On a sparse random graph with minimum degree three: Likely Posa's sets are large.

Alan Frieze* and Boris Pittel†

Carnegie Mellon University and Ohio State University

af1p@andrew.cmu.edu and bgp@math.ohio-state.edu

August 6, 2011

Mathematics Subject Classifications: 05C80, 05C30, 34E05, 60C05

Keywords: random, sparse graphs, degrees, longest path, Posa sets

Abstract

We consider the likely size of the endpoint sets produced by Posa rotations, when applied to a longest path in a random graph with cn, $c \geq 2.7$ edges that is conditioned to have minimum degree at least three.

1 Introduction

In the pioneering paper [10] Erdős and Rényi asked how large m, the number of edges, should be for the uniformly random graph on n vertices ($G(n, m)$) with high probability (whp) to have a Hamilton cycle. The problem was vigorously attacked by the various authors, see references in Bollobás [6]; in particular, Komlós and Szemerédi [17] showed that $m = n^{1+\varepsilon}$ suffices. A critical breakthrough was achieved by Pósa [22]; he showed that $m = cn \ln n$, $c > 30$, is enough. Qualitatively this is the best possible, since $m = \Theta(n \ln n)$ edges

*Frieze's research supported in part by NSF

†Pittel's research supported in part by NSF Grant DMS-0805996

are needed for $G(n,m)$ to be connected whp, [10]. Progressively stronger extensions of Pósa's result for $G(n,m)$ were achieved by Korshunov[18], Komlós and Szemerédi [19], Ajtai, Komlós and Szemerédi [1], Bollobás [7], Bollobás, Fenner and Frieze [9], and Bollobás and Frieze [3]. The proofs frequently used a deceptively simple but surprisingly potent "Pósa's Lemma" from [22].

Here is the Lemma. Given a graph G and a vertex x_0, let $P = x_0x_1\dots x_h$ be a longest path from x_0. If $(x_h, x_i) \in E(G)$ for some $i < h-1$, then $P' = x_0\dots x_ix_hx_{h-1}\dots x_{i+1}$ is also a path in G from x_0, of the same edge length h. In words, P' is obtained for P by rotation via the edge (x_i, x_h). Let S consist of x_h and the set of endpoints of all paths obtainable from P through any number of rotations. Let T be the set of all outside neighbors of S in G, $T = N(S)$ in short. $T \subset \{x_0, \dots, x_h\}$ as P is a longest path. Pósa's Lemma states that there are no edges between S and the vertices of $P \setminus (S \cup T)$.

Pósa's Lemma implies that $|T| < 2|S|$. Now if a graph G is not too sparse, one may expect that the not-too-large vertex sets A are sufficiently expanding, so that $|N(A)| \geq 2|A|$. If that is the case, then it follows from Pósa's Lemma that $|S|$ has to be "large", $|S| \geq s(G)$. Bollobás's next crucial observation was that, if a path P cannot be extended via a sequence of rotations at either of its ends and there is no cycle of length $h(P)+1$, then there are at least $\binom{s(G)}{2}$ "non-edges" with a property: adding any such non-edge to $E(G)$ creates a cycle out of a properly rotated P. Using these fundamental properties of Pósa's sets, and also de la Vega's theorem on whp existence of long paths in $G(n,p)$, $(np > 4\ln 2)$, [12], Bollobás [7], [6] found a surprisingly direct proof of Korshunov's, Komlós-Szemerédi's result on a sharp threshold value of m and p for whp-Hamiltonicity of $G(n,m)$ and $G(n,p)$.

Specifically, he proved that, for p in question, whp $s(G(n,p)) = \Theta(n)$, that is the likely number of those beneficial non-edges is quadratic in n. Using this, he identified a sequence $G(n,p_0) \subset G(n,p_1) \subset G(n,p_k) \subset G(n,p)$, $k = k(n)$, such that de la Vega's result applies to $G(n,p_0)$ with room to spare, and for each j, with the conditional probability $1 - O(n^{-2})$, the length of the longest path in $G(n,p_{j+1})$ strictly exceeds that in $G(n,p_j)$, if the latter is non-Hamiltonian. That $G(n,p_k)$ is whp Hamiltonian was then immediate.

Later we used a broadly similar argument to show that a random graph with minimum degree at least 2, $G_2(n,m)$ in short, whp has a perfect matching iff $G_2(n,m)$ has no isolated odd cycles, Frieze and Pittel [15]. A counterpart of Pósa's Lemma in our case was a lemma inspired by Gallai–Edmonds

Structure Theorem and Edmonds' Matching algorithm. This lemma allowed us to prove that, in absence of a perfect matching, with high conditional probability there exist $\Theta(n^2)$ non-edges each of which would increase the maximum matching number. And the place of de la Vega's algorithm was taken up by Karp-Sipser Matching Greedy [16], analyzed in detail in our earlier paper [2].

As the title shows, the core of Bollobás' paper [7] was a proof that the uniformly random d-regular graph on $[n]$, $G_d(n)$, is whp Hamiltonian, if $d > 10^7$. Fenner and Frieze [11] independently proved that $d > 796$ suffices and then Frieze [13] came up with an algorithmic proof of a better bound $d > 85$. It was commonly believed that $d \geq 3$ suffices, and indeed Robinson and Wormald [23] settled this conjecture affirmatively. Their non-algorithmic proof was based on a refined version of the second order moment prompted by their discovery that, for $d \geq 3$, $\mathrm{E}[X_n^2] = O(\mathrm{E}^2[X_n])$, X_n being the number of Hamilton cycles. (The random graphs $G(n,p), G(n,m)$, with p, m in question, lack this remarkable property.) Very recently, Bohman and Frieze [3] proved that another well-known random graph, $G_{d-\mathrm{out}}(n)$ is whp Hamiltonian, if $d \geq 3$, in which case the average vertex degree is asymptotic to 6.

Now, consider the whp Hamiltonicity of $G_3(n,m)$, the random graph on $[n]$ with m edges and minimum degree 3, at least. Of course, $m \geq 3n/2$, and we had better assume $m > 3n/2$, since equality implies that $G_3(n,m) = G_3(n)$. From a more general result in Bollobás, Cooper, Fenner and Frieze [8] it follows that whp $G_3(n,m)$ is Hamiltonian if $m > 128n$. Our ultimate goal is to push this bound down, close to the best possible $m > 1.5n$, and to construct an algorithm that finds a Hamilton cycle in $O_p\big(n^{1+o(1)}\big)$ running time for m/n in the arising range.

Theorem 1.1. *Introduce* $f_k(x) = \sum_{j\geq k} x^j/j!$, *a tail of the series for* e^x, $x^* = 4.789771\ldots$, *a unique positive root of*

$$\frac{x^3 f_1(x)}{f_2(x)^2} = 1,$$

and

$$a^* = \frac{x^* f_2(x^*)}{2 f_3(x^*)} \approx 2.6616.$$

If $m \geq (a^* + \varepsilon)n$, $\varepsilon > 0$, *i. e. the average vertex degree exceeds* 5.32, *then* *whp for each pair of Pósa's sets* (S,T),

$$|S|+|T|\geq n^{1-\delta_n},\quad \delta_n=(\ln\ln n)^{-1/2}.$$

In words, the likely Pósa's sets are, at least, almost linear in size.

Remark 1.1. *If the vertex degree range is an arbitrary* $D\subseteq[3,\infty)$, $3\in D$, *then the above assertion continues to hold if we replace* $f_3(x)$ *with*

$$f_D(x)=\sum_{j\in D}\frac{x^j}{j!},$$

and $f_2(x)$, $f_1(x)$ *with* $f_D'(x)$ *and* $f_D''(x)$ *respectively. For instance, if* $D=\{3,4\}$, *then the likely Pósa's sets are almost linear in size if* $m\geq 17/9\approx 1.9$.

The proof of this claim takes up the rest of the paper. It is quite technical, apparently due to exceeding sparseness of $G_3(n,m)$ for m/n that close to the best 1.5. We firmly believe that, in a complete analogy with Bollobás' proof of Hamiltonicity of $G(n,p)$ and $G(n,m)$, and our result on the existence of a perfect matching in $G_3(n,m)$, the random graph $G_3(n,m)$ is whp Hamiltonian if $m\geq(a^*+\varepsilon)n$. In fact, we have already found a necessary modification of Karp-Sipser Matching Greedy. This algorithm builds a 2-matching, viz. a spanning subgraph of $G_3(n,m)$ with maximum degree at most two. A companion paper [14] proves that whp this algorithm produces a 2-matching M with $O(\log n)$ components provided $c\geq 15$. It then shows how to transform M efficiently into a Hamilton cycle.

2 Posa's sets in a graph with minimum degree 3 at least

Consider a graph G. Suppose that P is a longest path from a fixed vertex v_0. Posa's set is a pair (S,T). S is the set of endpoints of paths obtainable from P through rotations via edges connecting an endpoint and another vertex of a current path; $T=N(S)$, the set of neighbors of S, outside of S. The vertices of T are all in P, since otherwise P would not be a longest path from v_0. Posa established a key property of (S,T), namely

$$|T|<2|S|. \tag{2.1}$$

Given a graph $G=(V,E)$ and $A\subseteq V$, we use $G(A)$ to denote the subgraph of G induced by A, and $e(A)$ to denote the number of edges in $G(A)$.

Lemma 2.1. *Suppose that the minimum degree of G is 3 at least. Then*

$$e(S \cup T) > |S \cup T|, \tag{2.2}$$

i. e. the edge density of $G(S \cup T)$ strictly exceeds 1.

Proof of Lemma 2.1 Let Q be any path obtained from P by rotations. Posa observed that

$$\text{every } t \in T \text{ has an } S\text{-neighbor on } Q. \tag{2.3}$$

Introduce T_1, a set of all vertices $t \in T$ such that t has only one neighbor $s \in S$. Posa's observation implies that, for every such pair (s, t), t and s are neighbors on every path Q. In particular, when s is an endpoint, t is next to s. It follows then that for any other vertex $t' \in T_1$ with a single neighbor $s' \in S$ we have $s' \neq s$. Therefore $|T_1| \leq |S|$.

Let $D(S)$ denote the total degree of vertices in S, and let $D_S(T)(\leq D(S))$ denote the total number of neighbors of T in S. Since each $t \in T \setminus T_1$ has at least two neighbors in S, and $|T_1| \leq |S|$, we have

$$D_S(T) \geq |T_1| + 2(|T| - |T_1|) \geq 2|T| - |S|. \tag{2.4}$$

Hence, as each $s \in S$ has degree 3 at least,

$$D(S) + D_S(T) \geq 3|S| + 2|T| - |S| = 2(|S| + |T|). \tag{2.5}$$

As

$$2e(S \cup T) = D(S) + D_S(T) + 2e(T), \tag{2.6}$$

$e(T)$ being the number of edges in the subgraph $G(T)$ induced by T, we see that

$$e(S \cup T) \geq |S \cup T|.$$

The rest of the argument is needed to upgrade this to the strict inequality.

From the proof of (2.5), and (2.6), it follows that the edge density of $G(S \cup T)$ may be equal 1 only if

(a)

$$D(S) + D_S(T) = 2(|S| + |T|);$$

(b) there are no edges in T;
(c) $|T_1| = |S|$;
(d) each vertex in S has exactly two neighbors in $S \cup (T \setminus T_1)$;
(e) each vertex in $T_2 := T \setminus T_1$ has exactly two neighbors in S.
If one of (c), (d), (e) is violated then $D(S) + D_S(T) > 2(|S| + |T|)$.

Case $T_2 = \emptyset$. Given a path P, the S-vertices are distributed over P as subpaths of vertex length $i \geq 1$, next neighbors of subpaths being T_1-vertices. Since T_1-vertices remain the neighbors of their single S-neighbors on every path, and $|T_1| = |S|$, there can be only paths of length 1 and 2, "monomers" and "dimers". An endpoint s of P is a monomer, as its left neighbor is $t \in T_1$. There are no other monomers in P, since an interior monomer would be flanked by two T_1-vertices, sharing a common neighbor in S, which is impossible. Consequently, the leftmost subpath of P is a dimer s_1, s_2, sandwiched between two vertices $t_1, t_2 \in T_1$. No rotation from P can use either t_i, as s_i is the only S-neighbor of t_i, or s_2, as $t_2 \notin S$. If the rotation uses s_1, then t_1, s_1, s, t becomes the new leftmost dimer with s_1 retaining the left position. Of course, if a rotation does not use s_1, then t_1, s_1, s_2, t_2 remains the leftmost subpath. So no sequence of rotations will make s_1 an endpoint. Contradiction.

Case $T_2 \neq \emptyset$. By (d)-(e), the graph $G(S \cup T_2)$ is a disjoint union of cycles. By (b), each cycle contains at least two vertices from S. In fact, there is just one cycle, since otherwise there would exist two vertices $s_1, s_2 \in S$ such that no sequence of rotations starting with a path with the endpoint s_1 would lead to a path ending at s_2.

Since there are no edges between vertices in T_2, two vertices from T_2 cannot be neighbors on the cycle. And no two vertices from S can be neighbors either. Otherwise, there is an arc $s_1 s_2 t$, with $s_1, s_2 \in S$, $t \in T_2$. Consider a path Q that ends at s_1. We know that the left neighbor of s_1 in Q is a $t_1 \in T_1$. By considering the rotation from Q via the edge (s_1, s_2) we see that s_2 has another neighbor $s_3 \in S$ distinct from s_1. Hence s_2 has at least three neighbors in $S \cup T_2$, namely s_1, s_3, t. This violates (d).

Therefore the vertices from S and from T_2 alternate on the cycle. Hence $|T_2| = |S|$, whence

$$|T| = |T_1| + |T_2| = 2|S|,$$

which violates Posa's inequality (2.1).

So the edge density of $G(S \cup T)$ exceeds 1. $\square$

Remark 2.1. *The above argument needs to be refined if we want to put a bound on the time taken construct the end-point sets. In this case suppose that we are doing a sequence of rotations with fixed endpoint v_0. If a rotation would produce an endpoint that has been produced before in this sequence, then we do not do this rotation. This limits the time spent producing endpoints, but it will reduce the number of endpoints, but we will now argue that Lemma 2.1 continues to hold. Indeed, all we have to observe is that* (2.3) *continues to hold. The argument being identical to Posá's argument.*

In the course of the proof, having assumed that the edge density of $G(S \cup T)$ is 1, we saw that then $G(S \cup T)$ must be quite special. Namely $|T_1| = |S|$, and either (1) $T_2 = \emptyset$ and $G(S \cup T)$ is a cycle on S, with each of T_1-vertices attached to its own S-vertex, or (2) $|T_2| = |S|$, and $G(S \cup T)$ is an alternating cycle on a bipartition (S, T_2), with each of S-vertices hosting its own pendant vertex from T_1. The punch line was that neither of these two graphs, each of edge density 1, can be a Posa's graph $G(S \cup T)$.

In the next section we will show that in the random graph $G^{(3)}(n, m)$ whp no vertex subset A, with $|A| \le \varepsilon_0 \ln n$ can induce a subgraph of edge density exceeding 1. So, by Lemma 2.1, whp $|S| > \varepsilon_0 \ln n$. We will also show that whp the edge density of the induced subgraph is $1 + o(1)$, for every A, with $\varepsilon_0 \ln n < |A| \le n^{1-o(1)}$. It is natural then to focus on the $o(n)$-Posa's sets of edge density close to 1, anticipating that the induced subgraphs $G(S \cup T)$ should interpolate between those two special, impossible, graphs. To prepare, let us have a look at the deterministic properties of $G(S \cup T)$ with an edge density close to 1.

Introduce $G^* = G^*(S \uplus (T \setminus T_1))$, a subgraph on the vertex set $S \cup (T \setminus T_1)$ whose edges have at least one end in S; so we disregard edges of $G(S \cup T)$ between vertices of T, and also edges joining the pendant vertices of T_1 to their respective S–"hosts". For $v \in S \cup (T \setminus T_1)$, let $d(v; G^*)$ denote the degree of v in G^*; by the definition of G^*, $\min_v d(v; G^*) \ge 2$. Introduce

$$S_2 = \{v \in S \,:\, d(v; G^*) = 2\}.$$

Lemma 2.2.

(i) No vertex from S_2 can be a neighbor of both a vertex in S_2 and a vertex in $T \setminus T_1$.

(ii) Suppose that

$$e(S \cup T) = (1+\sigma)(s+t), \quad s := |S|,\, t := |T|, \tag{2.7}$$

for some $\sigma > 0$. Then, denoting $|T_1| = t_1$,

$$s - 2\sigma(s+t) \le t_1 \le s, \tag{2.8}$$

$$\sum_{v \in S \cup (T \setminus T_1)} \big[d(v; G^*) - 2\big] \le 2\sigma(s+t). \tag{2.9}$$

Remark. Recalling that $t < 2s$, the bound (2.8) is not vacuous if $\sigma \le 1/6$.

Proof of Lemma 2.2. (i) Suppose that there are $s_1, s_2 \in S_2$ and $t \in T \setminus T_1$ such that (s_1, s_2) and (s_1, t) are edges in G^*. s_2 has a neighbor $t_1 \in T_1$, since s_2's degree in $G(S \cup T)$ is at least, whence exactly, 3. Consider a path P with s_2 as its endpoint. t_1 is necessarily a penultimate vertex of P. Rotating P via the edge (s_1, s_2) must make the right P-neighbor of s_1 a new endpoint. So s_1 has a neighbor in S distinct from s_2, and $d(s_1; G^*) \ge 3$. Contradiction.
(ii) First, using (2.4), $D(S) \ge 3s$, and (2.6), we obtain

$$2(s+t) + (s - t_1) \le 2e(S \cup T) \le 2(1+\sigma)(s+t),$$

which implies (2.8) as $s - t_1 \ge 0$. Second, the total vertex degree of G^* is

$$\sum_{v \in S \cup (T \setminus T_1)} d(v; G^*) = 2e(S \cup T) - 2t_1 - 2e(T).$$

Therefore

$$\begin{aligned} \sum_{v \in S \cup (T \setminus T_1)} \big[d(v; G^*) - 2\big] &= 2e(S \cup T) - 2t_1 - 2e(T) - 2(s + t - t_1) \\ &= 2e(S \cup T) - 2(s+t) - 2e(T) \\ &\le 2(1+\sigma)(s+t) - 2(s+t), \end{aligned}$$

which implies (2.9). $\square$

Introduce

$$\begin{aligned} &S_3 := S \setminus S_2, \\ &T_2 := \{v \in T \setminus T_1 : d(v; G^*) = 2\}, \quad T_3 := (T \setminus T_1) \setminus T_2, \end{aligned} \tag{2.10}$$

and denote $s_i = |S_i|$, $t_i = |T_i|$; so $s = s_2 + s_3$, $t = t_1 + t_2 + t_3$. It follows from (2.9) that

$$\sum_{v \in S_3 \cup T_3} \big[d(v; G^*) - 2\big] \le 2\sigma(s+t), \tag{2.11}$$

and then

$$|S_3 \cup T_3| = s_3 + t_3 \le 2\sigma(s+t). \tag{2.12}$$

Let μ_1 denote the total number of edges in the subgraph of $G^*(S \uplus (T \setminus T_1))$ induced by S, and let μ_2 denote the total number of the remaining edges of $G^*(S \uplus (T \setminus T_1))$, those joining vertices of S and $T \setminus T_1$, and set $\mu = \mu_1 + \mu_2$. Clearly

$$2\mu_1 + \mu_2 = \sum_{v \in S} d(v; G^*), \tag{2.13}$$

$$\mu_2 = \sum_{v \in T \setminus T_1} d(v; G^*). \tag{2.14}$$

Adding the equations (2.13) and (2.14),

$$\begin{aligned} \mu := \frac{1}{2} \sum_{v \in (S \cup T) \setminus T_1} d(v; G^*) = \; & s_2 + t_2 + \frac{1}{2} \sum_{v \in S_3 \cup T_3} d(v; G^*) \\ = s + t - t_1 + \frac{\xi_1 + \xi_2}{2}, & \end{aligned} \tag{2.15}$$

where

$$\xi_1 := \sum_{v \in S_3} \big[d(v; G^*) - 2\big] \ge s_3, \quad \xi_2 := \sum_{v \in T_3} \big[d(v; G^*) - 2\big] \ge t_3. \tag{2.16}$$

It follows from (2.14) and (2.13) that

$$\begin{aligned} &\mu_1 = s - t + t_1 + \frac{\xi_1 - \xi_2}{2}, \\ &\mu_2 = 2(t - t_1) + \xi_2. \end{aligned} \tag{2.17}$$

From (2.11),

$$\xi_1 + \xi_2 \le 2\sigma(s+t). \tag{2.18}$$

□

Remark 2.2. *We note here that Lemma 2.2 continues to hold under the restrictions described in Remark 2.1.*

Let two disjoint sets, S and T, the partitions $S = S_2 \cup S_3$, $T = T_1 \cup T_2 \cup T_3$, and ξ_1, ξ_2 be given. Let $\mathcal{N}(\mathbf{S}, \mathbf{T}, \boldsymbol{\xi})$ denote the total number of the subgraphs $G^*(S \uplus (T \setminus T_1))$, with μ_1, μ_2 determined by (2.17), such that the constraints (2.8), whence the constraints (2.11), (2.12) and (2.18) hold for some $\sigma > 0$.

Lemma 2.3.

(i)

$$\begin{aligned}
&\mathcal{N}(\mathbf{S}, \mathbf{T}, \boldsymbol{\xi}) \le \mathcal{N}_1(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}),\\
&\mathcal{N}_1(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}) := 2^{-s_2-t_2-\mu_1} \frac{(2\mu_1+\mu_2)!}{\mu_1!} \exp\left[O\big(\sigma(s+t)\big)\right].
\end{aligned} \tag{2.19}$$

(ii) There exists $\sigma_0 \in (0,1)$ such that, for $\sigma \le \sigma_0$ and

$$(1+\sigma^{1/2})\, s \le t \le 2(1-\sigma^{1/2})\, s, \tag{2.20}$$

a stronger bound holds:

$$\begin{aligned}
\mathcal{N}(\mathbf{S}, \mathbf{T}, \boldsymbol{\xi}) &\le \mathcal{N}_2(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}) := \mathcal{N}_1(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})(s+t)^2\\
&\times \exp\left[-(2s-t)\ln\frac{s}{2s-t} - (t-s)\ln\frac{s}{t-s} + O\big(\sigma^{1/2}(s+t)\big)\right].
\end{aligned} \tag{2.21}$$

Proof of Lemma 2.3. It is well known, Bollobás [4], that $g(\mathbf{d})$, the total number of graphs on $[\nu]$ with vertex degrees $d_1, \dots, d_\nu$, and total vertex degree $2M := \sum_i d_i$ satisfies

$$g(\mathbf{d}) \le (2M-1)!! \prod_{i=1}^{\nu} \frac{1}{d_i!}. \tag{2.22}$$

Here is a bipartite counterpart of (2.22). Let ν_1, ν_2, and $\mathbf{d}' = (d'_1, \dots, d'_{\nu_1})$, $\mathbf{d}'' = (d''_1, \dots, d''_{\nu_2})$ be such that

$$\sum_{i \in [\nu_1]} d'_i = \sum_{j \in [\nu_2]} d''_j = M.$$

Denote by $g(\mathbf{d}', \mathbf{d}'')$ the total number of bipartite graphs on a bipartition $[\nu_1] \uplus [\nu_2]$, with the left vertices and the right vertices having degrees $\mathbf{d}'$ and $\mathbf{d}''$. Then

$$g(\mathbf{d}', \mathbf{d}'') \le M! \prod_{i \in [\nu_1]} \frac{1}{d_i'!} \prod_{j \in [\nu_2]} \frac{1}{d_j''!}. \tag{2.23}$$

(i) Let $\mathbf{d} = \{d_v\}_{v \in S}$ be the (generic) vertex degrees of a subgraph of $G^*(S \uplus (T \setminus T_1))$ induced by S; so

$$\sum_{v \in S} d_v = 2\mu_1. \tag{2.24}$$

Let $\mathbf{d}' = \{d_v'\}_{v \in S}$ and $\mathbf{d}'' = \{d_v''\}_{v \in T_2 \cup T_3}$ denote the vertex degrees of the complementary bipartite graph on the bipartition $S \uplus (T_2 \cup T_3)$; so

$$\sum_{v \in S} d_v' = \sum_{v \in T_2 \cup T_3} d_v'' = \mu_2. \tag{2.25}$$

Here $\mu_1, \mu_2, \mu = \mu_1 + \mu_2$ are given by (2.13), (2.14) and (2.15). In addition,

$$d_v + d_v' \begin{cases} = 2, & v \in S_2, \\ \ge 3, & v \in S_3, \end{cases} \tag{2.26}$$

and

$$d_v \begin{cases} = 2, & v \in T_2, \\ \ge 3, & v \in T_3. \end{cases} \tag{2.27}$$

Using (2.22) and (2.23), we get an upper bound for the number of the graphs with vertex degrees $\mathbf{d}, \mathbf{d}', \mathbf{d}''$:

$$(2\mu_1 - 1)!!\, \mu_2!\, 2^{-s_2 - t_2} \prod_{v \in S_3} \frac{1}{d_v!\, d_v'!} \prod_{v \in T_3} \frac{1}{d_v''!}.$$

Introducing

$$f_k(x) = \sum_{j \ge k} \frac{x^j}{j!},$$

we have then

$$\sum_{\substack{\mathbf{d},\mathbf{d}',\mathbf{d}''\text{ meet}\\(2.24)-(2.27)}}\prod_{v\in S}\frac{1}{d_v!\,d'_v!}\prod_{v\in T_2\cup T_3}\frac{1}{d''_v!}$$

$$\begin{aligned}
&= 2^{-t_2}\,[x^{2\mu_1}y^{\mu_2}]\left(\sum_{d+d'=2}\frac{x^d y^{d'}}{d!\,d'!}\right)^{s_2}\left(\sum_{d+d'\geq 3}\frac{x^d y^{d'}}{d!\,d'!}\right)^{s_3}\cdot[z^{\mu_2-2t_2}]\left(\sum_{d''\geq 3}\frac{z^{d''}}{d''!}\right)^{t_3}\\
&= 2^{-t_2}\,[x^{2\mu_1}y^{\mu_2}]\left[\frac{(x+y)^2}{2}\right]^{s_2}\big[f_3(x+y)\big]^{s_3}\cdot[z^{\mu_2-2t_2}]\,\big[f_3(z)\big]^{t_3}\\
&= 2^{-s_2-t_2}\binom{2\mu_1+\mu_2}{\mu_2}[\xi^{2\mu_1+\mu_2}]\,\xi^{2s_2}f_3(\xi)^{s_3}\cdot[z^{\mu_2-2t_2}]\,\big[f_3(z)\big]^{t_3}\\
&= 2^{-s_2-t_2}\binom{2\mu_1+\mu_2}{\mu_2}[\xi^{2\mu_1+\mu_2-2s_2}]\,f_3(\xi)^{s_3}\cdot[z^{\mu_2-2t_2}]\,\big[f_3(z)\big]^{t_3}\\
&\leq 2^{-s_2-t_2}\binom{2\mu_1+\mu_2}{\mu_2}f_3(1)^{s_3+t_3}.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathcal{N}(\mathbf{S},\mathbf{T},\boldsymbol{\xi}) &\leq (2\mu_1-1)!!\,\mu_2!\,2^{-s_2-t_2}\binom{2\mu_1+\mu_2}{\mu_2}f_3(1)^{s_3+t_3}\\
&= 2^{-s_2-t_2}\frac{(2\mu_1-1)!!}{(2\mu_1)!}(2\mu_1+\mu_2)!\,f_3(1)^{s_3+t_3}.
\end{aligned}$$

So, by (2.16) and (2.18),

$$\mathcal{N}(\mathbf{S},\mathbf{T},\boldsymbol{\xi}) \leq \mathcal{N}_1(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) := 2^{-s_2-t_2-\mu_1}\frac{(2\mu_1+\mu_2)!}{\mu_1!}\exp\big[O(\sigma(s+t))\big]. \quad (2.28)$$

(ii) Let $\{d_v\}_{v\in S_2}$, $\{d_v\}_{v\in S_3}$ be the vertex degrees of the subgraphs of $G^*(S\uplus(T\setminus T_1))$ induced by S_2 and S_3 respectively. Let $\{\delta_v\}_{v\in S_2}$, $\{\delta_v\}_{v\in S_3}$ denote the vertex degrees of a bipartite graph induced by the bipartition $S_2\uplus S_3$. Finally, let $\{d'_v\}_{v\in S}$, $\{d''_v\}_{v\in T\setminus T_1}$ denote the vertex degrees of a bipartite graph induced by the bipartition $S\uplus(T\setminus T_1)$. By the definition,

$$\begin{aligned}
&d_v+\delta_v+d'_v=2,\ (v\in S_2),\quad d_v+\delta_v+d'_v\geq 3,\ (v\in S_3),\\
&d''_v=2,\ (v\in T_2),\quad d''_v\geq 3,\ (v\in T_3),
\end{aligned}$$

and by Lemma 2.2

$$d_v \cdot d'_v = 0, \quad v \in S_2. \quad (!)$$

Denote

$$\sum_{v\in S_2} d_v = 2\nu_2, \quad \sum_{v\in S_2} \delta_v = \nu_{2,3}, \quad \sum_{v\in S_2} d'_v = \mu_{2,2},$$
$$\sum_{v\in S_3} d_v = 2\nu_3, \quad \sum_{v\in S_3} \delta_v = \nu_{3,2}, \quad \sum_{v\in S_3} d'_v = \mu_{3,2};$$

then $\nu_{2,3} = \nu_{3,2}$, and

$$\begin{aligned} &2\nu_2 + \nu_{2,3} + \mu_{2,2} = 2|S_2| = 2s_2, \\ &2\nu_3 + \nu_{3,2} + \mu_{3,2} = 2\mu_1 + \mu_2 - 2s_2, \\ &\mu_{2,2} + \mu_{3,2} = \mu_2, \\ &\nu_2 + \nu_{2,3} + \nu_3 = \mu_1. \end{aligned} \tag{2.29}$$

Given the values of $\nu_2, \nu_3, \nu_{2,3}$, and $\mu_{2,2}, \mu_{3,2}$, the number of the corresponding subgraphs $G^*(S \uplus (T \setminus T_1))$ is bounded, as in part (i), by

$$\begin{aligned} &(2\nu_2 - 1)!!\,(2\nu_3 - 1)!!\,\nu_{2,3}!\,\mu_2!\,2^{-t_2}\,f_3(1)^{t_3} \\ &\qquad \times \left[x_1^{2\nu_2}\,x_2^{\mu_{2,2}}\,x_3^{\nu_{2,3}}\right] \left(\sum_{\substack{d+d'+\delta=2 \\ d\cdot d'=0}} \frac{x_1^d\,x_2^{d'}\,x_3^\delta}{d!\,d'!\,\delta!} \right)^{s_2} \\ &\qquad\qquad \times \left[y_1^{2\nu_3}\,y_2^{\mu_{3,2}}\,y_3^{\nu_{3,2}}\right] \left(\sum_{d+d'+\delta\ge 3} \frac{y_1^d\,y_2^{d'}\,y_3^\delta}{d!\,d'!\,\delta!} \right)^{s_3}. \end{aligned} \tag{2.30}$$

The last line factor is

$$\begin{aligned} &\binom{2\nu_3 + \mu_{3,2} + \nu_{3,2}}{2\nu_3,\ \mu_{3,2},\ \nu_{3,2}} \left[y^{2\nu_3+\mu_{3,2}+\nu_{3,2}}\right] f_3(y)^{s_3} \\ &\qquad \le 3^{2\nu_3+\mu_{3,2}+\nu_{3,2}}\,f_3(1)^{s_3} = 3^{2\mu_1+\mu_2-2s_2}\,f_3(1)^{s_3}. \end{aligned} \tag{2.31}$$

By (2.17) and (2.18), this is

$$2\mu_1 + \mu_2 - 2s_2 = 2s_3 + \xi_1. \tag{2.32}$$

Now $s_3 \le \xi_1$, see (2.16), and so the RHS of (2.31) is bounded by

$$3^{2s_3+\xi_1}\,f_3(1)^{s_3} \le \big(27 f_3(1)\big)^{\xi_1} \le \big(27 f_3(1)\big)^{2\sigma(s+t)}. \tag{2.33}$$

The second line factor in (2.30) is bounded by

$$\frac{\left(\frac{x_1^2}{2}+x_1x_3+\frac{(x_2+x_3)^2}{2}\right)^{s_2}}{x_1^{2\nu_2}\,x_2^{\mu_{2,2}}\,x_3^{\nu_{2,3}}}, \tag{2.34}$$

for all $x_1, x_2, x_3 > 0$. The challenge is to select the "best" x_1, x_2, x_3. First of all $\mu_{3,2}$, $\nu_{2,3} \le 3\xi_1$, since $\nu_{2,3} = \nu_{3,2}$ and by (2.16), (2.29) and (2.32)

$$2\nu_3 + \mu_{3,2} + \nu_{3,2} \le 3\xi_1.$$

Therefore, by (2.8), (2.11), (2.17), (2.18) and (2.29),

$$\nu_2 = \mu_1 - \nu_{2,3} - \nu_3 \le 2s - t + O(\sigma(s+t)), \tag{2.35}$$
$$\mu_{2,2} = \mu_2 - \mu_{3,2} = 2(t-s) + O(\sigma(t+s)). \tag{2.36}$$

By the condition in the Lemma, the explicit terms are of order $\sigma^{1/2}(s+t)$ at least, thus dwarf the remainders if σ is small. We pick

$$x_1 = (2\nu_2)^{1/2}, \quad x_2 = (\mu_{2,2})^{1/2}, \quad x_3 = (\nu_{2,3})^{1/2}.$$

Then

$$\begin{aligned}\frac{x_1^2}{2}+x_1x_3+\frac{(x_2+x_3)^2}{2} &= \frac{1}{2}(x_1^2+x_2^2+x_3^2)+(x_1+x_2)x_3\\ &= s_2 + O\big(\sqrt{(s+t)\xi}\big)\\ &= s + O\big(\sigma^{1/2}(s+t)\big),\end{aligned}$$

as $s_2 = s + O\big(\sigma(s+t)\big)$. So the fraction in (2.34) can be bounded from above by

$$\begin{aligned}&\frac{1}{2^{s_2}}\sqrt{\frac{(2s_2)^{2s_2}}{(2\nu_2)^{2\nu_2}\,\mu_{2,2}^{\mu_{2,2}}\,\nu_{2,3}^{\nu_{2,3}}}}\ \exp\big(O(\sigma^{1/2}(s+t))\big)\\ &= \frac{1}{2^{s_2}}\exp\left(\nu_2\ln\frac{s_2}{\nu_2}+\frac{\mu_{2,2}}{2}\ln\frac{s_2}{\mu_{2,2}/2}+\frac{\nu_{2,3}}{2}\ln\frac{s_2}{\nu_{2,3}/2}\right)\exp\big(O(\sigma^{1/2}(s+t))\big)\\ &= \frac{1}{2^{s_2}}\exp\left[(2s-t)\ln\frac{s}{2s-t}+(t-s)\ln\frac{s}{t-s}+O\big(\sigma^{1/2}(s+t)\big)\right].\end{aligned} \tag{2.37}$$

Turn to the first line of (2.30). Using $(2a-1)!! \le 2^a a!$, and (2.29), (2.35), we get

$$\begin{aligned}(2\nu_2-1)!!\,(2\nu_3-1)!!\,\nu_{2,3}!\,\mu_2! \le\ & 2^{\nu_2+\nu_3}\nu_2!\,\nu_{2,3}!\,\nu_3!\,\mu_2! \le 2^{\nu_2+\nu_3}\mu_1!\,\mu_2! \\ =\ & 2^{\mu_1}\mu_1!\,\mu_2!\,\exp\bigl(O(\sigma(s+t))\bigr). \qquad (2.38)\end{aligned}$$

Putting together (2.30), (2.33), (2.37) and (2.38), we conclude that the number of subgraphs $G^*(S \uplus (T \setminus T_1))$ with parameters $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ is bounded by

$$\frac{\mu_1!\,\mu_2!}{2^{s_2+t_2-\mu_1}}\exp\left[H(s,t)+O\bigl(\sigma^{1/2}(s+t)\bigr)\right], \qquad (2.39)$$

where

$$H(s,t) = (2s-t)\ln\frac{s}{2s-t} + (t-s)\ln\frac{s}{t-s}. \qquad (2.40)$$

We emphasize that the remainder term estimate is uniform over the range of the parameters $\boldsymbol{\xi}$, $\boldsymbol{\nu}$.

Let us compare the bounds (2.28) and (2.39). We have

$$\frac{2^{-s_2-t_2-\mu_1}\frac{(2\mu_1+\mu_2)!}{\mu_1!}}{2^{-s_2-t_2+\mu_1}\,\mu_1!\,\mu_2!} = 2^{-2\mu_1}\binom{2\mu_1+\mu_2}{\mu_1,\ \mu_1,\ \mu_2}.$$

Using

$$\mu_1 = 2s-t+O\bigl(\sigma(s+t)\bigr), \quad \mu_2 = 2(t-s)+O\bigl(\sigma(s+t)\bigr),$$

(cf. (2.35), (2.36)), and the Lemma condition on $2s-t$, $t-s$, it is simple to show that the last expession is

$$\exp\left[2H(s,t)+O\bigl(\sigma^{1/2}(s+t)\bigr)\right].$$

Thus the bound (2.39) can be written as

$$\mathcal{N}_1(\mathbf{s},\mathbf{t})\exp\left[-2H(s,t)+O\bigl(\sigma^{1/2}(s+t)\bigr)\right]. \qquad (2.41)$$

The bound (2.41) implies (2.21), since the factor $(s+t)^2$ is an upper bound for the number of solutions $(\nu_2, \nu_{2,3}, \nu_3, \mu_{2,2}, \mu_{3,2})$ of (2.29). □

Let $G^{**}(S\cup T)$ denote $G^*(S\cup(T\setminus T_1))$ adorned with t_1 pendant T_1-vertices attached to some t_1 S-vertices. That is, $G^{**}(S\cup T)$ is $G(S\cup T)$ without the edges joining T-vertices to each other.

Given two disjoint sets S and T, let $N(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})$ denote the total number of graphs $G^{**}(S \cup T)$ with $|S_2| = s_2$, $|T_1| = t_1$, $|T_2| = t_2$ and the parameters ξ_1, ξ_2. The number of ways to select $T_1 \subset T$ of cardinality t_1 and then to match the vertices of T_1 with some t_1 vertices in S is $t_1!\binom{s}{t_1}\binom{t}{t_1}$. The number of ways to select $S_2 \subset S$ of cardinality s_2 and to select $T_2 \subset T \setminus T_1$ of cardinality t_2 is $\binom{s}{s_2}\binom{t-t_1}{t_2}$, at most. (We neglect the constraint that S_2 needs to be a subset of the set of S-partners of T_1-vertices.) The total count of possibilities is bounded by the product of those two.

Lemma 2.4. *In the notations of Lemma 2.3,*

(i)

$$N(\mathbf{s}, \mathbf{t}) \le t_1! \binom{s}{t_1}\binom{t}{t_1}\binom{s}{s_2}\binom{t-t_1}{t_2} \mathcal{N}_1(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}); \tag{2.42}$$

(ii) if σ is small enough, and

$$(1 + \sigma^{1/2})\, s \le t \le 2(1 - \sigma^{1/2})\, s,$$

then

$$N(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}) \le t_1! \binom{s}{t_1}\binom{t}{t_1}\binom{s}{s_2}\binom{t-t_1}{t_2} \mathcal{N}_2(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi}). \tag{2.43}$$

□

Motivated by Lemma 2.1 and Lemma 2.2, in the next section we will focus on subgraphs of $G^{(3)}(m, n)$ of size not exceeding $n^{1-o(1)}$, showing as promised above that the likely edge density of such subgraphs is asymptotic to 1. It will remain to prove in the last section that whp there are no Posa's sets S, T of low edge density, with $s + t = O\big(n^{1-o(1)}\big)$. Lemma 2.4 will be a key ingredient of that argument.

3 Edge density of subgraphs of $G^{(3)}(n, m)$.

Recalling the notation

$$f_k(x) = \sum_{j \ge k} \frac{x^k}{k!},$$

introduce λ, a unique positive root of

$$\frac{\lambda f_2(\lambda)}{f_3(\lambda)} = c := \frac{2m}{n}, \tag{3.1}$$

Lemma 3.1.

(i) For $\varepsilon_0 = (1/3)\ln^{-1}\bigl[27 f_3(8\lambda)/2^{10}\lambda f_2(\lambda)\bigr]$,

$$P\bigl(\exists\, A \subset [n],\, |A| \le \varepsilon_0 \ln n \,:\, e(A) > |A|\bigr) \to 0. \tag{3.2}$$

Consequently, whp there does not exist an endpoint set S of size below $\varepsilon_0 \ln n$.

(ii) Let $\sigma_n \to 0$, but $(\sigma_n \ln n)/\ln\ln n \to \infty$, and let $\rho_n = (\sigma_n \ln n)^{-1/2}$, so that $\rho_n \to 0$. Then

$$P\bigl(\exists\, A \subset [n],\, \varepsilon_0 \ln n \le |A| \le n^{1-\rho_n} \,:\, e(A) \ge (1+\sigma_n)|A|\bigr) \to 0. \tag{3.3}$$

In words, with high probability, the edge density of subgraphs induced by sets A, of size from $\varepsilon_0 \ln n$ to $n^{1-o(1)}$, is $1+o(1)$ at most.

Proof of Lemma 3.1. Our random graph $G^{(3)}(n,m)$ is distributed uniformly on the set of all $C(n,m)$ graphs of minimum degree at least 3, with m edges and n vertices. From a more general result in Pittel and Wormald [21], for $c > 3/2$ we have: for $n \to \infty$,

$$C(n,m) \sim (2\pi n \mathrm{Var}[Z])^{-1/2}(2m-1)!!\frac{f_3(\lambda)^n}{\lambda^{2m}}\, \exp\bigl(-\eta - \eta^2/2\bigr); \tag{3.4}$$

here λ is the root of (3.1), and Z is Poisson(λ) conditioned on being at least 3. Probabilistically, (3.1) says that $\mathrm{E}[Z] = 2m/n$. Also $\eta := c^{-1}\mathrm{E}\bigl[\binom{Z}{2}\bigr]$. Constant factors aside, the claim is that

$$C(n,m) = \Theta\left(n^{-1/2}(2m-1)!!\,\frac{f_3(\lambda)^n}{\lambda^{2m}}\right). \tag{3.5}$$

Let $d_1, \dots, d_n \ge 3$, meeting $\sum_i d_i = 2m$, be such that there exists a graph with the degree sequence $\mathbf{d} = (d_1, \dots, d_n)$; we call such $\mathbf{d}$ graphical. Existence of graphical $\mathbf{d}$'s for large n, m is a weak consequence of (3.5). Let $g(\mathbf{d})$ denote the total number of graphs for a graphical $\mathbf{d}$. Introduce $G_{\mathbf{d}}$, a random graph distributed uniformly on the set of all $g(\mathbf{d})$ graphs of a given graphical $\mathbf{d}$; obviously, $G_{\mathbf{d}}$ equals, in distribution, $G^{(3)}(n,m)$, conditioned on $\{\mathbf{d}(G^{(3)}(n,m)) = \mathbf{d}\}$. To handle $G_{\mathbf{d}}$ we use a random pairing model, Bollobás [4], defined as follows.

Introduce a partition of $[2m]$ into n disjoint subsets $Q_1, \dots, Q_n$, $|Q_i| = d_i$, and a set Ω of all $(2m-1)!!$ pairings ω of $2m$ points in $[2m]$. Each ω induces a unique *multigraph*: a pair $(u,v) \in \omega$ with $u, v \in Q_i$ becomes a loop at vertex i; a pair $(u,v) \in \omega$ with $u \in Q_i$, $v \in Q_j$ becomes an edge (i,j). Let ω be random, distributed uniformly on Ω. Let $MG_{\mathbf{d}} = MG_{\mathbf{d}}(\omega)$ denote the random multigraph induced by ω. And let $\Omega_g(\mathbf{d})$ be the set of all graphical ω's, those for which $MG_{\mathbf{d}}(\omega)$ is a simple graph, i.e. has neither loops nor multiple edges. Then $MG_{\mathbf{d}}(\omega)$, conditioned on $\omega \in \Omega_g(\mathbf{d})$, coincides, in distribution, with $G_{\mathbf{d}}$. This implies that, for any graph property $\mathcal{G}$,

$$\mathrm{P}(G_{\mathbf{d}} \in \mathcal{G}) = \frac{\mathrm{P}\big(\{MG_{\mathbf{d}} \in \mathcal{G}\} \cap \Omega_g(\mathbf{d})\big)}{\mathrm{P}(\Omega_g(\mathbf{d}))} \le \frac{\mathrm{P}\big(MG_{\mathbf{d}} \in \mathcal{G}\big)}{\mathrm{P}(\Omega_g(\mathbf{d}))}.$$

Crucially,

$$g(\mathbf{d}) = \frac{(2m-1)!!}{\prod_i d_i!}\ \mathrm{P}(\Omega_g(\mathbf{d})), \tag{3.6}$$

[4]. We conclude that

$$\begin{aligned}\mathrm{P}(G^{(3)}(n,m) \in \mathcal{G}) &= C(n,m)^{-1} \sum_{\mathbf{d}}\ \mathrm{P}(G_{\mathbf{d}} \in \mathcal{G}) g(\mathbf{d}) \\ &\le \frac{(2m-1)!!}{C(n,m)} \sum_{\mathbf{d}}\ \mathrm{P}(MG_{\mathbf{d}} \in \mathcal{G}) \prod_{i \in [n]} (1/d_i!);\end{aligned}$$

the sums are over all admissible graphical $\mathbf{d}$. So, by (3.5), uniformly for all graph properties $\mathcal{G}$,

$$\mathrm{P}(G^{(3)}(n,m) \in \mathcal{G}) \le_b n^{1/2} \frac{\lambda^{2m}}{f_3(\lambda)^n} \sum_{\mathbf{d}}\ \mathrm{P}(MG_{\mathbf{d}} \in \mathcal{G}) \prod_{i \in [n]} (1/d_i!). \tag{3.7}$$

(For brevity, we write $A \le_b B$ when $A = O(B)$ uniformly over parameters involved, and B is too long to compose nicely with the big O notation.) This bound is perfectly tailored for $\mathcal{G}$'s implicit in (3.2) and (3.3).

Part (i): Denote the probability in (3.2) by P_{n1}. Suppose that for some ω and $A = A(\omega) \subseteq [n]$, $|A| \le \varepsilon_0 \ln n$, the sub(multi)graph of $MG_{\mathbf{d}}(\omega)$ induced by A has more edges than vertices. Then there exists $k = k(\omega) \in [2, \varepsilon_0 \ln n]$

and point sets $A_{i_1}, \dots, A_{i_k}$ such that the pairing ω contains $(k+1)$ pairs of points from $A_{i_1} \cup \dots \cup A_{i_k}$.

Combining (3.7) and the union bound, we have then

$$P_{n1} \le_b n^{1/2} \frac{\lambda^{2m}}{f_3(\lambda)^n} \sum_{4 \le k \le \varepsilon_0 \ln n} \binom{n}{k} \frac{\big(2(k+1)-1\big)!!\,\big(2(m-k-1)-1\big)!!}{(2m-1)!!} \times \sum_{\mathbf{d}} \binom{d_{1:k}}{2(k+1)} \prod_{i=1}^{n} \frac{1}{d_i!}, \quad (3.8)$$

$d_{1:k} := d_1 + \dots + d_k$. The second line sum in (3.8) is

$$\sum_{3k \le d \le 2m} \binom{d}{2(k+1)} \sum_{\substack{d_1+\dots+d_k=d \\ d_i \ge 3}} \prod_{i=1}^{k} \frac{1}{d_i!} \sum_{\substack{d_{k+1}+\dots+d_n=2m-d \\ d_i \ge 3}} \prod_{i=k+1}^{n} \frac{1}{d_i!} \le_b \sum_{3k \le d \le 2m} \binom{d}{2(k+1)} \frac{f_3(x)^k}{x^d} \frac{f_3(\lambda)^{n-k}}{\lambda^{2m-d}},$$

for every $x > 0$. (We have used a bound

$$\sum_{\substack{\delta_1+\dots+\delta_j=\delta \\ \delta_i \ge 3}} \prod_{i=1}^{j} \frac{1}{\delta_i!} = [y^\delta]\, f_3(y)^j \le \frac{f_3(y)^j}{y^\delta}, \quad \forall\, y > 0.) \quad (3.9)$$

The ratio of two consecutive terms in the last sum is

$$\frac{d+1}{d-2k-1}\,\frac{\lambda}{x} \le \frac{3k+1}{k-1}\,\frac{\lambda}{x} \le 7\frac{\lambda}{x} = \frac{7}{8} < 1,$$

if

$$x = 8\lambda.$$

So the sum is of order

$$\binom{3k}{2(k+1)} \frac{f_3(8\lambda)^k}{(8\lambda)^{3k}} \frac{f_3(\lambda)^{n-k}}{\lambda^{2m-3k}}.$$

Using this bound, $\binom{n}{k} \le n^k/k!$, $2m/n = \lambda f_2(\lambda)/f_3(\lambda)$, and

$$\big(2(k+1)-1\big)!! = \frac{\big(2(k+1)\big)!}{2^{k+1}(k+1)!},$$

we easily transform (3.8) into

$$P_{n1} \le_b m^{-1/2} \sum_{k\le\varepsilon_0 \ln n} \frac{k\,(3k)!}{2^k(k!)^3}\left[\frac{n}{2m}\,\frac{\lambda^3}{(8\lambda)^3}\,\frac{f_3(8\lambda)}{f_3(\lambda)}\right]^k$$
$$\le_b n^{-1/2} \sum_{k\le\varepsilon_0 \ln n} k\left[\frac{27}{2^{10}}\,\frac{f_3(8\lambda)}{\lambda f_2(\lambda)}\right]^k \to 0, \quad (3.10)$$

as

$$\varepsilon_0 = (1/3)\ln^{-1}\left[\frac{27\,f_3(8\lambda)}{2^{10}\,\lambda f_2(\lambda)}\right].$$

Part (ii): Let P_{n2} be the probability in (3.3). This time we need to bound the probability that there exists $A \subset [n]$, of cardinality $k \in [\varepsilon_0 \ln n, n^{1-\rho_n}]$ that has at least $\ell = \ell(k) = \lceil (1+\sigma_n)k \rceil$ edges. The counterpart of (3.8) is

$$P_{n2} \le_b n^{1/2}\frac{\lambda^{2m}}{f_3(\lambda)^n} \sum_{\varepsilon_0 \ln n\le k\le n^{1-\rho_n}} \binom{n}{k}\frac{\big(2\ell-1\big)!!\,\big(2(m-\ell)-1\big)!!}{(2m-1)!!}$$
$$\times \sum_{\mathbf{d}} \binom{d_{1:k}}{2\ell}\prod_{i=1}^n \frac{1}{d_i!}. \quad (3.11)$$

The bottom sum is bounded by a sum

$$\sum_{3k\le d\le 2m} \binom{d}{2\ell}\,\frac{f_3(x)^k}{x^d}\,\frac{f_3(\lambda)^{n-k}}{\lambda^{2m-d}},$$

with the consecutive terms ratio bounded by $3.01\lambda/x \le 0.76$, if $x = 4\lambda$. So, like (3.10),

$$P_{n2} \le_b n^{1/2} \sum_{\varepsilon_0 \ln n\le k\le n^{1-\rho_n}} \binom{n}{k}\frac{\big(2\ell-1\big)!!\,\big(2(m-\ell)-1\big)!!}{(2m-1)!!}$$
$$\times \binom{3k}{2\ell}\gamma^k, \qquad \gamma := 4^{-3}\frac{f_3(4\lambda)}{f_3(\lambda)}. \quad (3.12)$$

Here

$$\binom{3k}{2\ell} \le \binom{3k}{2k} \le_b \left(\frac{3^3}{2^2}\right)^k,$$

and

$$\frac{(2\ell-1)!!\,\big(2(m-\ell)-1\big)!!}{(2m-1)!!}=\frac{\binom{m}{\ell}}{\binom{2m}{2\ell}}\le\binom{m}{\ell}^{-1}.$$

Using the last two bounds and

$$\binom{n}{k}\le\Big(\frac{en}{k}\Big)^k,\quad m^{-1/2}\Big(\frac{m}{\ell}\Big)^{\ell}\le_b\binom{m}{\ell},$$

we simplify (3.12) to

$$P_{n2}\le_b n\sum_{\varepsilon_0\ln n\le k\le n^{1-\rho_n}}\Big(\frac{n}{k}\Big)^k\Big(\frac{m}{\ell}\Big)^{-\ell}\gamma_1^k,\quad \gamma_1:=e\,3^3\,2^{-2}\,\gamma.$$

Here, since $\ell\ge(1+\sigma_n)k$,

$$\begin{aligned}\Big(\frac{n}{k}\Big)^k\Big(\frac{m}{\ell}\Big)^{-\ell}\gamma_1^k&\le\Big(\frac{n}{k}\Big)^k\Big(\frac{n}{k(1+\sigma_n)}\Big)^{-k(1+\sigma_n)}\\&\le\Big(\frac{n}{k}\Big)^{-k\sigma_n}\cdot\big[(1+\sigma_n)^{-(1+\sigma_n)}\big]^k\gamma_1^k.\end{aligned}$$

The last expression is decreasing for $k\le n^{1-\rho_n}$, because its logarithmic derivative is

$$\begin{aligned}&-\sigma_n\ln\frac{n}{k}+\sigma_n-(1+\sigma_n)\ln(1+\sigma_n)+\ln\gamma_1\\&\qquad\le-\sigma_n\ln\frac{n}{k}+\ln\gamma_1\le-\sigma_n\rho_n\ln n+\ln\gamma_1\\&\qquad\qquad=-(\sigma_n\ln n)^{1/2}+\ln\gamma_1\to-\infty,\end{aligned}$$

as $\sigma_n\ln n\to\infty$. So

$$P_{n2}\le_b\exp\big[-\varepsilon_0\sigma_n(\ln n)^2+O\big((\ln n)\ln\ln n\big)\big]\to 0,$$

as $\sigma_n\ln n\gg\ln\ln n$. □

4 Moderately large, sparse Posa's sets are unlikely.

Let $d_{\max}=d_{\max}(n,m)$ denote the largest vertex degree in $G^{(3)}(n,m)$. Then let S,T be disjoint subsets of $[n]$, of cardinalities s and t, with $t<2s$. In view of Lemma 3.1, part (i), we may and will confine ourselves to $s+t \ge \varepsilon_0 \ln n$. Lemma 2.4 asserts two upper bounds for the total number of subgraphs $G^{**}(S\cup T)$ with parameters $\mathbf{s}$, $\mathbf{t}$ and $\boldsymbol{\xi}$. (See (2.16) for definition of ξ_1 and ξ_2.) Let us bound the number of ways to extend this subgraph to a graph on $[n]$, of minimum degree 3 at least, with m edges.

Recall that the edge set of $G^{**}(S\cup T)$ does not contain edges between T-vertices. So any such extension of $G^{**}(S\cup T)$ is determined by an induced subgraph $G(S^c)$. Let d_i denote the degree of vertex $i\in S^c$ in $G(S^c)$. An admissible $\mathbf{d}=\{d_i\}_{i\in S^c}$ meets the conditions

$$d_i \ge \begin{cases} 3, & i\in S^c\setminus T,\\ 3-i, & i\in T_i,\ i=1,2,3,\end{cases} \tag{4.1}$$

and

$$\sum_{i\in S^c} d_i = 2m-2D, \quad D:=\mu+t_1=o(n). \tag{4.2}$$

Then, by (2.22) and the definition of $f_k(y)$, the number of ways to extend a given $G^{**}(S\cup T)$ is bounded above by

$$\begin{aligned}
&\big(2(m-D)-1\big)!!\sum_{\substack{\mathbf{d}\text{ meets}\\ (4.1)-(4.2)}}\prod_{i\in S^c}\frac{1}{d_i!}\\
&\quad= \big(2(m-D)-1\big)!!\,[y^{2(m-D)}]\prod_{i=1}^3\Bigg(\sum_{d\ge 3-i}\frac{y^d}{d!}\Bigg)^{t_i}\Bigg(\sum_{d\ge 3}\frac{y^d}{d!}\Bigg)^{n-s-t}\\
&\quad= \big(2(m-D)-1\big)!!\,[y^{2(m-D)}]\prod_{i=1}^3 f_{3-i}(y)^{t_i}\cdot f_3(y)^{n-s-t}.
\end{aligned}$$

By the Cauchy integral formula,

$$\begin{aligned}
&[y^{2(m-D)}]\prod_{i=1}^3 f_{3-i}(y)^{t_i}\cdot f_3(y)^{n-s-t}\\
&\qquad= \frac{1}{2\pi i}\oint\limits_{|y|=r}\frac{1}{y^{2(m-D)+1}}\prod_{i=1}^3 f_{3-i}(y)^{t_i}\cdot f_3(y)^{n-s-t}\,dy.
\end{aligned}$$

Here $n-s-t\sim n$. Using $|f_k(y)|\le f_k(|y|)$, an inequality ([20])

$$|f_3(y)| \le f_3(|y|)\exp\left(-\frac{|y|-\text{Re }y}{4}\right),$$

and selecting $r = \lambda$, we obtain

$$\begin{aligned}
&\left|\frac{1}{2\pi i}\oint\limits_{|y|=r}\frac{1}{y^{2(m-D)+1}}\prod_{i=1}^{3}f_{3-i}(y)^{t_i}\cdot f_3(y)^{n-s-t}\,dy\right|\\
&\le_b \frac{1}{\lambda^{2(m-D)}}\prod_{i=1}^{3}f_{3-i}(\lambda)^{t_i}\cdot f_3(\lambda)^{n-s-t}\int_{\theta=-\pi}^{\pi}e^{-(n-o(n))\lambda(1-\cos\theta)/4}d\theta\\
&\le_b \frac{1}{n^{1/2}\lambda^{2(m-D)}}\prod_{i=1}^{3}f_{3-i}(\lambda)^{t_i}\cdot f_3(\lambda)^{n-s-t}
\end{aligned}$$

And so the number of extensions of a given $G^{**}(S\cup T)$ is of order

$$N_{\rm ext}(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) := \frac{\big(2(m-D)-1\big)!!}{n^{1/2}\lambda^{2(m-D)}}\prod_{i=1}^{3}f_{3-i}(\lambda)^{t_i}\cdot f_3(\lambda)^{n-s-t},$$

at most. Then, multiplying $N_{\rm ext}(\mathbf{s},\mathbf{t},\boldsymbol{\xi})$ by $N_1(\mathbf{s},t,\boldsymbol{\xi})$, the first bound given in Lemma 2.4, we get an upper bound for the total number of graphs on $[n]$ with m edges, such that $S\cup T$ induces a subgraph $G(S\cup T)$ with parameters $\mathbf{s},\mathbf{t},\boldsymbol{\mu}$. Multiplying $N(\mathbf{s},t,\boldsymbol{\xi})N_{\rm ext}(\mathbf{s},\mathbf{t},\boldsymbol{\xi})$ by $\binom{n}{s,t}\le n^{s+t}/s!t!$, and dividing by $C(n,m)$, the total number of the (n,m)-graphs of minimum degree 3 at least, we obtain a bound $O\big(E_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi})\big)$ for the expected number of Posa's subgraphs with parameters $\mathbf{s},\mathbf{t},\boldsymbol{\mu}$, where

$$\begin{aligned}
E_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) = n^{s+t}\,&\frac{(2\mu_1+\mu_2)!}{2^{\mu_1}\mu_1!}\,\frac{\big(2(m-D)-1\big)!!}{(2m-1)!!}\exp\big[O\big(\sigma(s+t)\big)\big]\\
&\times\frac{\lambda^{2D}}{2^{s_2+t_2}f_3(\lambda)^{s+t}}\prod_{i=1}^{3}f_{3-i}(\lambda)^{t_i}\cdot[f_3(1)]^{s_3+t_3}\\
&\qquad\times\frac{1}{s!\,t!}\,t_1!\binom{s}{t_1}\binom{t}{t_1}\binom{s}{s_2}\binom{t-t_1}{t_2}. \qquad (4.3)
\end{aligned}$$

(See (2.15) and (2.29) for μ_1,μ_2 expressed through ξ_1 and ξ_2.) In view of (2.12)-(2.11) and Lemma 3.1, part (ii), if we allow only $s+t\le n^{1-\rho_n}$, $\rho_n\to 0$, which we do, we need to consider only $\mathbf{s},\mathbf{t},\boldsymbol{\xi}$ such that

$$0\le s-t_1\le 2\sigma_n(s+t),\quad s_3+t_3\le 2\sigma_n(s+t),\quad \xi_1+\xi_2\le 2\sigma_n(s+t), \quad (4.4)$$

where $\sigma_n \to 0$. (See Lemma 3.1 for a more precise definition of σ_n, ρ_n.) Our remaining task is to show that the sum of $E_{n,m}(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})$ over the admissible $(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})$ approaches zero.

To this end, let us first bound $E_{n,m}(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})$ by a simpler $E^*_{n,m}(\mathbf{s}, \mathbf{t}, \boldsymbol{\xi})$ times $\exp\bigl(o(s+t)\bigr)$. First, by (4.2) and (4.4), in the second line of (4.4)

$$\begin{aligned}\frac{\lambda^{2D}}{2^{s_2+t_2}}\, f_0(\lambda)^{t_3}\, [f_3(1)]^{s_3+t_3} &= \frac{\lambda^{2(s+t)}}{2^t}\exp\bigl(O(s_3+t_3+s-t_1)\bigr)\\ &= \frac{\lambda^{2(s+t)}}{2^t}\exp\bigl(O(\sigma_n(s+t)\bigr). \qquad (4.5)\end{aligned}$$

Next, using

$$(2a-1)!! = \frac{(2a)!}{2^a a!} = \Theta\left[\left(\frac{2a}{e}\right)^a\right],$$

we obtain that the second fraction fraction in the first line of (4.3) is of order

$$\begin{aligned}\left(\frac{e}{2m}\right)^D (1-D/m)^{m-D} &= (2m)^{-D}\, e^{O(D^2/m)}\\ &= (2m)^{-t_1-\mu}\exp\bigl(O(n^{-\rho_n}(s+t))\bigr)\\ &=(2m)^{-(s+t+\xi/2)}\exp\bigl(O(n^{-\rho_n}(s+t))\bigr), \qquad (4.6)\end{aligned}$$

$\xi := \xi_1 + \xi_2$. Further, by (2.8), (2.17) and (2.18),

$$\frac{(2\mu_1+\mu_2)!}{2^{\mu_1}\mu_1!} = \frac{(2s+\xi_1)!}{2^{2s-t}\bigl[s-t+t_1+(\xi_1-\xi_2)/2\bigr]!}\exp\bigl(O(\sigma_n(s+t))\bigr). \qquad (4.7)$$

Given ξ, the last fraction attains its *maximum* at $\xi_1 = \xi$, $\xi_2 = 0$, and it is

$$\begin{aligned}&\frac{(2s+\xi)!}{2^{2s-t}(s-t+t_1+\xi/2)!}\\ &\qquad = \frac{(s+t-t_1)!\,(\xi/2)!}{2^{2s-t}}\binom{2s+\xi}{s-t+t_1+\xi/2,\ \xi/2,\ s+t-t_1}. \qquad (4.8)\end{aligned}$$

The reason behind (4.8) is that the multinomial coefficients are amenable to easy but sharp estimates. The factorial $(s+t-t_1)!$ combined with $(t_1!/s!\,t!)\binom{s}{t_1}\binom{t}{t_1}$ in (4.3) will later produce another friendly trinomial coefficient.

Using an inequality

$$\binom{a+b+c}{a, b, c} \le \frac{(a+b+c)^{a+b+c}}{a^a\, b^b\, c^c}, \qquad (4.9)$$

the trinomial coefficient in (4.8) is bounded above by $e^{H_1(s,\mathbf{t},\xi)}$, where

$$H_1(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) = (s-t+t_1+\xi/2)\ln\frac{2s+\xi}{s-t+t_1+\xi/2} + \xi/2\ln\frac{2s+\xi}{\xi/2} + (s+t-t_1)\ln\frac{2s+\xi}{s+t-t_1}. \quad (4.10)$$

Consider the first summand. Notice that

$$s-t+t_1+\xi/2 = 2s-t-(s-t_1)+\xi/2 \le 2s-t+\xi/2,$$

and $2s-t>0$. Suppose that $2s-t \ge \sigma_n^{1/2}(s+t)$. Then, as $s-t_1$ and ξ are of order $O\big(\sigma_n(s+t)\big)$, the summand is

$$(2s-t)\ln\frac{2s}{2s-t} + O\big(\sigma_n^{1/2}(s+t)\big).$$

If $2s-t \le \sigma_n^{1/2}(s+t)$, then, as $x\ln(a/x)$ is increasing for $x \le a/e$, the summand is bounded above crudely by

$$(2s-t+\xi/2)\ln\frac{2s+\xi}{2s-t+\xi/2} \le 2\sigma_n^{1/2}(s+t)\ln\frac{3s}{\sigma_n^{1/2}(s+t)} \le (2s-t)\ln\frac{2s}{2s-t} + 2\sigma_n^{1/2}(\ln(1/\sigma_n))(s+t).$$

Thus the summand is always

$$(2s-t)\ln\frac{2s}{2s-t} + 2\sigma_n^{1/2}(\ln(1/\sigma_n))(s+t),$$

at most. For the second summand in (4.10),

$$\xi/2\ln\frac{2s+\xi}{\xi/2} \le \sigma_n(s+t)\ln\frac{3s}{\sigma_n(s+t)} = O\big(\sigma_n(\ln 1/\sigma_n)(s+t)\big).$$

The third summand in (4.10) is

$$t\ln\frac{2s}{t} + O\big(\sigma_n(s+t)\big).$$

Thus

$$H_1(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) \le (2s-t)\ln\frac{2s}{2s-t} + t\ln\frac{2s}{t} + O\big(\sigma_n^{1/2}(\ln(1/\sigma_n))(s+t)\big), \quad (4.11)$$

uniformly for $2s-t>0$. The equation (4.7) becomes

$$\frac{(2\mu_1+\mu_2)!}{2^{\mu_1}\mu_1!} \le \frac{(s+t-t_1)!\,(\xi/2)!}{2^{2s-t}} \times \exp\left[(2s-t)\ln\frac{2s}{2s-t}+t\ln\frac{2s}{t}\right]\cdot\exp\left[O\big(\sigma_n^{1/2}(\ln(1/\sigma_n))(s+t)\right]. \quad (4.12)$$

Collecting (4.5), (4.6) and (4.12), we conclude that

$$E_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) \le E^*_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi})\exp\left[O\big(\sigma_n^{1/2}(\ln(1/\sigma_n))(s+t)\right],$$

where

$$\begin{aligned} E^*_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi}) = {} & \frac{(\xi/2)!}{m^{\xi/2}}\,\frac{n^{s+t}\lambda^{2(s+t)}f_2(\lambda)^s f_1(\lambda)^{t-t_1}}{(2m)^{s+t}f_3(\lambda)^{s+t}\,2^{2s}} \\ & \times\exp\left[(2s-t)\ln\frac{2s}{2s-t}+t\ln\frac{2s}{t}\right] \\ & \times\frac{(s+t-t_1)!}{s!\,t!}\;t_1!\binom{s}{t_1}\binom{t}{t_1}\binom{s}{s_2}\binom{t-t_1}{t_2}. \end{aligned} \quad (4.13)$$

Here, recalling again (4.4),

$$s_3,\,t_3\le 2\sigma_n(s+t), \quad \xi\le 2\sigma_n(s+t).$$

Subject to this constraint, let us bound $\sum_{s_2,t_2,\xi} E^*_{n,m}(\mathbf{s},\mathbf{t},\boldsymbol{\xi})$. First of all,

$$\begin{aligned} \sum_{\xi}\frac{(\xi/2)!}{m^{\xi/2}} & \le \sum_{\xi}\left(\frac{e\xi}{2m}\right)^{\xi/2} \\ & \le \sum_{\xi\ge 0}\left(\frac{e\sigma_n(s+t)}{m}\right)^{\xi/2}\to 1. \end{aligned}$$

Secondly,

$$\begin{aligned} \sum_{\substack{s_2+s_3=s\\ s_3\le 2\sigma_n(s+t)}}\binom{s}{s_2} & = \sum_{s_3\le 2\sigma_n(s+t)}\binom{s}{s_3} \\ & \le_b \binom{s}{2\sigma_n(s+t)} \le \left(\frac{es}{2\sigma_n(s+t)}\right)^{2\sigma_n(s+t)} \\ & = \exp\left[O\big(\sigma_n(\ln(2/\sigma_n))(s+t)\big)\right]. \end{aligned}$$

Likewise

$$\sum_{\substack{t_2+t_3=t-t_1 \\ t_3\le 2\sigma_n(s+t)}} \binom{t-t_1}{t_2} = \exp\left[O\big(\sigma_n(\ln(2/\sigma_n))(s+t)\big)\right].$$

Observing also that

$$\frac{(s+t-t_1)!}{s!\,t!}\; t_1!\binom{s}{t_1}\binom{t}{t_1} = \binom{s+t-t_1}{s-t_1,\, t_1,\, t-t_1},$$

we then have

$$\sum_{s_2,t_2,\xi} E^*_{n,m}(\mathbf{s},\mathbf{t},\xi) \le E_{n,m}(s,t,t_1)\exp\left[O\big(\sigma_n(\ln(2/\sigma_n))(s+t)\big)\right],$$

where

$$\begin{aligned} E_{n,m}(s,t,t_1) := {} & \frac{n^{s+t}\lambda^{2(s+t)} f_2(\lambda)^s f_1(\lambda)^{t-t_1}}{(2m)^{s+t} f_3(\lambda)^{s+t}\, 2^{2s}} \\ & \times \exp\left[(2s-t)\ln\frac{2s}{2s-t} + t\ln\frac{2s}{t}\right]\cdot\binom{s+t-t_1}{s-t_1, t_1, t-t_1}. \end{aligned} \tag{4.14}$$

The trinomial coefficient in (4.14) is bounded above by

$$\exp\left[(s-t_1)\ln\frac{s+t-t_1}{s-t_1} + t_1\ln\frac{s+t-t_1}{t_1} + (t-t_1)\ln\frac{s+t-t_1}{t-t_1}\right]. \tag{4.15}$$

Recall that

$$s-2\sigma_n(s+t)\le t_1\le s \Longrightarrow 0\le s-t_1\le 2\sigma_n(s+t). \tag{4.16}$$

Since

$$(s-t_1)\ln\frac{s+t-t_1}{s-t_1} = (s-t_1)\ln\frac{t}{s-t_1} + O\big((s-t_1)^2/t\big),$$

and $x\ln(t/x)$ is increasing for $x<t/e$, we obtain

$$\begin{aligned} (s-t_1)\ln\frac{s+t-t_1}{s-t_1} &\le 2\sigma_n(s+t)\ln\frac{t}{2\sigma_n(s+t)} + O\big(\sigma_n^2(s+t)\big) \\ &= O\big(\sigma_n\ln(1/\sigma_n)(s+t)\big), \end{aligned} \tag{4.17}$$

where $\sigma_n \ln(1/\sigma_n) \to 0$, as $\sigma_n \to 0$. Furthermore,

$$t_1 \ln \frac{s+t-t_1}{t_1} = s \ln \frac{t}{s} + O(\sigma_n(s+t)). \tag{4.18}$$

Turn to the last summand in (4.15). By (4.16),

$$0 \le t - t_1 \le t - s + 2\sigma_n(s+t) < t - s + 3\sigma_n(s+t).$$

Further

$$\begin{aligned}(t-t_1) \ln \frac{s+t-t_1}{t-t_1} &= (t-t_1) \ln \frac{t}{t-t_1} + O(s-t_1) \\ &= (t-t_1) \ln \frac{t}{t-t_1} + O(\sigma_n(s+t)).\end{aligned}$$

Suppose that

$$t - s + 3\sigma_n(s+t) < t/e \tag{4.19}$$

which is equivalent to

$$t < \frac{s(1-3\sigma_n)}{1-e^{-1}+3\sigma_n}.$$

Then, for $t > s$,

$$\begin{aligned}(t-t_1) \ln \frac{t}{t-t_1} &\le \big(t-s+3\sigma_n(s+t)\big) \ln \frac{t}{t-s+3\sigma_n(s+t)} \\ &\le (t-s) \ln \frac{t}{t-s} + O\left[\sigma_n(s+t) \ln \frac{t}{\sigma_n(s+t)}\right] \\ &= (t-s) \ln \frac{t}{t-s} + O(\sigma_n \ln(1/\sigma_n)(s+t)).\end{aligned} \tag{4.20}$$

If $t \le s$, then (4.19)

$$\begin{aligned}(t-t_1) \ln \frac{t}{t-t_1} &\le \big(t-s+3\sigma_n(s+t)\big) \ln \frac{t}{t-s+3\sigma_n(s+t)} \\ &\le 3\sigma_n(s+t) \ln \frac{t}{3\sigma_n(s+t)} \\ &= O(\sigma_n \ln(1/\sigma_n)(s+t)).\end{aligned} \tag{4.21}$$

Suppose that

$$t \ge \frac{s(1-3\sigma_n)}{1-e^{-1}+3\sigma_n}.$$

Then $t - s = \Theta(s)$, and so

$$(t - t_1) \ln \frac{t}{t - t_1} - (t - s) \ln \frac{t}{t - s} = O\big(s - t_1\big) = O\big(\sigma_n(s + t)\big). \qquad (4.22)$$

Combining (4.14)-(4.22), we obtain

$$E_{n,m}(s,t) := \sum_{t_1} E_{n,m}(s,t,t_1) \le E^*_{n,m}(s,t) \exp\left[O\big(\sigma_n(\ln 1/\sigma_n)(s+t)\big)\right],$$

where

$$E^*_{n,m}(s,t) := t \frac{n^{s+t} \lambda^{2(s+t)}}{(2m)^{s+t} f_3(\lambda)^{s+t}\, 2^{2s}} f_2(\lambda)^s f_1(\lambda)^{t-s}$$
$$\times \exp\left[(2s - t) \ln \frac{2s}{2s - t} + t \ln \frac{2s}{t} + s \ln \frac{t}{s} + (t - s)^+ \ln \frac{t}{(t - s)^+}\right], \qquad (4.23)$$

where we define $x^+ = \max\{0, x\}$, and $0 \ln(t/0) = 0$.

The rest is a bit of calculus. Recalling that $2m/n = \lambda f_2(\lambda)/f_3(\lambda)$, and setting $t = xs$, we write

$$E^*_{n,m}(s,t) = sxe^{sH_1(x)},$$

where

$$H_1(x) = (1 + x) \ln \lambda - x \ln f_2(\lambda) + (x - 1) \ln f_1(\lambda)$$
$$+ (2 - x) \ln \frac{1}{2 - x} + (x - 1) \ln \frac{1}{x} + (x - 1)^+ \ln \frac{x}{(x - 1)^+}.$$

The exponent in (4.23) is

$$(2s - t) \ln \frac{2s}{2s - t} + t \ln \frac{2s}{t} + s \ln \frac{t}{s}$$
$$= 2s \ln 2 + (2s - t) \ln \frac{s}{2s - t} + t \ln \frac{s}{t} + s \ln \frac{t}{s}$$
$$= 2s \ln 2 + (2s - t) \ln \frac{s}{2s - t} + (t - s) \ln \frac{s}{t},$$

and the term $2s\ln 2$ cancels 2^{2s} in the first line fraction denominator. Since

$$H_1'(x) = \begin{cases} \ln\left(\frac{\lambda f_1(\lambda)}{f_2(\lambda)}\frac{2-x}{x}\right) + \frac{1}{x}, & x < 1, \\ \ln\left(\frac{\lambda f_1(\lambda)}{f_2(\lambda)}\,\frac{2-x}{x-1}\right), & x \in (1,2), \end{cases}$$

and $\lambda f_1(\lambda)/f_2(\lambda) > 2$, we see that $H_1(x)$ is unimodal on $(0,2)$, and attains its maximum at $x^* \in (1,2)$

$$x^* = \frac{1 + 2\lambda f_1(\lambda)/f_2(\lambda)}{1 + \lambda f_1(\lambda)/f_2(\lambda)},$$

and

$$H_1(x^*) = \ln\left[\frac{\lambda^2}{f_2(\lambda)}\big(1 + \lambda f_1(\lambda)/f_2(\lambda)\big)\right] \tag{4.24}$$

Maple shows that the function on the RHS of (4.24) increases with λ and it is zero at $\lambda^* = 5.162717...$. At the first glance it would seem necessary to put a constraint $\lambda > \lambda^*$in order to claim that , for those λ's, whp there are no Posa's sets of cardinality $|S| + |T| \le n^{1-o(1)}$.

We can do better though! Indeed, by the unimodality of $H_1(x)$,

$$\begin{aligned} \max\{H_1(x)\,:\, x \in [0, 1+\sigma_n^{-1/2}] \cup [2-\sigma_n^{1/2}, 2]\} \\ = \max\{H_1(1+\sigma_n^{-1/2}),\, H_1(2-\sigma_n^{1/2})\} \\ = \max\left\{\ln\frac{\lambda^2}{f_2(\lambda)},\, \ln\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2}\right\} + O\big(\sigma_n^{1/2}\ln(1/\sigma_n)\big) \\ = \ln\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2} + O\big(\sigma_n^{1/2}\ln(1/\sigma_n)\big), \end{aligned} \tag{4.25}$$

as $\lambda f_1(\lambda)/f_2(\lambda) > 2$. As for $x = t/s \in [1+\sigma_n^{1/2}, 2-\sigma_n^{1/2}]$, we use (2.43) instead of (2.42) and improve the bound (4.23) by the factor

$$(s+t)^2 \exp\left[-(2s-t)\ln\frac{s}{2s-t} - (t-s)\ln\frac{s}{t-s}.\right]$$

So we can re-define

$$E^*_{n,m}(s,t) = s^3 x(1+x)^2 e^{sH_2(x)},$$

where

$$H_2(x) = (1+x)\ln\lambda - x\ln f_2(\lambda) + (x-1)\ln f_1(\lambda),$$

a linear function! Now

$$\max\{H_2(x)\,:\,x\in[1,2]\} = H_2(2) = \ln\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2},$$

and this function decreases with λ.

Indeed, introducing $F(\lambda) = \lambda/f_1(\lambda)$ that decreases from 1 at 0+ to 0 at ∞, we have

$$\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2} = \frac{\lambda^3 f_1(\lambda)}{\big(f_1(\lambda)-\lambda\big)^2} = \lambda^2\frac{F(\lambda)}{\big(1-F(\lambda)\big)^2}.$$

So

$$\begin{aligned}\frac{d}{d\lambda}\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2} =& 2\lambda\frac{F(\lambda)}{\big(1-F(\lambda)\big)^2} + \lambda^2\frac{1+F(\lambda)}{\big(1-F(\lambda)\big)^3}\,F'(\lambda)\\ &\left(\text{using } F'(\lambda) = \frac{1}{f_1(\lambda)} - \frac{\lambda e^\lambda}{f_1(\lambda)^2} = \lambda^{-1}\big(F(\lambda) - e^\lambda F(\lambda)^2\big)\right)\\ =& \frac{\lambda F(\lambda)}{\big(1-F(\lambda)\big)^3}\left[3 - F(\lambda) - e^\lambda F(\lambda)\big(1+F(\lambda)\big)\right]\\ =& \frac{\lambda F(\lambda)}{\big(1-F(\lambda)\big)^3(e^\lambda-1)}\,D(\lambda);\end{aligned}$$

here

$$D(\lambda) = (3-\lambda)e^{2\lambda} - (6+\lambda^2)e^\lambda + \lambda + 3 = \sum_{j\ge 4} d_j\lambda^j,$$

and

$$d_j = 3\cdot 2^j - j2^{j-1} - j(j-1) - 6,\quad j\ge 4.$$

By induction on $j\ge 4$, $d_j<0$ for all $j\ge 4$. Hence

$$\frac{d}{d\lambda}\frac{\lambda^3 f_1(\lambda)}{f_2(\lambda)^2} < 0,\quad \forall\,\lambda>0.$$

Maple shows that $\lambda^3 f_1(\lambda)/f_2(\lambda)^2$ attains value 1 at

$$\lambda^{**} = 4.789771...$$

The corresponding average vertex degree

$$c^{**} = \frac{\lambda^{**} f_2(\lambda^{**})}{f_3(\lambda^{**})} = 5.323132...$$

It follows that for $m \geq 2.662n$ the expected number of the likely Posa's sets (S,T) of size $|S| + |T| \leq n^{1-o(1)}$ approaches zero as $n, m \to \infty$.

Remark 4.1. *As a final remark, observe that within the constraints on the $G(S \cup T)$, the dominant contribution to the total number of sparse Posa sets (S,T) comes from $G(S \cup T)$ very close to an alternating cycle on $S + T_2, |T_2| = |S| = s$, with the s pendant T_1 vertices attached to to S-vertices. It is not difficult to get directly the asymptotic expected number of such extreme subgraphs in our random graph, and it turns out essentially the same as the current estimate.*

What this likely means is that it is fruitless to search for another constraint on $G(S \cup T)$ with a potential to further decrease λ^{**} via a sharper bound for the expected number of Posa sets (S,T).

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, First occurrence of Hamilton cycles in random graphs, *Cycles in graphs (Burnaby, B.C.)* (1982) 173-178.

[2] J. Aronson, A.M. Frieze and B. Pittel, Maximum matchings in sparse random graphs: Karp-Sipser revisited, *Random Structures and Algorithms* **12** (1998) 111-178.

[3] T. Bohman and A.M. Frieze, Hamilton cycles in 3-out, *Random Structures and Algorithms*, **35** (2009) 393-417.

[4] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Comb.*, **1**, (1980) 311–316.

[5] B. Bollobás, The evolution of random graphs, *Trans. Amer. Math. Soc.*, **286**, (1984) 257–274.

[6] B. Bollobás, Random Graphs, Second Edition *Cambridge University Press*, (2001).

[7] B. Bollobás, Almost all regular graphs are Hamiltonian, *European Journal on Combinatorics* **4** (1983) 97-106.

[8] B. Bollobás, C. Cooper, T. Fenner and A.M. Frieze, On Hamilton cycles in sparse random graphs with minimum degree at least k, *Journal of Graph Theory* **34** (2000) 42-59.

[9] B. Bollobás, T. Fenner and A.M. Frieze, Hamilton cycles in random graphs with minimal degree at least k, *in A tribute to Paul Erdos, edited by A.Baker, B.Bollobas and A.Hajnal* (1990) 59-96.

[10] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Hungar. Acad. Sci.*, **5**, (1960) 17–61.

[11] T. Fenner and A.M. Frieze, Hamiltonian cycles in random regular graphs, *Journal of Combinatorial Theory B* **40** (1984) 103-112.

[12] W. Fernandez de la Vega, Long paths in sparse random graphs, *Studia Sci. Math. Hungar.* **14** (14) (1979) 335-340.

[13] A.M. Frieze, Finding hamilton cycles in sparse random graphs, *Journal of Combinatorial Theory B*, **44** (1988) 230-250.

[14] A.M. Frieze, On Hamilton Cycles in Random Graphs with Minimum Degree at Least Three.

[15] A.M. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, *Trends in Mathematics, Birkhauser Verlag, Basel*, (2004) 95-132.

[16] R.M. Karp and M. Sipser, Maximum matchings in sparse random graphs, *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computing* (1981) 364-375.

[17] J. Komlós and E. Szemerédi, Hamilton cycles in random graphs, *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdos on his 60th birthday)*, (1975) 1003-1010.

[18] A. Korshunov, A solution of a problem of P. Erdős and A. Rényi about Hamilton cycles in non-oriented graphs, *Metody Diskr. Anal. Teoriy Upr. Syst. Sh. Trudov.* **31** (1977) 17-56 (in Russian).

[19] J. Komlós and E. Szemerédi, Limit distribution for the existence of Hamiltonian cycles in a random graph, *Discrete Mathematics* **43** (1983) 55-63.

[20] B. Pittel, On tree census and the giant component in sparse random graphs, *Random Structures and Algorithms* **1** (1990) 311-342.

[21] B. Pittel and N. C. Wormald, Counting connected graphs inside-out, *J. Comb. Theory, Ser. B*, **93**, (2005) 122–172.

[22] L. Pósa, Hamiltonian circuits in random graphs, *Discrete Mathematics* **14** 359-364.

[23] R. Robinson and N. Wormald, Almost all regular graphs are Hamiltonian, *Random Structures and Algorithms* **5**m (1994) 363-374.