_n \to \infty$, $\{p_l(\xi_n)\}$ is asymptotically normal as $\sigma_n \to \infty$.*

In contrast to the situation for graphs, this is not true if $\delta_n \not\to \infty$. In fact, for low degree $\mathcal{H}$ with large σ, the distribution of ξ need not even be unimodal. (This strange phenomenon has something to do with expanders, but we will not go into details.)

Furthermore, $\nu_n - \mu_n \to \infty$ does not imply $\sigma_n \to \infty$ (even assuming $\delta_n \to \infty$). So it may be that the whole business should be viewed with some suspicion; nonetheless we will hazard one more:

Conjecture 10.4 *If $\mathcal{H}_n$ is simple, k-uniform and d_n-regular with $d_n \to \infty$, then $\{p_l(\xi_n)\}$ is asymptotically normal.*

(The natural analogue of Theorem 1.9 for hypergraphs as in Conjecture 10.3 is also a possibility; see [27, Conjecture 1.4].)

These conjectures, which may (again) be thought of as expressions of approximate independence in $\mathcal{M}(\mathcal{H})$, seem to require (at least) an understanding of some questions close to those concerning phase transition in the "exclusion" or "hard-core lattice gas" model of statistical physics (see e.g. [2] for a start; h.c.d.'s on matchings are special cases of this model); for instance:

It can be shown that for a simple graph G with uniform matching M and $A, B \in E(G)$

$$\Pr(A, B \in M) \sim \Pr(A \in M) \Pr(B \in M) \quad \text{as} \quad \Delta(A, B) \to \infty, \tag{91}$$

where Δ denotes distance. But (91) is not true for (simple, k-uniform) hypergraphs in general; and though it seems likely to be true when degrees are large, we do not yet see how to prove it:

Conjecture 10.5 *For simple, k-bounded $\mathcal{H}$, $A, B \in \mathcal{H}$ and M uniform from $\mathcal{M}(\mathcal{H})$,*

$$\Pr(A, B \in M) \sim \Pr(A \in M) \Pr(B \in M) \quad \text{as} \quad \delta(\mathcal{H}) \to \infty,\ \Delta(A, B) \to \infty.$$

This looks considerably more delicate—possibly a euphemism for "less true"—than (91). For instance, in a graph (or even multigraph), the event $\{A \in M\}$ can be shown to be nearly independent of the *entire* restriction of M to edges far from A, where "far from" now depends on maximum degree (results in this vein for h.c.d.'s are central to [26], [23], [24]); but this is not true for hypergraphs. Roughly, (91) is true (for graphs) because there is almost no interaction between the events $\{A \in M\}$, $\{B \in M\}$; but for hypergraphs there is, in general, substantial interaction, so that Conjecture 10.5 requires a balance of positive and negative effects (see e.g. [2, Theorem 5.4]).

Matroids

For our purposes it's convenient to regard a matroid on set E as its ideal $\mathcal{I} \subseteq 2^E$ of independent sets. (Recall that $\mathcal{I} \subseteq 2^E$ is an *ideal* if $A \subseteq B \in \mathcal{I} \Rightarrow A \in \mathcal{I}$, and is the ideal of independent sets of a matroid if for any $A \subseteq E$, any two maximal independent subsets of A have the same cardinality. For matroid background see e.g. [49]).

Notice that the notion of a hard-core distribution extends to ideals (actually to arbitrary set systems and more): the h.c.d. p associated with $\alpha : E \to \mathbf{R}^+$ is given by

$$w(I) = \prod_{e\in I} \alpha_e, \qquad p(I) = w(I)/\sum_{I'\in\mathcal{I}} w(I').$$

(We again speak of the *weighted set* E and, if $E = E(G)$, of the *weighted graph* G.)

We are interested in the behavior of ξ, the size of a set drawn uniformly, or, more generally, according to some h.c.d., from $\mathcal{I}$.

Suppose for example that $\mathcal{I} = \mathcal{F} = \mathcal{F}(G)$, the collection of forests of some multigraph G. Then (recalling that r is rank) we expect:

Conjecture 10.6 *For a sequence $\{G_n\}$ of (unweighted) multigraphs, the following are equivalent:*

(a) $\{p_k(\xi_n)\}$ *is asymptotically normal;*

(b) $\sigma_n \to \infty$;

(c) $r_n - \mu_n \to \infty$.

Here we again lose (3) and so the easy equivalence of (a) and (b). Also—in contrast with questions and results mentioned above—membership in $\mathcal{F}$ is not a "local" property (a formal definition of which seems unnecessary here). Nonetheless, the conjecture suggests that forests are actually nicer than matchings, in that we get the described equivalences even for multigraphs (equivalently integer-weighted graphs).

For arbitrary h.c.d.'s, $r-\mu$ is again fairly meaningless, but we expect that the natural analogue of γ as in (90) (that is, with ν replaced by r and M by a random forest) is the proper substitute. Moreover, here we have a seemingly reasonable replacement for the λ of Theorem 1.10, as follows. For weighted G, let F be chosen from the associated h.c.d. on $\mathcal{F} = \mathcal{F}(G)$, and let F' be chosen from the natural (induced) h.c.d. on $\mathcal{F}(G/F)$ (G/F the usual contraction). Our replacement for λ is then $\lambda'(G) = \mathbf{E}[|F'|]$. (So very roughly, λ' is large if our first random choice, F, tends to leave substantial uncertainty about the value of the 2-stage random forest $F \cup F'$. For unweighted matroids, λ' and λ are easily seen to differ by at most a factor of 2, so are equivalent for our purposes.)

Conjecture 10.7 *For a sequence $\{G_n\}$ of weighted graphs, asymptotic normality of $\{p_k(\xi_n)\}$ is equivalent to each of the conditions $\sigma_n \to \infty$, $\gamma_n \to \infty$, $\lambda'_n \to \infty$.*

Conjecture 10.6 is not true for general matroids: for, say, the uniform matroids U^n_{2n}—those with $|E| = 2n$ and $\mathcal{I}$ consisting of all sets of size at most n—we have (b), (c), but not (a). It may be, however, that the conjecture is true for matroids representable over a fixed $GF(q)$, or, more generally, excluding a fixed minor U^2_k.

On the other hand, with regard to (b) vs. (c) we have the following curious situation. Anders Johansson [19] discovered a neat proof, suggested by the proof of Lemma 5.1, that for general matroids we have $\sigma^2 \geq (r-\mu)/2$. So in particular (c) $\Rightarrow$ (b), which is the analogue of what we considered the central result of the present paper. But at this writing I don't know how to show the reverse implication (the trivial one if we had (3), and also an easy consequence of Conjecture 10.8 below), even for forests.

There is also the very intriguing possibility that all of the conditions $\sigma \to \infty$, $\gamma \to \infty$, $\lambda' \to \infty$ (with the definitions of γ, λ extended in the obvious way) are equivalent for hard-core distributions on matroids *in general.* (This would actually characterize matroids in the sense that, for a family $\mathcal{F}$ of ideals closed under minors and direct sums—see below—neither of $\sigma \to \infty$, $\gamma \to \infty$ can imply $\lambda' \to \infty$ unless $\mathcal{F}$ is a family of matroids.)

Log-concavity

It is worth recalling here that the numbers $p_k = p_k(\xi)$ are the subject of the well-known log-concavity conjectures of Mason [35], the weakest of which is

Conjecture 10.8 *For any matroid $\mathcal{I}$, the sequence $\{p_k\}_{k=0}^r$ is log-concave.*

(The strongest says that even $\{p_k/\binom{n}{k}\}$, where $n = |E|$, is log-concave.) We may think of Conjecture 10.6 as a "global" counterpart of Mason's "local" conjecture for forests. Wandering a bit further, could (something like) the following be true?

Conjecture 10.9 *For each k there exists D such that if $\mathcal{H}$ is a simple k-uniform hypergraph with $\delta(\mathcal{H}) > D$ and ξ the size of a uniform matching of $\mathcal{H}$, then the distribution of ξ is log-concave.*

Small perturbations

A natural way to begin attacking some of the problems just mentioned is to look for analogues of Lemma 4.1. In addition to being nice indicators of approximate independence, such results would—again, in the absence of (3)—be concrete first steps in the direction of asymptotic normality, as follows.

Suppose again that G is a graph and ξ the size of M drawn from some h.c.d. on $\mathcal{M}(G)$. Let $E(G) = \{A_1, \ldots, A_m\}$ and set $\omega_i = 1_{\{A_i \in M\}}$, $\xi_i = \mathbf{E}[\xi|\omega_1, \ldots, \omega_i]$. Then $\{\xi_i\}_{i=0}^m$ is a martingale. Moreover, it follows from Lemma 4.1 that the differences $\xi_i - \xi_{i-1}$ are small ($|\xi_i - \xi_{i-1}| \leq 2$, though the precise bound isn't important). In such a situation we may hope to show asymptotic normality by proving suitable concentration for the conditional variance $\sum \mathbf{E}[\xi_i^2|\omega_1, \ldots, \omega_{i-1}]$ or sum of squares $\sum \xi_i^2$; see e.g. [15, Section 3.2]. (This does not seem easy, however, and even a ((3)-free) proof of Theorem 1.6 (even, say, for regular graphs) along these lines would be quite interesting.)

For matroids there is no constant upper or lower bound on $\mathbf{E}[\xi|e] - \mathbf{E}[\xi]$, but it may be that there are bounds which are small compared to the rank of the matroid (this would still be enough for the proposed application).

Problem 10.10 *Estimate $f(r) := \sup |\mathbf{E}[\xi|e] - \mathbf{E}[\xi]|$ as $\mathcal{I}$ ranges over matroids of rank r and e over elements of $\mathcal{I}$. In particular, is it true that $f(r) = O(\log r)$?*

(One can give examples to show that both $\mathbf{E}[\xi|e] - \mathbf{E}[\xi]$ and $\mathbf{E}[\xi] - \mathbf{E}[\xi|e]$ can be $\Omega(\log r)$.)

For forests there is still no constant lower bound, but we expect $\mathbf{E}[\xi|e] - \mathbf{E}[\xi] \leq 1 - p(e)$ and, more precisely,

Conjecture 10.11 *For graphic $\mathcal{I}$ and e, f distinct elements of E, $p(f|e) \leq p(f)$.*

The analogous statement for a random *spanning tree* is an old result of Tutte [46], based on ideas from electric networks (e.g. [5]).

It may also be that there are rough versions of Lemma 4.1 for, e.g., hypergraphs as in Conjecture 10.3, but we have not really thought about concrete possibilities for this.

Ideals

Most of the above questions can be formulated at the level of general ideals. All the parameters r, μ, σ, γ, λ' make sense for the size, ξ, of I chosen according to some h.c.d. from ideal $\mathcal{I}$. For meaningful conjectures we should then restrict to $\mathcal{I}_n$ drawn from some *family* of ideals (matchings in graphs, matroid ideals...), where, for example, one might take "family" to imply closure under *minors* (meaning ideals $(\mathcal{I}\setminus B)/A := \{X \subseteq E\setminus(A\cup B) : A \cup X \in \mathcal{I}\}$, with A, B disjoint subsets of E and $A \in \mathcal{I}$) and direct sums (obvious definition). (Note, though, that hypergraphs as in Conjecture 10.3 will have low-degree hypergraphs as minors.) It would be extremely nice if there were general conditions on a family of ideals sufficient for conclusions like those conjectured above.

One (perhaps far-fetched) possibility: could there be some connection between our desired conclusions and some kind of nice behavior of the polytopes

$$K(\mathcal{I}) = \text{conv}\{1_A : A \in \mathcal{I}\} \subseteq \mathbf{R}^E \text{ ?}$$

For multigraph matchings and matroids these much-studied polytopes are indeed nice in various ways; see e.g. [14].

(Note that each of the conditions (b)-(e) of Theorem 1.10 says that some linear objective function on $K(\mathcal{I})$ has average (over vertices) close to its maximum. It was this observation that originally suggested looking at $K(\mathcal{I})$ in the present context.)

We just mention one suggestive connection: The truth of conclusions like those of Conjectures 10.6-10.8 for some ideal $\mathcal{I}$ would seem to be a reflection of something like expansion in some underlying graph Γ on $\mathcal{I}$, e.g. that with

$$I \sim_\Gamma I' \quad \text{iff} \quad |I \setminus I'| = |I' \setminus I| = 1 \text{ or } |I \Delta I'| = 1.$$

But for a *matroid* $\mathcal{I}$, this graph is precisely the 1-skeleton of $K(\mathcal{I})$, good expansion of which is a special case of the well-known "Polytope Conjecture" of M. Mihail and U. Vazirani. (See [36]. The conjecture is actually for general 0-1 polytopes, so in particular applies to any $K(\mathcal{I})$; but for general $\mathcal{I}$ the 1-skeleton of $K(\mathcal{I})$ need not have much to do with the combinatorics of $\mathcal{I}$.)

Acknowledgments

Though not as a way of spreading the blame, I would like to thank Tom Bohman, Anders Johansson, Gil Kalai, Vera Sós and Jeff Vogel for helpful conversations on the material of Section 10. I would also like to thank Laci Babai for many helpful suggestions on the exposition.

References

[1] A.D. Barbour, L. Holst and S. Janson, *Poisson Approximation*, Clarendon Press, Oxford, 1992.

[2] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, *Stochastic Processes and Appl.* **49** (1994), 179-197.

[3] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998.

[4] B. Bollobás and S.E. Eldridge, Maximal matchings in graphs with given minimal and maximal degrees, *Math. Proc. Cambridge Philos. Soc.* **79** (1976), 221-234.

[5] P.G. Doyle and J.L. Snell, *Random Walks and Electric Networks*, MAA, Washington, 1984.

[6] J. Edmonds, Maximum matching and a polyhedron with $0,1$-vertices, *J. Res. Nat. Bureau of Standards* (B) **69** (1965), 125-130.

[7] G. Ewald, *Combinatorial Convexity and Algebraic Geometry*, Springer 1996.

[8] W. Feller, *An introduction to probability theory and its applications, Vol. 2*, Wiley, New York, 1971.

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

[10] A. Godbole, personal communication.

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

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

[13] J. Goldman, J. Joichi and D. White, Rook polynomials, Möbius inversion and the umbral calculus, *J. Combinatorial Th. (A)* **21** (1976), 230-239.

[14] M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, Berlin, 1988.

[15] P. Hall and C.C. Heyde, *Martingale Limit Theory and its Application*, Academic Press, Boston, 1980.

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

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

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

[19] A. Johansson, personal communication.

[20] 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. (Distributed by AMS.)

[21] J. Kahn, On some hypergraph problems of Paul Erdős and the asymptotics of matchings, covers and colorings, pp.345–371 in *The Mathematics of Paul Erdős, v.1*, Springer, Berlin, 1997.

[22] J. Kahn, Asymptotics of hypergraph matching, covering and coloring problems, pp.1353-1362 in *Proc. Int'l. Cong. of Mathematicians, Zürich, 1994*, Birkhäuser Verlag, Basel, 1995.

[23] J. Kahn, Asymptotics of the chromatic index for multigraphs, *J. Combinatorial Th. (B)* **68** (1996), 233–254.

[24] J. Kahn, Asymptotics of the list-chromatic index for multigraphs, submitted.

[25] J. Kahn and P.M. Kayll, On the stochastic independence properties of hard-core distributions, *Combinatorica* **7** (1997), 369–391.

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

[27] J. Kahn and J.H. Kim, Random matchings in regular graphs, *Combinatorica* **18** (1998), 201-226.

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

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

[30] C.W. Lee, Some recent results on convex polytopes, pp. 3-19 in *Contemporary Mathematics*, AMS, 1990.

[31] P. Lévy, Sur une propriété de la loi de Poisson relative aux petites probabilités, *Soc. Math. de France, Comptes rendus des scéances de l'annee 1936* (1937), 29.

[32] P. Lévy, *Théorie de l'Addition des Variables Aléatoires, 2nd ed.*, Gauthiers-Villars, Paris, 1954.

[33] G.M. Louth, *Stochastic Networks: Complexity, Dependence and Routing*, thesis, Cambridge University, 1990.

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

[35] J.H. Mason, Matroids: unimodal conjectures and Motzkin's theorem, pp. 207-221 in *Combinatorics* (D.J.A. Welsh and D.R. Woodall, eds.), Inst. Math. & Appl., 1972.

[36] M. Mihail, On the expansion of combinatorial polytopes, A survey, in *Mathematical Foundations of Computer Science, Prague 1992*, Springer-Verlag Lecture Notes in Computer Science, vol. 629, pp. 37-49.

[37] A. Nijenhuis, On permanents and zeros of rook polynomials, *J. Combinatorial Th. (A)* **21** (1976), 240-244.

[38] J. Pitman, Probabilistic bounds on the coefficients of polynomials with only real zeros, *J. Combinatorial Th. (A)* **77** (1997), 279-303.

[39] Y. Rabinovich, A. Sinclair and A. Wigderson, Quadratic Dynamical Systems, *Proc. 33rd IEEE Symposium on Foundations of Computer Science* (1992), 304-313.

[40] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.

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

[42] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, Chichester, 1986.

[43] C.E. Shannon, A theorem on coloring the lines of a network, *J. Math. Phys.* **28** (1949), 148-151.

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

[45] C. Stein, *Approximate Computation of Expectations*, IMS, Hayward CA, 1986.

[46] W.T. Tutte, A problem on spanning trees, *Quart. J. Math. Oxford* **25** (1974), 253-255.

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

[48] D.G. Wagner, The partition polynomial of a finite set system, *J. Combinatorial Th.* (A) **56** (1991), 138-159.

[49] D.J.A. Welsh, *Matroid Theory*, Academic Press, London, 1976.