Intersections of Randomly Embedded Sparse Graphs are Poisson

Edward A. Bender
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112
ebender@math.ucsd.edu

E. Rodney Canfield
Department of Computer Science
The University of Georgia
Athens, GA 30602, USA
erc@cs.uga.edu

Submitted: August 3, 1999; Accepted: September 26, 1999

Abstract

Suppose that $t \geq 2$ is an integer, and randomly label t graphs with the integers $1 \ldots n$. We give sufficient conditions for the number of edges common to all t of the labelings to be asymptotically Poisson as $n \to \infty$. We show by example that our theorem is, in a sense, best possible. For G_n a sequence of graphs of bounded degree, each having at most n vertices, Tomescu [7] has shown that the number of spanning trees of K_n having k edges in common with G_n is asymptotically $e^{-2s/n}(2s/n)^k/k! \times n^{n-2}$, where $s = s(n)$ is the number of edges in G_n. As an application of our Poisson-intersection theorem, we extend this result to the case in which maximum degree is only restricted to be $o(n \log\log n/ \log n)$. We give an inversion theorem for falling moments, which we use to prove our Poisson-intersection theorem.

AMS-MOS Subject Classification (1990): 05C30; Secondary: 05A16, 05C05, 60C05

1. Introduction and Statement of Graphical Results

This paper considers random embeddings of an m-vertex graph G into the complete graph K_n where $m \le n$. With no loss, assume the vertices of G are $\{1, 2, \ldots m\}$. The number of injections of an m-set into an n-set is $(n)_m$, the falling factorial

$$(n)_m \quad = \quad n(n-1)\cdots(n-m+1).$$

By a *random embedding* of G into K_n, we mean that one of the above injections is chosen from the uniform distribution.

Tomescu [7] showed that the number of edges a randomly embedded graph G_n has in common with a random spanning tree of K_n is asymptotically Poisson when the degree of the graph is bounded. (The result had been conjectured in [6], and proven there for a special case.) This can be interpreted in terms of random embeddings of pairs of graphs in K_n. Theorem 1, discusses random embeddings of t-tuples of graphs. In Theorem 2, we use this to extend Tomescu's result from graphs of bounded degree to those whose degrees may grow as fast as $o(n \log\log n/\log n)$.

Theorem 1. *Let $t \ge 2$ be an integer. Suppose that for each i, $1 \le i \le t$, we have a sequence $G_n(i)$, of graphs, each having at most n vertices and at least one edge. Let $s_n(i)$ and $\Delta_n(i)$ be the number of edges and the maximum degree, respectively, for $G_n(i)$. Let Y_n equal the number of edges common to t randomly chosen embeddings of the $G_n(i)$ into K_n. Let*

$$\lambda_n \quad = \quad \frac{\prod_{i=1}^{t} s_n(i)}{\binom{n}{2}^{t-1}} \tag{1}$$

and

$$\rho_n \quad = \quad \prod_{i=1}^{t} \frac{(\Delta_n(i))^2}{s_n(i)}. \tag{2}$$

If $\min(\lambda_n, \rho_n) \to 0$, *then*

$$\Big(\mathrm{Prob}\big(Y_n = k\big) \;-\; e^{-\lambda_n}\,\lambda_n^k\,/\,k!\Big) \quad \to \quad 0 \qquad \textit{for each fixed } k. \tag{3}$$

Theorem 2. *Let G_n be a sequence of graphs, each having at most n vertices. Let s_n and Δ_n be the number of edges and the maximum degree, respectively, for G_n. Let $T(G_n; n, k)$ be the number of spanning trees of K_n having k edges in common with G_n. If*

$$\Delta_n \quad = \quad o(n \log\log n/\log n), \tag{4}$$

then

$$\Big(T(G_n; n, k)/n^{n-2} \;-\; e^{-2s_n/n}\,(2s_n/n)^k\,/\,k!\Big) \quad \to \quad 0 \qquad \textit{for each fixed } k. \tag{5}$$

To what extent are the constraints on the sequences Δ_n needed in Theorems 1 and 2? Some condition is needed in Theorem 2 since $T(G_n; n, 0) = 0$ for the n-vertex star; however, we do not know if (4) is best possible. For Theorem 1 we have the following result.

Theorem 3. *We cannot replace the condition* $\min(\lambda_n, \rho_n) \to 0$ *in Theorem 1 with* $\min(\lambda_n, \rho_n) = \mathcal{O}(1)$*:*

(a) If $G_n(1)$ *is an* n*-cycle and* $G_n(2)$ *is an* n*-vertex star, then* $\mathrm{Prob}(Y_n = 2) = 1$.

(b) If $G_n(1)$ *and* $G_n(2)$ *are both caterpillars with* $b = \lfloor n^{1/2} \rfloor$ *nonleaf vertices, each of degree* b*, then*

$$\sum_{k=0}^{\infty} \lim_{n\to\infty} \mathrm{Prob}(Y_n = k) z^k \quad = \quad e^{z-1} \exp\bigl\{(e^{z-1} - 1)\bigr\}.$$

The two examples are extreme: in (a) the ratios $\Delta_n(i)^2/s_n(i)$ differ greatly; in (b) they are equal.

2. A Theorem on Convergence to Poisson

The proof of Theorem 1, in the next section, requires an inversion theorem for falling moments. Inverting estimates for moments into estimates for the underlying probability distribution is a classical technique. In [2] this is done when the moment generating function has positive radius of convergence. An inversion theorem more useful in some circumstances is stated in [3, p. 75]: If there is a λ such that for every k we have $\mathbf{E}\bigl((Y_n)_k\bigr) \to \lambda^k$, then also we have $\mathrm{Prob}(Y_n = k) \to e^{-\lambda}\lambda^k/k!$. (Strictly, the theorem is stated in terms of factorial cumulants (see page 50), but is equivalent to what is stated here.) No proof or reference is given, but this assertion is a corollary of Theorem 4 below. Similar inversion theorems are found in [1, p. 491] and [5, p. 22], phrased in inclusion-exclusion terms. Of the above, only [1] does not require that $\mathbf{E}(Y_n) \to \lambda$. The following theorem is similar to that in [1], but differs sufficiently that it is inappropriate to refer to that paper for a proof.

Theorem 4. *Let* $Y_1, Y_2, \ldots$ *be a sequence of nonnegative integer valued random variables, each of which has falling moments of all orders. Let* λ_n *be the expected value* $\mathbf{E}(Y_n)$ *of the random variable* Y_n*. For each* $\epsilon > 0$ *define the sets* $\mathcal{A}_\epsilon, \mathcal{B}_\epsilon \subseteq \{1, 2, \ldots\}$ *by*

$$\begin{aligned} \mathcal{A}_\epsilon &= \{n : \lambda_n > 1/\epsilon\} \\ \mathcal{B}_\epsilon &= \{n : 1/\epsilon > \lambda_n > \epsilon\}. \end{aligned}$$

Suppose that for each real $\epsilon > 0$ *we have*

$$\mathbf{E}\bigl((Y_n)_2\bigr) \sim \lambda_n^2 \quad \text{as } n \to \infty \text{ through } \mathcal{A}_\epsilon, \tag{6}$$

and that for each real $\epsilon > 0$ *and integer* $k > 0$ *we have*

$$\mathbf{E}\bigl((Y_n)_k\bigr) \sim \lambda_n^k \quad \text{as } n \to \infty \text{ through } \mathcal{B}_\epsilon. \tag{7}$$

Then it follows that

$$\Bigl(\mathrm{Prob}\bigl(Y_n = k\bigr) - e^{-\lambda_n}\,\lambda_n^k\,/\,k!\Bigr) \to 0 \quad \text{for each } k.$$

To prove this, we require Bonferroni's inequalities:

Theorem 5. *Let Y be a random variable taking on nonnegative integer values. Suppose $\mathbf{E}((Y)_k)$ exists for $k \le K$. Then, for $0 \le J \le K-k$*

$$\sum_{j=0}^{J} \frac{(-1)^j\, \mathbf{E}((Y)_{k+j})}{k!\, j!} \tag{8}$$

is an over-estimate of $\mathrm{Prob}(Y=k)$ *when J is even and an under-estimate when J is odd. Furthermore, in absolute value the last term in the sum is a bound on the difference between the sum over $0 \le j < J$ and* $\mathrm{Prob}(Y=k)$.

Proof of Theorem 5. The last sentence in the theorem follows immediately from the over- and under-estimate claim concerning (8). We now prove the over- and under-estimate claim. Let $y_k(N) = \sum_{n=k}^{N} \mathrm{Prob}(Y=n)\,(n)_k$. Note that $\mathbf{E}((Y)_k) = \lim_{N\to\infty} y_k(N)$. We have

$$\begin{aligned}\sum_{j=0}^{J} \frac{(-1)^j y_{k+j}(N)}{k!\, j!} &= \sum_{j\le J}\sum_{n=k}^{N} \frac{(-1)^j\, \mathrm{Prob}(Y=n)\,(n)_{k+j}}{k!\, j!} \\ &= \sum_{n=k}^{N} \frac{\mathrm{Prob}(Y=n)\,(n)_k}{k!} \left(\sum_{j\le J} \frac{(-1)^j (n-k)_j}{j!}\right).\end{aligned}$$

The parenthesized sum is $\sum_{j\le J}\binom{n-k}{j} = (-1)^J\binom{n-k-1}{J}$ when $n>k$, and it equals 1 when $n=k$. Letting $N\to\infty$ proves the theorem. ∎

Proof of Theorem 4. Let an integer k and an $\epsilon>0$ be given. We must exhibit N such that

$$n \ge N \quad\Rightarrow\quad \Big(\,\big|\,\mathrm{Prob}(Y_n=k) - e^{-\lambda_n}\lambda_n^k/k!\,\big| \;<\; \epsilon\Big). \tag{9}$$

We separate the proof into three cases, and exhibit four constants N_1, N_2, N_3, and L such that each of the conditions

$$\begin{aligned} &n \ge N_1 \;\text{ and }\; \lambda_n < \epsilon \\ &n \ge N_2 \;\text{ and }\; \lambda_n > L \\ &n \ge N_3 \;\text{ and }\; \epsilon \le \lambda_n \le L \end{aligned}$$

implies the desired conclusion appearing on the right side of (9).

When $k\ge 1$ and $\lambda>0$, we easily have

$$1-\lambda \;\le\; e^{-\lambda} \;\le\; 1 \;\text{ and }\; 0 \;\le\; e^{-\lambda}\lambda^k/k! \;\le\; \lambda e^{-\lambda}(\lambda^{k-1}/(k-1)!) \;<\; \lambda.$$

For any random variable Y taking non negative integer values, we have, for $k\ge 1$,

$$0 \;\le\; \mathrm{Prob}(Y=k) \;\le\; \mathrm{Prob}(Y\ge 1) \;\le\; \mathbf{E}(Y),$$

whence also

$$1 - \mathbf{E}(Y) \;\le\; \mathrm{Prob}(Y = 0) \;\le\; 1.$$

Combining these with the previous estimates, we obtain the first part of our proof simply by taking $N_1 = 1$.

For the second part of the proof, we employ Chebyshev's inequality. Recall that k is fixed. Choose L so large that

$$L \ge 2k, \quad 1/L < \epsilon/2, \qquad \text{and} \qquad \lambda > L \;\Rightarrow\; e^{-\lambda}\lambda^k/k! < \epsilon.$$

Let σ_n^2 be the variance of Y_n. Chebyshev's inequality gives

$$\mathrm{Prob}(Y_n = k) \;\le\; \Big(\frac{\sigma_n}{\lambda_n - k}\Big)^2,$$

and the latter is no greater than $4(\sigma_n/\lambda_n)^2$ since $L \ge 2k$. However, by hypothesis, if n becomes infinite through values such that $\lambda_n \ge L$, we have

$$\begin{aligned} \sigma_n^2 &= \mathbf{E}\big((Y_n)_2\big) + \lambda_n - \lambda_n^2 \\ &= (1 + o(1))\lambda_n^2 + \lambda_n - \lambda_n^2 \\ &= \lambda_n + o(\lambda_n^2), \end{aligned}$$

whence

$$(\sigma_n/\lambda_n)^2 \;=\; (\lambda_n)^{-1} + o(1).$$

The first term on the right is less than $\epsilon/2$ by our choice of L; hence, choosing N_2 so large that the $o(1)$ term is less than $\epsilon/2$ for $n \ge N_2$ completes the second part of the proof.

For the third and final part of the proof, we use Theorem 5. Choose J sufficiently large that

$$\frac{(\epsilon^{-1})^{k+J}}{k!\,J!} \;<\; \epsilon/3 \quad \text{and} \quad J \;>\; L.$$

Note that

$$\Big| \sum_{0 \le j < J} (-1)^j \lambda^j/j! \;-\; e^{-\lambda} \Big| \;\le\; \lambda^J/J! \tag{10}$$

since, for $J > \lambda$, the absolute values of the terms with $j \ge J$ are decreasing and so the error is at most the first neglected term. There are three errors, E_1, E_2, E_3, to bound:

$$\begin{aligned} \mathrm{Prob}(Y_n = k) &= \sum_{0 \le j < J} \frac{(-1)^j \mathbf{E}\big((Y_n)_{k+j}\big)}{k!\,j!} + E_1 \\ \sum_{0 \le j < J} \frac{(-1)^j \mathbf{E}\big((Y_n)_{k+j}\big)}{k!\,j!} &= \sum_{0 \le j < J} \frac{(-1)^j \lambda_n^{j+k}}{k!\,j!} + E_2 \\ \sum_{0 \le j < J} \frac{(-1)^j \lambda_n^{j+k}}{k!\,j!} &= \frac{\lambda_n^k}{k!} e^{-\lambda_n} + E_3. \end{aligned}$$

- By Theorem 5, E_1 is smaller than $\mathbf{E}\big((Y_n)_{k+J}\big)/k!\,J!$, which by assumption is $(1+o(1))\lambda_n^{J+k}/k!\,J!$. By choice of J the latter is less than $\epsilon/3$ for n sufficiently large.
- E_2 is bounded in absolute value by $o(1)\sum_{0\le j<J} 1/k!\,j! < o(1)e/k!$, the $o(1)$ term being the maximum of the finitely many differences $|\big((Y_n)_{k+j}\big) - \lambda_n^{k+j}|$. For n sufficiently large, $o(1)e/k!$ is smaller than $\epsilon/3$.
- Finally, by (10), E_3 is smaller in absolute value than $\lambda_n^{k+J}/k!\,J!$, which as noted already is less than $\epsilon/3$ for n sufficiently large.

This concludes the proof. ∎

3. Proof of Theorem 1

Using Theorem 4, we now prove Theorem 1. At times we drop subscripts and superscripts and refer to a graph $G \subseteq K_n$ with s edges and maximum degree Δ. Throughout the proof, we speak of the probability of various events, and evaluate the expected value of some random variables. The underlying probability space for all this is the set of all t-tuples of embeddings of the graphs $G(i)$ into K_n with the uniform distribution.

If $\omega = \omega_1, \ldots, \omega_t$ is such an embedding, then $(Y_n(\omega))_k$ is the number of ways to choose a sequence $\mathbf{e}$ of k distinct edges all of which lie in every embedding. Let $\chi(S)$ be 1 if the statement S is true and 0 otherwise. With a sum on ω running over all embeddings and a sum on $\mathbf{e}$ running over all k-tuples of distinct edges,

$$\begin{aligned}
\mathbf{E}((Y_n)_k) &= \sum_{\omega} \mathrm{Prob}(\omega) \sum_{\mathbf{e}} \chi(\mathbf{e} \text{ is in every } \omega) \\
&= \sum_{\mathbf{e}} \sum_{\omega} \prod_{i=1}^{t} \Big(\mathrm{Prob}(\omega_i)\, \chi(\mathbf{e} \text{ is in every } \omega)\Big) \\
&= \sum_{\mathbf{e}} f_n(\mathbf{e}), \quad \text{where } f_n(\mathbf{e}) = \prod_{i=1}^{t} p_n(G_n(i) \supset \mathbf{e}), \qquad (11)
\end{aligned}$$

and $p_n(G \supset \mathbf{e})$ is the probability that a random embedding of G in K_n contains $\mathbf{e}$.

Partition the k-tuples of distinct edges in K_n into two classes, $\mathcal{I}$ and $\mathcal{D}$, where $\mathcal{I}$ contains all k-tuples of *independent* edges and $\mathcal{D}$ contains all other k-tuples (the *dependent* sets). Thus $\sum_{\mathbf{e}}$ in (11) can be partitioned into sums over $\mathcal{I}$ and $\mathcal{D}$.

Here is a way to compute $p_n(G \supset \mathbf{e})$. Imagine G as a subgraph of K_n. Now choose edges of K_n to be relabeled as $\mathbf{e}$, preserving whatever incidences are required among the ends of the e_i by the names of their vertices. The probability that these k chosen edges lie in G is $p_n(G \supset \mathbf{e})$.

We now consider $\sum_{\mathbf{e}\in\mathcal{I}} f_n(\mathbf{e})$, using the method in the previous paragraph to estimate $p_n(G \supset \mathbf{e})$. The edges chosen to be $\mathbf{e}$ can be any independent set in K_n, of which there are $(n)_{2k}$ if the edges are directed and so $(n)_{2k}/2^k$ if the edges are not directed. Hence $|\mathcal{I}| = (n)_{2k}/2^k$ and

$$p_n(G \supset \mathbf{e}) = \frac{2^k I(G,k)}{(n)_{2k}}, \tag{12}$$

where $I(G,k)$ is the number of k-long sequences of independent edges in G. Thus

$$\sum_{\mathbf{e}\in\mathcal{I}} f_n(e) \;=\; \left(\frac{2^k}{(n)_{2k}}\right)^{t-1} \sum_{\mathbf{e}\in\mathcal{I}} \prod_{i=1}^{t} I(G_n(i),k). \tag{13}$$

When $k=1$, we have $\mathcal{D}=\emptyset$ and $I(G_n(i),1)=s_n(i)$. Thus, from (11) and (13),

$$\mathbf{E}(Y_n) = \frac{1}{\binom{n}{2}} \prod_{i=1}^{t} \frac{s_n(i)}{\binom{n}{2}}. \tag{14}$$

This shows that the λ_n of Theorem 1 is $\mathbf{E}(Y_n)$. By the hypotheses of Theorem 4, we may restrict our attention to n with $\lambda_n > \epsilon$, which we do from now on. By hypothesis $\min(\rho_n, \lambda_n) \to 0$, and so

$$\rho_n \to 0 \text{ as } n \to \infty \text{ through } \mathcal{A}_\epsilon \cup \mathcal{B}_\epsilon. \tag{15}$$

We detour briefly to prove a bound on the growth of the Δ's that is needed later: For each i,

$$\Delta_n(i)/s_n(i) \to 0 \quad \text{as } n \to \infty \text{ through } \mathcal{A}_\epsilon \cup \mathcal{B}_\epsilon. \tag{16}$$

For all i, $n\Delta_n(i) \geq 2s_n(i)$ by a simple counting argument. Hence

$$\begin{aligned} \frac{\rho_n}{\lambda_n} &= \frac{\prod_{i=1}^{t}(\Delta_n(i)^2/s_n(i))}{\binom{n}{2}\prod_{i=1}^{t}\left(s_n(i)/\binom{n}{2}\right)} \\ &\sim \left(\frac{\Delta_n(j)}{s_n(j)}\right)^2 \prod_{i\neq j} \frac{\left(n\Delta_n(i)/s_n(i)\right)^2}{2} \;\geq\; 2^{1-t}\left(\frac{\Delta_n(j)}{s_n(j)}\right)^2. \end{aligned}$$

By (15) and $\lambda_n > \epsilon$, (16) follows.

We have

$$s^k \geq I(G,k) \geq s(s-\Delta)\cdots(s-(k-1)\Delta).$$

Since

$$\begin{aligned} \frac{s^k}{s(s-\Delta)\cdots(s-(k-1)\Delta)} &< \left(\frac{s}{s-k\Delta}\right)^k \\ &= \left(1+\frac{k\Delta}{s-k\Delta}\right)^k \;<\; \exp\left(\frac{k^2\Delta}{s-k\Delta}\right), \end{aligned}$$

it follows from (16) that $I(G_n(i), k) \sim s_n(i)^k$ for each fixed k. Hence $f_n(\mathbf{e}) \sim (2\lambda/n^2)^k$ when $\mathbf{e} \in \mathcal{I}$. Since $|\mathcal{I}| = (n)_{2k}/2^k \sim n^{2k}/2^k$, it follows from (11) and (12) that

$$\sum_{\mathbf{e}\in\mathcal{I}} f_n(\mathbf{e}) \;\sim\; (n^{2k}/2^k)\prod_{i=1}^{t}(s_n(i)^k 2^k/n^{2k}) \;\sim\; \lambda_n^k. \tag{17}$$

We will show that

$$\sum_{\mathbf{e}\in\mathcal{D}} f_n(\mathbf{e}) = o(\lambda_n^k). \tag{18}$$

When (17) and (18) are combined with $\lambda_n = \mathbf{E}(Y_n)$, we obtain (7) with $\mathcal{B}_\epsilon$ replaced by $\mathcal{A}_\epsilon \cup \mathcal{B}_\epsilon$ and hence (6) follows as well. Thus, proving (18) will complete the proof of Theorem 1.

Suppose that the k edges in $\mathbf{e}$ form a graph H with v vertices and c components. Since $\mathbf{e} \in \mathcal{D}$,

$$c < v/2 < k. \tag{19}$$

Fix a spanning forest F of H. The edges of $\mathbf{e}$ are relabeled in the following order:

1. One edge in each tree in the spanning forest. The probability that each such edge lies in G, conditioned on edges already relabeled is bounded above by $2s/(n-2k)^2$.
2. Additional edges that grow each tree in the spanning forest in a connected fashion. The probability that each such edge lies in G, conditioned on edges already relabeled is bounded above by $\Delta/(n-2k)$. To see this, note that one vertex on each such edge has already been embedded.
3. The remaining edges of $\mathbf{e}$. Here we use the trivial bound of 1 for the conditional probability.

Since the number of edges in a spanning forest on v vertices and c components is $v-c$, there are $v-2c$ edges in Step 2 and so

$$p_n(G \supset \mathbf{e}) \le \left(\frac{2s}{(n-2k)^2}\right)^c \left(\frac{\Delta}{n-2k}\right)^{v-2c}.$$

Constructing possible $\mathbf{e}$'s with the given values of k, v, and c in a similar manner, we see that there are at most

$$k!\,(n^2/2)^c\,n^{v-2c} \;=\; (2^c\,k!)\,n^v.$$

Hence the contribution of such $\mathbf{e}$ to the sum over $\mathcal{D}$ is at most

$$\begin{aligned}
&(2^c\,k!)\,n^v\prod_{i=1}^{t}\left(\frac{2s_n(i)}{(n-2k)^2}\right)^c\left(\frac{\Delta_n(i)}{n-2k}\right)^{v-2c}\\
&\quad= \mathcal{O}(1)\,n^v\prod_{i=1}^{t}\left(\frac{2s_n(i)}{n^2}\right)^c\left(\frac{\Delta_n(i)}{n}\right)^{v-2c}\\
&\quad= \mathcal{O}(1)\left\{n^2\prod_{i=1}^{t}\frac{s_n(i)}{n^2}\right\}^{v/2}\left(\prod_{i=1}^{t}\frac{\Delta_n(i)^2}{s_n(i)}\right)^{v/2-c}.
\end{aligned}$$

It follows from (15) and (19) that all $\mathbf{e} \in \mathcal{D}$ with a given set of values for v and c contribute $\mathcal{O}(\lambda_n^{v/2})$ to $\mathbf{E}((Y_n)_k)$. Since the number of choices for v and c is bounded, $\lambda_n > \epsilon$ and $v/2 < k$ by (19), we are done.

4. Proof of Theorem 2

Apply Theorem 1 with $t = 2$, $G_n(1) = G_n$, and $G_n(2) = T_n$, a spanning tree of K_n. In the next paragraph we show that, if $k \log\log n/\log n \to \infty$, then almost all spanning trees of K_n have $\Delta < k$. Thus $\rho_n \to 0$ for almost all spanning trees T_n provided

$$\frac{\Delta_n^2(\log n/\log\log n)^2}{ns_n} = o(1).$$

Averaging Theorem 1 over almost all T_n, eliminating those of high degree, proves Theorem 2.

The maximum degree bound follows from [4], but we include a simple proof here for completeness. Consider the Prüfer sequence for a tree. If the maximum degree is k, no number appears more than k times. An upper bound on sequences with at least $t = k+1$ appearances of some number is obtained by choosing (i) a number from $\{1, \ldots, n\}$ to appear at least t times, (ii) t locations for it in the Prüfer sequence, and (iii) the remaining $n - t - 2$ sequence elements. Hence we have the upper bound

$$n\binom{n-2}{t}n^{n-t-2} \;<\; n^{n-2}(n/t!) \;<\; n^{n-2}(n/k!).$$

Since there are n^{n-2} trees, the maximum degree is almost surely less than k if $n/k! = \mathcal{O}(1)$. The claim in the previous paragraph follows from Stirling's formula.

5. Proof of Theorem 3.

For part (a), suppose that the star has been embedded in K_n and let v be the vertex that is connected to the other $n-1$ vertices. When the n-cycle is embedded, one vertex will map to v. The two edges of the cycle that contain v as an end point also lie in the star and no other edges do.

Part (b) involves somewhat more calculation. Suppose the first caterpillar, $G_n(1)$, has been embedded and let $\mathcal{V} = \{v_1, \ldots, v_b\}$ be the vertices of degree b. There are two sources of common edges: First, when the second caterpillar has vertices of degree b in $\mathcal{V}$. Second, when vertices of degree 1 in $G_n(2)$ lie in $\mathcal{V}$. In our computations, we will ignore some dependencies that become insignificant as $n \to \infty$.

We consider the first case. The number of vertices in $G_n(2)$ of degree b that lie in $\mathcal{V}$ is asymptotically Poisson and its expected value is

$$b \times \text{Prob}(v_1 \text{ has degree } b \text{ in } G_n(2)) \;=\; b \times (b/n) \;\sim\; 1.$$

If v_k has degree b in $G_n(2)$, the number of common edges between $G_n(1)$ and $G_n(2)$ that share v_k is Poisson and its expected value is

$$\sum_{v\neq v_k}\prod_{i=1}^{2} p_n\Big(G_n(i)\supset\{v_k,v\}\mid \deg(v_k)=b\Big) \;=\; (n-1)\times(b/(n-1))^2 \;\sim\; 1.$$

Hence the generating function for the number of such vertices in common is the composition of two Poisson distributions of mean 1.

We now consider the second case. Since nearly all vertices have degree 1, the number of degree 1 vertices of $G_n(2)$ that lie in $\mathcal{V}$ is asymptotic to $|\mathcal{V}|=b$. For each such vertex v, the probability that its edge in $G_n(2)$ is also in $G_n(1)$ is asymptotic to b/n since v has degree b in $G_n(1)$. Hence the number of such common edges is asymptotically Poisson with mean $b\times(b/n)\sim 1$.

Combining the results of the two previous paragraphs, we obtain Theorem 3(b).

References

1. E. A. Bender, Asymptotic methods in enumeration, *Siam Review* **16** (1974) 485–515.
2. J. H. Curtiss, A note on the theory of moment generating functions, *Annals of Mathematical Statistics* **13** (1942) 430–433.
3. F. N. David and D. E. Barton *Combinatorial Chance*, Griffin, London, 1962.
4. A. Meir and J. W. Moon, A note on trees with concentrated maximum degrees, *Utilitas Mathematica* **42** (1992) 61–64.
5. Joel Spencer, *Ten Lectures on the Probabilistic Method*, 2nd ed., SIAM, Philadelphia, 1994.
6. I. Tomescu, On the number of trees having k edges in common with a caterpillar of moderate degree, *Ann. Discrete Math.* **28** (1985) 305–310.
7. I. Tomescu, On the number of trees having k edges in common with a graph of bounded degrees, *Discrete Math.* **169** (1997) 283–286.