Hamilton cycles in random graphs with minimum degree at least 3: an improved analysis

Michael Anastos and Alan Frieze*
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA15213
U.S.A.

October 2, 2020

Abstract

In this paper we consider the existence of Hamilton cycles in the random graph $G = G^{\delta\geq 3}_{n,m}$. This random graph is chosen uniformly from $\mathcal{G}^{\delta\geq 3}_{n,m}$, the set of graphs with vertex set $[n]$, m edges and minimum degree at least 3. Our ultimate goal is to prove that if $m = cn$ and $c > 3/2$ is constant then G is Hamiltonian w.h.p. In an earlier paper [4], the second author showed that $c \geq 10$ is sufficient for this and in this paper we reduce the lower bound to $c > 2.662...$. This new lower bound is the same lower bound found in Frieze and Pittel [6] for the expansion of so-called Pósa sets.

1 Introduction

In this paper we consider the existence of Hamilton cycles in the random graph $G = G^{\delta\geq 3}_{n,m}$. This random graph is chosen uniformly from $\mathcal{G}^{\delta\geq 3}_{n,m}$, the set of graphs with vertex set $[n]$, m edges and minimum degree at least 3. If $c = 3/2$ then $G^{\delta\geq 3}_{n,m}$ is precisely the random 3-regular graph which is proven, via the small cycle conditioning method, to be Hamiltonian [11]. However as $G^{\delta\geq 3}_{n,3n/2} \not\subset G^{\delta\geq 3}_{n,cn}$ for every $c > 0$ we cannot directly infer Hamiltonicity for larger values of c. In addition, due to the increase in the variance of the degree sequence, the method itself cannot be transfered directly. Our ultimate goal is to prove that if $m = cn$ and $c > 3/2$ is constant then G is Hamiltonian w.h.p. In an earlier paper [4], the second author showed that $c \geq 10$ is sufficient for this and in this paper we reduce the lower bound

*Research supported in part by NSF Grant DMS1363136

to $c > 2.662...$. This new lower bound is the same lower bound found in Frieze and Pittel [6] for expansion of so-called Posá sets i.e. sets of endpoints that may be formed via the application of Pósa rotations, Pósa [10]. In summary we prove,

Theorem 1.1. *W.h.p. $G_{n,m}^{\delta\geq 3}$ is Hamiltonian for $m = cn, c > 2.662...$.*

One of the motivations for studying this problem arises from the fact that the 3-core of the random graph $G_{n,m}$ is distributed precisely as $G_{\nu,\mu}^{\delta\geq 3}$, where ν, μ are the (random) number of vertices and edges in the 3-core and w.h.p. ν is known to be linear in n. In particular, it is plausible that the first non-empty 3-core in the random graph process is Hamiltonian w.h.p. To prove this to be true, we would need to reduce the lower bound on c to the edges to vertices ratio of the corresponding 3-core which is known to be w.h.p. about 1.8 [7]. In addition, we note that Krivelevich, Lubetzky and Sudakov [8] showed that w.h.p. the first non-empty k-core, $k \geq 15$, is Hamiltonian.

2 Proof of Theorem 1.1

2.1 The game plan

The key to the proof Theorem 1.1 is the following lemma:

Lemma 2.1. *Let $V = [n]$ and $G = (V, E)$ where $E = E_1 \cup E_2$ and $E_2 = \{e_1, ..., e_a\} \subset \binom{V}{2} \setminus E_1$. Let $G_1 = (V, E_1)$ and let $\mathcal{P}$ be a set of vertex disjoint paths in G_1 that covers V. Suppose that for some $0 < \beta < 1$,*

P1 $|\mathcal{P}| \leq \min\left\{\frac{|E_2|}{n^{2-2\beta}\log^2 n}, \frac{n^\beta}{4}\right\}$.

P2 *Given $e_1, e_2, \ldots, e_{i-1}$, the edge e_i is chosen uniformly from $\binom{V}{2} \setminus \left(E_1 \cup \{e_1, \ldots, e_{i-1}\}\right)$.*

P3 *$X \subseteq V$, $|X| \leq n^\beta$ implies that either $e(X \cup N(X)) \leq |X \cup N(X)|$ or $|N(X)| \geq 2|X|$. (Here $N(X) = \{y \in V \setminus X : \exists x \in X$ such that $\{x, y\} \in E_1\}$. In addition $e(X \cup N(X))$ denotes the number of edges spanned by $X \cup N(X)$.)*

Then G is Hamiltonian with probability $1 - o(n^{-3})$.

We later apply Lemma 2.1 with $\beta = 0.99$ and $|E_2| = n^{1/2-o(1)}$.

Proof. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_\ell\}$ be a minimum cardinality set of vertex disjoint paths in G_1 that covers V (and satisfies **P1**). Let the endpoints of P_i be $v_{(i,1)}$ and $v_{(i,2)}$ for $i \in [\ell]$. Because $\mathcal{P}$ is of minimum cardinality we have that $\{v_{(i,2)}v_{(i+1,1)}\} \notin E_1$ for $i \in [\ell]$ (here we identify $v_{(\ell+1,1)}$ with $v_{(1,1)}$). In addition, $H_0 = v_{(1,1)}, P_1, v_{(1,2)}v_{(2,1)}, P_2, v_{(2,2)}v_{(3,1)}P_3, \ldots, v_{(\ell,1)}P_\ell$

$v_{(\ell,2)}v_{(1,1)}, v_{(1,1)}$ is a Hamilton cycle in the graph $\Gamma_0 = (V, E_1 \cup R)$ where $R = \{\{v_{(i,2)}v_{(i+1,1)}\} : i \in [\ell]\}$.

Starting with H_0, we find a Hamilton cycle in G by removing the edges of R from our cycle. We do this with at most ℓ rounds of an extension-rotation procedure. Fix $i \geq 0$ and suppose then that after i rounds, we have a Hamilton cycle H_i in the graph $\Gamma_i = (V, E_1 \cup R_i \cup F_i)$ where $R_i \subseteq R$ and $|R_i| \leq \ell - i$. Here $F_i = \{e_1, e_2, \dots, e_b\}$ are the edges of E_2 that have been *revealed* so far. We explain revealed momentarily.

We start round $i+1$ by deleting an edge e from R_i to create a Hamilton path Q_1. We then use Pósa rotations to try to find a Hamilton cycle in $\Gamma_i - e$. Given a path $P = (x_1, x_2, \dots, x_s)$ and an edge $\{x_s, x_j\}$ where $1 < j < s-1$, the path $(x_1, \dots, x_j, x_s, x_{s-1}, \dots, x_{j+1})$ is said to be obtained from P by a rotation with x_1 as the fixed end vertex. The edge $\{x_s, x_j\}$ will be called the rotating edge.

First consider all Hamilton paths obtainable from Q_1 by a sequence of rotations with x_1 fixed. In these rotations, we are only allowed to use edges from $E(\Gamma_i) \setminus \{e\}$ as rotating edges. Next let $END(Q_1, x_1)$ denote the set of end vertices of these paths, other than x_1. If there exists $y \in END(Q_1, x_1)$ such that $\{x_1, y\} \in E(\Gamma_i) \setminus \{e\}$ then this round is complete. We have a Hamilton cycle containing one less member of R. Thus we can define $R_{i+1} = R_i \setminus \{e\}$ and $F_{i+1} = F_i$.

In the event there is no such y, we proceed as follows: Let $END(Q_1, x_1) = \{z_1, z_2, \dots, z_q\}$ and let $Q_j, j = 2, \dots, q$ denote a path from x_1 to z_j found by rotations. Then, for $1 \leq j \leq q$, we let $END(Q_j, z_j)$ denote the set of end vertices of paths obtainable from Q_j by a sequence of rotations with z_j fixed. If for some j we find $y \in END(Q_j, z_j)$ such that $\{z_j, y\} \in E(\Gamma_i) \setminus \{e\}$ then, as before, this round is complete. We have a Hamilton cycle containing one less member of R. We can then define $R_{i+1} = R_i \setminus \{e\}$ and $F_{i+1} = F_i$.

Failing this, we start revealing the edges of $e_{b+1}, e_{b+2}, ..., e_a$, in this order, to search for an edge of the form $\{z_j, y_j\}$ where $y_j \in END(Q_j, z_j)$. If e_c is the first such edge, $b \leq c \leq a$, then we let $R_{i+1} = R_i \setminus \{e\}$, $F_{i+1} = F_i \cup \{e_{b+1}, e_{b+2}, \dots, e_c\}$, $\Gamma_{i+1} = (V, E_1 \cup R_{i+1} \cup F_{i+1})$ and H_{i+1} be a Hamilton cycle in Γ_{i+1}. Pósa's lemma states that $|N(END(Q_j, z_j))| < 2|END(Q_j, z_j)|$ (see Corollary 6.7 of [5]) and Lemma 2.1 of [6] that $e\big(N(END(Q_j, z_j)) \cup END(Q_j, z_j)\big) > |N(END(Q_j, z_j)) \cup END(Q_j, z_j)|$. Thus, **P3** implies that $|END(Q_j, z_j)| > n^\beta$ for all $1 \leq j \leq q$ and similarly that $q > n^\beta$.

For $1 \leq l \leq a = |E_2|$ let Y_l be the indicator for the event that either e_l is not revealed (in any round) in the above procedure or when it is revealed a new Hamilton cycle is identified. From **P2**, we have,

$$\mathbf{Pr}(Y_j = 1) \geq \frac{\binom{n^\beta - 2j}{2}}{\binom{n}{2}} \geq \frac{n^{2\beta-2}}{5},$$

for $j \leq n^\beta/4$.

In the event that G is not Hamiltonian all the edges in E_2 are revealed and for less than $|\mathcal{P}|$ of them a new Hamilton cycle is identified. Indeed, if we assume otherwise then $\Gamma_{|\mathcal{P}|} \subseteq \Gamma$ is

Hamiltonian. Hence, $Z \leq |\mathcal{P}|$. But $Y_l, 1 \leq l \leq a$ dominates a $Bernoulli(n^{2\beta-2}/5)$ random variable. This domination holds regardless of $Y_1, Y_2, \ldots, Y_{l-1}$. Hence, from **P1**, we have

$$\mathbf{Pr}(G \text{ is not Hamiltonian }) \leq \mathbf{Pr}(Binomial(n^{2-2\beta}|\mathcal{P}|\log^2 n, n^{2\beta-2}/5) \leq |\mathcal{P}|) = o(n^{-3}).$$

□

2.2 Choice of E_2

Let

$$s = n^{1/2}\log^{-2} n$$

and let

$$\Omega = \left\{(H,Y) : H \in \mathcal{G}^{\delta\geq 3}_{n,cn-s}, Y \subseteq \binom{[n]}{2}, |Y| = s \text{ and } E(H) \cap Y = \emptyset\right\}$$

where $\mathcal{G}^{\delta\geq 3}_{n,m} = \left\{G^{\delta\geq 3}_{n,m}\right\}$.

We consider two ways of randomly choosing an element of Ω.

(a) First choose G uniformly from $\mathcal{G}^{\delta\geq 3}_{n,cn}$ and then choose an s-set X uniformly from $E(G) \setminus E_3(G)$, where $E_3(G)$ is the set of edges of G that are incident with a vertex of degree 3. This produces a pair $(G - X, X)$. We let $\mathbf{Pr}_a$ denote the induced probability measure on Ω.

(b) Choose H uniformly from $\mathcal{G}^{\delta\geq 3}_{n,cn-s}$ and then choose an s-set Y uniformly from $\binom{[n]}{2} \setminus E(H)$. This produces a pair (H, Y). We let $\mathbf{Pr}_b$ denote the induced probability measure on Ω.

The following lemma implies that as far as properties that happen **whp** in G, we can use Method (b), just as well as Method (a) to generate our pair (H, Y). For a proof see Lemma 10.1 of [4].

Lemma 2.2. *There exists $\Omega_1 \subseteq \Omega$ such that*

(i) $\mathbf{Pr}_a(\Omega_1) = 1 - o(1)$.

(ii) $\omega = (H, Y) \in \Omega_1$ *implies that* $\mathbf{Pr}_a(\omega) = (1 + o(1))\,\mathbf{Pr}_b(\omega)$.

It follows that we can take E_2 as the set Y in the lemma and then we have $|E_2| = n^{0.5-o(1)}$ and this covers **P2** of Lemma 2.1.

2.3 P3 of Lemma 2.1

The main result of [6], (see Theorem 1.1 of that paper), is that if $m = cn$ and $c > 2.6616\ldots$ then w.h.p. if $e(S \cup N(S)) > |S \cup N(S)|$ then $|S| + |N(S)| \geq n^{1-o(1)}$. So, we see that we can take $\beta = 0.99$ in Lemma 2.1. This covers **P3**.

In [6] it is also shown that if G has minimum degree 3, P is a path of G and x an endpoint of P then the set $S = END(P, x)$, defined in the proof of Lemma 2.1, satisfies the relation $e(S \cup N(S)) > |S \cup N(S)|$.

P1 of Lemma 2.1 will follow from Lemma 5.2. The corresponding set of paths $\mathcal{P}$ satisfies $|\mathcal{P}| \leq n^{0.41} \leq \min\left\{\frac{n^{0.5-o(1)}}{n^{0.02}\log^2 n}, \frac{n^{0.99}}{4}\right\} = \min\left\{\frac{|E_2|}{n^{2-2\beta}}, \frac{n^\beta}{4}\right\}$.

3 Random Sequence Model

We must now take some time to explain the model we use for $G_{n,m}^{\delta\geq 3}$. We use a variation on the pseudo-graph model of Bollobás and Frieze [2] and Chvátal [3]. Given a sequence $\mathbf{x} = (x_1, x_2, \ldots, x_{2M}) \in [N]^{2M}$ of $2M$ integers between 1 and N we can define a (multi)-graph $G_\mathbf{x} = G_\mathbf{x}(N, M)$ with vertex set $[N]$ and edge set $\{(x_{2i-1}, x_{2i}) : 1 \leq i \leq M\}$. The degree $d_\mathbf{x}(v)$ of $v \in [N]$ is given by

$$d_\mathbf{x}(v) = |\,\{j \in [2M] : x_j = v\}\,|.$$

If $\mathbf{x}$ is chosen randomly from $[N]^{2M}$ then $G_\mathbf{x}$ is close in distribution to $G_{N,M}$. Indeed, conditional on being simple, $G_\mathbf{x}$ is distributed as $G_{N,M}$. To see this, note that if $G_\mathbf{x}$ is simple then it has vertex set $[N]$ and M edges. Also, there are $M!2^M$ distinct equally likely values of $\mathbf{x}$ which yield the same graph.

We will use the above variation on the pseudo-graph model to analyze 2GREEDY, an algorithm that finds 2-matchings, applied to $G_{n,m}^{\delta\geq 3}$. A 2-matching is a set of edges such that every vertex is incident to at most 2 edges in it. 2GREEDY is described in Section 4. As 2GREEDY progresses vertices become matched (incident with edges selected for the 2-matching), edges are deleted and vertices of small degree are identified. As such we will need to impose additional constrains on the vertex degrees and our situation becomes more complicated. At any step of the algorithm we keep track of 3 sets J_3, J_2 and J_0 that partition the current vertex set, say $[N]$. (A vertex that becomes incident with 2 edges of the 2-matching is not included in the current vertex set.) J_3 is a set of vertices of degree at least 3 and it consists of vertices that have not been matched yet. J_2 is a set of vertices of degree at least 2 and it consists of vertices that are incident to exactly 1 edge in the current 2-matching. Finally J_0 consists of the remaining vertices and whose sum of degrees will be proven to be $D = o(N)$.

So we let

$$[N]_{J_2,J_3;D}^{2M} = \{\mathbf{x} \in [N]^{2M} : d_\mathbf{x}(j) \geq i \text{ for } j \in J_i,\ i = 2, 3 \text{ and } \sum_{j\in J_0} d_\mathbf{x}(j) = D\}.$$

Let $G = G(N, M, J_2, J_3; D)$ be the multi-graph $G_{\mathbf{x}}$ for $\mathbf{x}$ chosen uniformly from $[N]^{2M}_{J_2,J_3;D}$. What we need now is a procedure that generates $G_{\mathbf{x}}$ conditioned on $G_{\mathbf{x}}$ being simple or equivalently a way to access the degree sequence of elements in $[N]^{2M}_{J_2,J_3;D}$. Such a procedure is given in [4] and it is justified by Lemmas 3.1, 3.2 and 3.3 that follow. In Lemma 3.1 it is proven that the degree sequence of $[N]^{2M}_{J_2,J_3;D}$ (restricted to the sets J_2, J_3) has the same distribution as the joint distribution of $\mathcal{P}_1, \mathcal{P}_2, ..., \mathcal{P}_{|J_2|+|J_3|}$ where (i) for $i \in J_\ell$, $\mathcal{P}_i$ is a $Poisson(\lambda)$ random variable condition on being at least ℓ for some carefully chosen value of λ and (ii) $\sum_{i=1}^{|J_2|+|J_3|} \mathcal{P}_i = 2M - D$. In Lemma 3.2 it is shown that the marginal of $d_{\mathbf{x}}(j)$ and joint of $(d_{\mathbf{x}}(j_1), d_{\mathbf{x}}(j_2))$ distributions are close to the marginal of $\mathcal{P}_i$ and joint of $(\mathcal{P}_i, \mathcal{P}_j)$ distributions respectively. This fact is used in Lemma 3.3 where we establish concentration of the number of vertices of degree k in $J_\ell, \ell = 2, 3$. For the proofs of Lemmas 3.1, 3.2 and 3.3 see [4].

Let
$$f_k(\lambda) = e^\lambda - \sum_{i=0}^{k-1} \frac{\lambda^i}{i!}$$
for $k \geq 0$.

Lemma 3.1. *Let $\mathbf{x}$ be chosen randomly from $[N]^{2M}_{J_2,J_3;D}$. For $i = 2, 3$ let Z_j $(j \in [J_i])$ be independent copies of a* truncated Poisson *random variable $\mathcal{P}_i$, where*
$$\mathbf{Pr}(\mathcal{P}_i = t) = \frac{\lambda^t}{t! f_i(\lambda)}, \qquad t = i, i+1, \dots\ .$$
Here λ satisfies
$$\sum_{i=2}^{3} \frac{\lambda f_{i-1}(\lambda)}{f_i(\lambda)} |J_i| = 2M - D. \tag{1}$$
For $j \in J_0$, $Z_j = d_j$ is a constant and $\sum_{j \in J_0} d_j = D$. Then $\{d_{\mathbf{x}}(j)\}_{j \in [N]}$ is distributed as $\{Z_j\}_{j \in [N]}$ conditional on $Z = \sum_{j \in [n]} Z_j = 2M$.

To use Lemma 3.1 for the approximation of vertex degrees distributions we need to have sharp estimates of the probability that Z is close to its mean $2M$. In particular we need sharp estimates of $\mathbf{Pr}(Z = 2M)$ and $\mathbf{Pr}(Z - Z_1 = 2M - k)$, for $k = o(N)$. These estimates are possible precisely because $\mathbf{E}(Z) = 2M$. Using the special properties of Z, a standard argument in an appendix of [4] shows that where $N_\ell = |J_\ell|$ and $N^* = N_2 + N_3$ and the variances are
$$\sigma_\ell^2 = \frac{f_\ell(\lambda)(\lambda^2 f_{\ell-2}(\lambda) + \lambda f_{\ell-1}(\lambda)) - \lambda^2 f_{\ell-1}(\lambda)^2}{f_\ell(\lambda)^2} \text{ and } \sigma^2 = \frac{1}{N^*} \sum_{\ell=2}^{3} N_\ell \sigma_\ell^2, \tag{2}$$
that if $N^* \sigma^2 \to \infty$ and $k = O(\sqrt{N^*}\sigma)$ then
$$\mathbf{Pr}\left(Z = 2M - k\right) = \frac{1}{\sigma\sqrt{2\pi N^*}} \left(1 + O\left(\frac{k^2+1}{N^*\sigma^2}\right)\right). \tag{3}$$

Given (3) and

$$\sigma_\ell^2 = O(\lambda), \qquad \ell = 2, 3,$$

we obtain

Lemma 3.2. *Let* $\mathbf{x}$ *be chosen randomly from* $[N]^{2M}_{J_2,J_3;D}$.

(a) *Assume that* $\log N^* = O((N^*\lambda)^{1/2})$. *For every* $j \in J_\ell$ *and* $\ell \le k \le \log N^*$,

$$\mathbf{Pr}(d_{\mathbf{x}}(j) = k) = \frac{\lambda^k}{k! f_\ell(\lambda)} \left(1 + O\left(\frac{k^2+1}{N^*\lambda}\right)\right). \tag{4}$$

Furthermore, for all $\ell_1, \ell_2 \in \{2, 3\}$ *and* $j_1 \in J_{\ell_1}, j_2 \in J_{\ell_2}$, $j_1 \neq j_2$, *and* $\ell_i \le k_i \le \log N^*$,

$$\mathbf{Pr}(d_{\mathbf{x}}(j_1) = k_1, d_{\mathbf{x}}(j_2) = k_2) = \frac{\lambda^{k_1}}{k_1! f_{\ell_1}(\lambda)} \frac{\lambda^{k_2}}{k_2! f_{\ell_2}(\lambda)} \left(1 + O\left(\frac{\log^2 N^*}{N^*\lambda}\right)\right). \tag{5}$$

(b)

$$d_{\mathbf{x}}(j) \le \frac{\log N}{(\log\log N)^{1/2}} \quad \mathbf{qs}^1 \tag{6}$$

for all $j \in J_2 \cup J_3$.

Let $\nu_{\mathbf{x}}^\ell(s)$ denote the number of vertices in $J_\ell, \ell = 2, 3$ of degree s in $G_{\mathbf{x}}$. Equation (3) and a standard tail estimate for the binomial distribution shows the following:

Lemma 3.3. *Suppose that* $\log N^* = O((N^*\lambda)^{1/2})$ *and* $N_\ell \to \infty$ *with* N. *Let* $\mathbf{x}$ *be chosen randomly from* $[N]^{2M}_{J_2,J_3;D}$. *Then* **qs**,

$$\mathcal{D}(\mathbf{x}) = \left\{\left|\nu_{\mathbf{x}}^\ell(j) - \frac{N_\ell \lambda^j}{j! f(\lambda)}\right| \le \left(1 + \left(\frac{N_\ell \lambda^j}{j! f(\lambda)}\right)^{1/2}\right) \log^2 N,\ k \le j \le \log N\right\}. \tag{7}$$

□

We can now show $G_{\mathbf{x}}$, $\mathbf{x} \in [n]^{2m}_{\emptyset,[n];0}$ is a good model for $G^{\delta \ge 3}_{n,m}$. For this we only need to show now that

$$\mathbf{Pr}(G_{\mathbf{x}} \text{ is simple}) = \Omega(1). \tag{8}$$

For this we can use a result of McKay [9]. If we fix the degree sequence of $\mathbf{x}$ then $\mathbf{x}$ itself is just a random permutation of the multi-graph in which each $j \in [n]$ appears $d_{\mathbf{x}}(j)$ times. This in fact is another way of looking at the configuration model of Bollobás [1]. The reference [9] shows that the probability $G_{\mathbf{x}}$ is simple is asymptotically equal to $e^{-(1+o(1))\rho(\rho+1)}$ where

[1] An event $\mathcal{E} = \mathcal{E}(N^*)$ occurs quite surely (**qs**, in short) if $\mathbf{Pr}(\mathcal{E}) = 1 - O(N^{-a})$ for any constant $a > 0$

$\rho = m_2/m$ and $m_2 = \sum_{j\in[n]} d_{\mathbf{x}}(j)(d_{\mathbf{x}}(j)-1)$. One consequence of the exponential tails in Lemma 3.3 is that $m_2 = O(m)$. This implies that $\rho = O(1)$ and hence that (8) holds. We can thus use the Random Sequence Model to prove the occurrence of high probability events in $G_{n,m}^{\delta\geq 3}$.

All that is left now is to find the collection of paths promised at the end of Section 2. For this we analyse algorithm 2GREEDY of [4], which is described in Section 4. Its performance is given in Lemma 5.2 in Section 5.

4 Greedy Algorithm

We now describe the algorithm 2GREEDY of [4]. Our algorithm will be applied to the random graph $G = G_{n,m}^{\delta\geq 3}$ and analyzed in the context of $G_{\mathbf{x}}$, with $N = n$ initially. As the algorithm progresses, it makes changes to G and we let Γ denote the current state of G. The algorithm grows a 2-matching M and for $v \in [n]$ we let $b(v)$ be the number of edges in M that are incident to v. We let

- $V_{0,j} = \{v \in [n] : d_\Gamma(v) = 0,\, b(v) = j\}$, $j = 0, 1$,
- $Y_k = \{v \in [n] : d_\Gamma(v) = k \text{ and } b(v) = 0\}$, $k = 1, 2$,
- $Z_1 = \{v \in [n] : d_\Gamma(v) = 1 \text{ and } b(v) = 1\}$,
- $Y = \{v \in [n] : d_\Gamma(v) \geq 3 \text{ and } b(v) = 0\}$, This is J_3 of Section 3.
- $Z = \{v \in [n] : d_\Gamma(v) \geq 2 \text{ and } b(v) = 1\}$, This is J_2 of Section 3.
- M is the set of edges in the current 2-matching.

Algorithm

Step 1 $Z_1 \cup Y_1 \cup Y_2 \neq \emptyset$
Choose a random vertex v from $Z_1 \cup Y_1 \cup Y_2$. Let w be a random neighbor of v. (We allow the case $v = w$ as we are analyzing the algorithm within the context of $G_{\mathbf{x}}$. This case is of course unnecessary when the input is simple i.e. for $G_{n,m}^{\delta\geq k}$). Add (v, w) to M and delete it from Γ. Update $b(v) = b(v) + 1$, $b(w) = b(w) + 1$. Delete all vertices in $V(\Gamma)$ satisfying $b(u) \geq 2$ and the edges incident to them. Delete any isolated vertices.

Step 2: $Y_1 \cup Y_2 \cup Z_1 = \emptyset$
Choose a random vertex v from $Z \cup Y$. Let w be a random neighbor of v. Add (v, w) to M and delete it to from Γ. Update $b(v) = b(v) + 1$, $b(w) = b(w) + 1$. Delete all vertices in $V(\Gamma)$ satisfying $b(u) \geq 2$ and the edges incident to them. Delete any isolated vertices.

The algorithm ends when no vertices left in Γ. The output of 2GREEDY is set of edges in M.

5 Analysis of 2greedy

We will use the following additional notation to that given in Section 4:

- m_i: number of edges at time i in Γ.
- $Z_j, j \geq 2$ and $Y_j, j \geq 3$ resp. are the subsets of Z and Y respectively consisting of vertices of degree j.
- $y_i = |Y|, z_i = |Z|$ at time i.
- $\zeta_i = |Y_1| + 2|Y_2| + |Z_1|$.
- $$p_{2,i} = \frac{2|Z_2|}{2m_i} \text{ and } p_{3,i} = \frac{3|Y_3|}{2m_i}.$$
- λ_i is the unique constant satisfying
$$\frac{\lambda_i f_2(\lambda_i)}{f_3(\lambda_i)}|Y| + \frac{\lambda_i f_1(\lambda_i)}{f_2(\lambda_i)}|Z| = 2m_i - \zeta_i.$$

Let $\epsilon = 10^{-9}$. We also define the stopping time

$$\tau := \min\{i : m_i \leq n^{0.4+2\varepsilon}\}.$$

Lemma 5.1. *W.h.p.*
$$\text{for } i < \tau \text{ we have } \zeta_i < n^{0.4+\varepsilon} = o(m_i). \tag{9}$$

We use Lemma 5.1 in the proof of Lemma 5.2 to bound the number of vertices of degree at most 1 in M. Its proof is given in subsection 5.1.

Lemma 5.2. *For $c > 3/2$ and $m = cn$, w.h.p* 2greedy *applied to $G = G_{n,m}^{\delta\geq 3}$ outputs a 2-matching of size at least $n - 4n^{0.4+3\varepsilon}$. In addition w.h.p. in $G_{n,m}^{\delta\geq 3}$ there exists a set of vertex disjoint paths of size at most $2n^{0.4+2\varepsilon}$ that covers $V = V(G)$.*

Proof. Every component in M defines a path and the union of the vertices of these paths is V. The number κ of components of the 2-matching M output by 2greedy can be bounded as follows. κ can be bounded by the number κ_1 of vertices of degree one or zero in M plus κ_2, the number of cycles. For every vertex $v \in V$ that contributes to κ_1 there exists a step i such that either (i) $v \in Z_1 \cup Y_1 \cup Y_2$ and at step i a neighbor of v is matched and then removed from Γ or (ii) $v \notin Z_1 \cup Y_1 \cup Y_2$, 2 neighbors of v are matched and then removed from Γ and as a result at least $d(v) - 2$ edges incident to v are removed. If the above occurs then we say that step i witnesses an increase of κ_1.

For the number of cycles spanned by M, observe that at step i, κ_2 can increase by one only if we add an edge $\{u, v\}$ to M where u is connected to v by a path in M. If the above occurs then we say that step i witnesses an increase of κ_2.

Since w.h.p the maximum degree of G, and hence of Γ during the execution of 2GREEDY, is $\log n$ we have that step i witnesses an increase of $\kappa_1 + \kappa_2$ of magnitude at most $2 \log n$ with probability at most $(2 \log n)\zeta_i/2m_i + O(1/m_i)$. If $\kappa_1 + \kappa_2$ reaches $n^{0.4+2\varepsilon}$ before time τ then, there are at least $\varepsilon^2 n^{0.4+2\varepsilon}/2 \log n$ steps with $m_i \in [n^{0.4+2\varepsilon+(r-1)\epsilon^2}, n^{0.4+2\varepsilon+r\epsilon^2}]$ for some integer $1 \le r \le 1/\varepsilon^2$ that witness an increase of $\kappa_1 + \kappa_2$. The probability that this occurs for a fixed r, while $\zeta_i \le n^{0.4+\varepsilon}$, is bounded by

$$\binom{n^{0.4+2\varepsilon+r\varepsilon^2}}{\frac{\varepsilon^2 n^{0.4+2\varepsilon}}{2\log n}} \left(\frac{2n^{0.4+\varepsilon} \log n}{n^{0.4+2\varepsilon+(r-1)\varepsilon^2}} \right)^{\frac{\varepsilon^2 n^{0.4+2\varepsilon}}{2\log n}} \le \left(\frac{en^{0.4+2\varepsilon+r\varepsilon^2}}{\frac{\varepsilon^2 n^{0.4+2\varepsilon}}{2\log n}} \cdot \frac{2n^{0.4+\varepsilon} \log n}{n^{0.4+2\varepsilon+(r-1)\varepsilon^2}} \right)^{\frac{\varepsilon^2 n^{0.4+2\varepsilon}}{2\log n}} \le n^{-5}.$$

Hence w.h.p. if $\zeta_i \le n^{0.4+\varepsilon}$ for $i < \tau$ then the total increase in $\kappa_1 + \kappa_2$ in the first $\tau - 1$ steps is bounded by $n^{0.4+2\varepsilon}$. Once $m_i \le n^{0.4+2\varepsilon}$, at most $n^{0.4+2\varepsilon}$ more components can be created, yielding in total at most $2n^{0.4+2\varepsilon}$ components.

Finally M is bounded below by $n - 2\kappa_1$ which is w.h.p. at least $n - 4n^{0.4+2\varepsilon}$. $\square$

5.1 Proof of Lemma 5.1

We study ζ_i in two regimes (captured by the event $\mathcal{A}_i$ defined below) depending on whether we can approximate $Z_j, j \ge 2$ and $Y_j, j \ge 3$ by the corresponding truncated poisson distribution as described in section 3. In both regimes, to bound ζ_i we show that it can be stochasticly dominated by either a random walk with significantly large negative drift or a lazy random walk whose drift is asymptotically 0. The first kind of behavior is associated with the study of the random variables X_i, X_i' defined below while the second one is associated with W_i and W_i'.

For $i < \tau$, we define the events

$$\mathcal{A}_i = \left\{ (z_j + y_j)\lambda_j \ge \log^3 n \text{ for } j \le i \right\} \text{ and } \mathcal{B}_i = \left\{ (\lambda_i \ge m_i^{-0.2}) \vee (y_i \ge m_i^{0.8}) \right\}.$$

For $i < \tau$, we also define the following random variables:

$$\begin{aligned}
X_i &= (\zeta_{i+1} - \zeta_i)\mathbb{I}(\mathcal{A}_i, \mathcal{B}_i, 0 < \zeta_i < n^{0.4+\varepsilon}). \\
W_i &= (\zeta_{i+1} - \zeta_i)\mathbb{I}(\mathcal{A}_i, \neg\mathcal{B}_i, 0 < \zeta_i < n^{0.4+\varepsilon}). \\
X_i' &= (\zeta_{i+1} - \zeta_i)\mathbb{I}(\neg\mathcal{A}_i, \mathcal{B}_i, 0 < \zeta_i < n^{0.4+\varepsilon}). \\
W_i' &= (\zeta_{i+1} - \zeta_i)\mathbb{I}(\neg\mathcal{A}_i, \neg\mathcal{B}_i, 0 < \zeta_i < n^{0.4+\varepsilon}).
\end{aligned}$$

For $0 < i < \tau$ we have that w.h.p.

$$\min\{\zeta_i, n^{0.4+\varepsilon}\} \le M + \sum_{j=0}^{i-1}(X_i + W_i + X_i' + W_i') \tag{10}$$

where $M = \log^2 n$ is such that the following holds: w.h.p. for every $i \ge 0$ with $\zeta_i = 0$ we have that $\zeta_{i+1} \le M$. Our bound for M is justified by the fact that the maximum degree in G is $o(\log n)$ w.h.p.

We use the inequality $i < \tau$, hence $m_i \ge n^{0.4+2\varepsilon}$, to impose that if $\zeta_i \le n^{0.4+\varepsilon}$ then almost all of the vertices belong to $Y \cup Z$. We will see from the analysis below that w.h.p.

$$m_i \ge n^{0.4+2\varepsilon} \text{ implies } \zeta_i \le n^{0.4+\varepsilon}. \tag{11}$$

Equation (80) of [4] states that if $\mathcal{H}_i$ denotes the history of the process up to the end of iteration i, assuming the event $\mathcal{A}_i$ occurs, then

$$\zeta_i > 0 \text{ implies } \mathbf{E}(\zeta_{i+1} - \zeta_i \mid \mathcal{H}_i) \le -\Omega(\min\{1, \lambda_i\}^2) + O\left(\frac{\log^2 m_i}{\lambda_i m_i}\right). \tag{12}$$

In the following cases we will assume that $i < \tau$ and $\zeta_i > 0$. The case $\zeta_i = 0$ is handled by M of (10).

Case 1: $\mathcal{A}_i \wedge \mathcal{B}_i$
Case 1a
If $\lambda_i \ge m_i^{-0.2}$ we have from (12) that

$$\mathbf{E}(X_i|\mathcal{H}_i) \le -c\lambda_i^2 \le -c_1 n^{-0.4}$$

for some constant $c_1 > 0$.

Case 1b:
Assume now that $\lambda_i \le m_i^{-0.2}$. In this case since $\mathcal{A}_i$ occurs we have that for $i \ge 2$, $|Z_i|$ is approximately equal to the sum of $|Z_i|$ independent random variables that follow Poisson(λ_i) conditioned on having value at least 2. More precisely, it follows from Lemma 3.3 of [4] that as long as $\mathcal{A}_i$ holds, we have

$$\begin{aligned} \frac{|Z_3|}{|Z_2|} &= \frac{\lambda_i}{3}\left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right), \\ \frac{|Z_4|}{|Z_2|} &= \frac{\lambda_i^2}{12}\left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right), \\ \sum_{i\ge 5} |Z_i| &\le |Z_2|\lambda_i^3. \end{aligned} \tag{13}$$

Similarly

$$\begin{aligned}
\frac{|Y_4|}{|Y_3|} &= \frac{\lambda_i}{4}\left(1+O(m_i^{1/2}\lambda_i\log^2 m_i)\right),\\
\frac{|Y_5|}{|Y_3|} &= \frac{\lambda_i^2}{20}\left(1+O(m_i^{1/2}\lambda_i\log^2 m_i)\right),\\
\sum_{i\geq 6}|Y_i| &\leq |Y_3|\lambda_i^3.
\end{aligned} \tag{14}$$

Recall that if $\zeta_i > 0$ then the algorithm will choose a vertex $v \in Z_1 \cup Y_1 \cup Y_2$ and it will match it to some vertex w. Thus initially ζ_i will decrease by 1.

For $w \in Z$ let $d(w, Y_3)$ and $d(w, Z_2)$ be the number of neighbors of w in Y_3 and $Z_2 \setminus \{v\}$. Also let $f(w)$ be the number of vertices that are connected to w by multiple edges. We consider the following cases:

Case a: $w \in Y_2 \cup Y_1 \cup Z_1$ then $\zeta_{i+1} - \zeta_i = -2$.
Case b: $w \in Y$ then $\zeta_{i+1} - \zeta_i = -1$.
Case c: $w \in Z_2$ and $d(w, Z_2) = 1$ then $\zeta_{i+1} - \zeta_i = 0$.
Case d: $w \in Z_2$ and $d(w, Y_3) = 1$ then $\zeta_{i+1} - \zeta_i = 1$.
Case e: $w \in Z_2$ and $d(w, Z_2) + d(w, Y_3) = 0$ then $\zeta_{i+1} - \zeta_i = -1$.
Case f: $w \in Z \setminus Z_2$ then $\zeta_{i+1} - \zeta_i \leq -1 + d(w, Z_2) + 2d(w, Y_3) + O(f(w))$.

Differentiating cases c,d,e,f will be helpful later when we bound $\sum_{i\geq 0} Y_i$.

Summarizing we have,

$$\zeta_{i+1} - \zeta_i \begin{cases} = -2, & \text{Case a: probability } (\zeta_i/2m_i)(1+O(m_i^{-1})). \\ = -1, & \text{Case b: probability } p_{3,i}(1+O(m_i^{-1})). \\ = 0, & \text{Case c: probability } p_{2,i}^2(1+O(m_i^{-1})). \\ = +1 & \text{Case d: probability } p_{2,i}p_{3,i}(1+O(m_i^{-1}). \\ = -1 & \text{Case e: probability } p_{2,i}(1-p_{2,i}-p_{3,i})(1+O(m_i^{-1})). \\ \leq -1 + d(w, Z_2) & \\ \ +2d(w, Y_3) + O(f(w)) & \text{Case f: otherwise} \end{cases} \tag{15}$$

The net contribution of Cases c,d,e to $\mathbf{E}(X_i|\mathcal{H}_i)$ is

$$-p_{2,i} + p_{2,i}(p_{2,i} + 2p_{3,i}) = -\mathbf{Pr}(w \in Z_2) + p_{2,i}(p_{2,i} + 2p_{3,i}). \tag{16}$$

Similarly, the contribution of Case f to $\mathbf{E}(X_i|\mathcal{H}_i)$ is at most

$$\begin{aligned}
&\mathbf{E}\big[[-1 + d(w, Z_2) + 2d(w, Y_3)) + O(f(w))]\mathbb{I}(w \in Z \setminus Z_2)|\mathcal{H}_i\big]\\
&= -\mathbf{Pr}(w \in Z \setminus Z_2) + 2(p_{2,i} + 2p_{3,i}) \times \frac{3|Z_3|}{2m_i} + 3(p_{2,i} + 2p_{3,i}) \times \frac{4|Z_4|}{2m_i}\\
&+ O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right)
\end{aligned}$$

$$= -\mathbf{Pr}(w \in Z \setminus Z_2) + p_{2,i}\left(\lambda_i + \frac{\lambda_i^2}{2}\right)(p_{2,i} + 2p_{3,i}) + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right). \qquad (17)$$

For the second line of the above calculation recall that w belongs to Z_j, $j \geq 3$, with probability $j|Z_j|/2m_i$. Thereafter out of its $j-1$ neighbors (other than v), $p_{2,i}$ (and $p_{3,i}$ resp.) fraction of them belong to Z_2 (Y_3 resp.). The terms associated with Z_j, $j \geq 5$, have been absorbed by the $O(\lambda_i^3)$ error term. To derive the last equality we used (13).

Finally observe that (13), (14) imply that

$$\begin{aligned}
1 &= \frac{2|Z_2| + 3|Z_3| + 4|Z_4|}{2m_i} + \frac{3|Y_3| + 4|Y_4| + 5|Y_5|}{2m_i} + \frac{\zeta_i}{2m_i} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right) \\
&= p_{2,i}\left(1 + \frac{\lambda_i}{2} + \frac{\lambda_i^2}{6}\right) + p_{3,i}\left(1 + \frac{\lambda_i}{3} + \frac{\lambda_i^2}{12}\right) + \frac{\zeta_i}{2m_i} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right). \qquad (18)
\end{aligned}$$

Therefore (15), (17) and $\mathbf{Pr}(w \in Z \setminus Z_2) \leq 1 - p_{2,i} - p_{3,i} - \frac{\zeta_i}{2m_i}$ give,

$$\begin{aligned}
\mathbf{E}(X_i|\mathcal{H}_i) &\leq \left(-\frac{2\zeta_i}{2m_i} - p_{3,i} + p_{2,i}p_{3,i} - p_{2,i}(1 - p_{2,i} + 2p_{3,i})]\right)\left(1 + O\left(\frac{1}{m_i}\right)\right) \\
&\quad + \left(-\left(1 - p_{2,i} - p_{3,i} - \frac{\zeta_i}{2m_i}\right) + p_{2,i}\left(\lambda_i + \frac{\lambda_i^2}{2}\right)(p_{2,i} + 2p_{3,i})\right) \\
&\quad + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right) \\
&= -1 - \frac{\zeta_i}{2m_i} + p_{2,i}\left(1 + \lambda_i + \frac{\lambda_i^2}{2}\right)(p_{2,i} + 2p_{3,i}) + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right).
\end{aligned}$$

In the last equality we used that $\mathcal{A}_i$ implies that $\frac{1}{m_i} \ll \frac{\lambda_i \log^2 m_i}{m_i^{1/2}}$. We now use (18) to replace -1 by the squared expression to obtain

$$\begin{aligned}
&\leq -\left[p_{2,i}\left(1 + \frac{\lambda_i}{2} + \frac{\lambda_i^2}{6}\right) + p_{3,i}\left(1 + \frac{\lambda_i}{3} + \frac{\lambda_i^2}{12}\right) + \frac{\zeta_i}{2m_i}\right]^2 \\
&\quad + p_{2,i}\left(1 + \lambda_i + \frac{\lambda_i^2}{2}\right)(p_{2,i} + 2p_{3,i}) - \frac{\zeta_i}{2m_i} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right) \\
&\leq -\frac{\lambda_i^2 p_{2,i}^2}{12} + 2p_{2,i}p_{3,i}\left(\frac{\lambda_i}{6} + \frac{\lambda_i^2}{12}\right) - p_{3,i}^2\left(1 + \frac{2\lambda_i}{3} + \frac{5\lambda_i^2}{18}\right) + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right) \\
&= -\left(\frac{\lambda_i p_{2,i}}{4} - p_{3,i}\left(\frac{2}{3} + \frac{\lambda_i}{3}\right)\right)^2 - \frac{\lambda_i^2 p_{2,i}^2}{48} - p_{3,i}^2\left(\frac{5}{9} + \frac{2\lambda_i}{9} + \frac{\lambda_i^2}{6}\right) \\
&\quad + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right) \\
&\leq -\frac{\lambda_i^2 p_{2,i}^2}{48} - \frac{5p_{3,i}^2}{9} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right). \qquad (19)
\end{aligned}$$

In Case 1b we have that the events $\mathcal{A}_i \wedge \mathcal{B}_i$ and $\lambda_i \leq m_i^{-0.2}$ occur. In addition $i < \tau$, hence $m_i \geq n^{0.4+2\varepsilon}$. $\mathcal{A}_i \wedge \mathcal{B}_i$ and $\lambda_i \leq m_i^{-0.2}$ imply that $y_i \geq m_i^{0.8}$ and so $p_{3,i} + p_{2,i} = \Omega(1)$ and $p_{3,i} \geq m_i^{-0.2}$. Therefore

$$\mathbf{E}(X_i|\mathcal{H}_i) \leq -c_2 m_i^{-0.4} \leq -c_3 n^{-0.4}$$

for some constants $c_2, c_3 > 0$.

Let $c_4 = \min\{c_1, c_3\}$. If Case 1 occurs we have by the Azuma inequality that

$$\sum_{\ell \geq 0} \mathbf{Pr}\left(\sum_{i=0}^{j} X_i \geq n^{0.4+\varepsilon/2} \right) \leq m_0 \max_{0 \leq j \leq m_0} \exp\left\{ - \frac{(n^{0.4+\varepsilon/2} + c_4 j n^{-0.4})^2}{j \log^2 n} \right\} + n^{-6} = o(1).$$

The n^{-6} term accounts for the probability that the degree of G exceeds $\log n$. The maximum degree bounds $|\zeta_{i+1} - \zeta_i|$.

Case 2: $\mathcal{A}_i \wedge \neg\mathcal{B}_i$
To bound $\sum_{i\geq 0}^{j} W_i$, let R_i be the indicator of the event that $\{\zeta_i \leq n^{0.4+\varepsilon}\}$ plus one of the cases (a),(b),(d),(e) and(f) from (15) occurs. Note that if case (c) occurs then $W_i = 0$, hence R_i is also the indicator of the event $W_i \neq 0$. Just as in Case 1, since the contribution of Case c to $\mathbf{E}(X_i|\mathcal{H}_i)$ is 0 and $W_i = 0$ if $\zeta_i \geq n^{0.4+\varepsilon}$, we have that

$$\begin{aligned} \mathbf{E}(W_i R_i|\mathcal{H}_i) &\leq -\frac{\lambda_i^2 p_{2,i}^2}{48} - \frac{5p_{3,i}^2}{9} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right) \\ &\leq -\frac{\lambda_i^2 p_{2,i}^2}{48} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right) \\ &\leq O(m_i^{-1} \log^4 m_i). \end{aligned} \tag{20}$$

For the last inequality we used that in the event $\mathcal{A}_i \wedge \neg\mathcal{B}_i$ (13), (14) and (18) imply that $p_{2,i} = 1 - o(1)$. In addition,

$$\begin{aligned} \mathbf{Pr}(R_i = 1) &\leq \mathbf{Pr}(\text{Case(a)}) + \mathbf{Pr}(\text{Case(b)}) + \mathbf{Pr}(\text{Case(d)}) + \mathbf{Pr}(\text{Case(e)}) + \mathbf{Pr}(\text{Case(f)}) \\ &= O\left(\frac{\zeta_i}{2m_i} + p_{3,i} + p_{2,i}p_{3,i} + p_{2,i}(1 - p_{3,i} - p_{2,i}) + \lambda_i \right) = O\left(\frac{\zeta_i}{2m_i} + p_{3,i} + \lambda_i \right). \end{aligned} \tag{21}$$

where we have used $1 - p_{3,i} - p_{2,i} = O(\lambda_i)$.

In the event $\neg\mathcal{B}_i$ we have that $\lambda_i \leq m^{-0.2}$ and $y_i \leq m_i^{0.8}$ and hence $p_{3,i} \leq m_i^{-0.2}$. Hence, if $\zeta_i \leq n^{0.4+\varepsilon}$ then $\mathbf{Pr}(R_i = 1) \leq m_i^{-0.2}$. Thus,

$$\begin{aligned} \sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} R_i > n^{0.8+\varepsilon/3} \right) &\leq \sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} R_i \mathbb{I}(m_i > n^{0.8}) > n^{0.8+\varepsilon/3} - n^{0.8} \right) \\ &\leq m_0 \exp\left\{ - \frac{(n^{0.8+\varepsilon/3} - n^{0.8} - \sum_{m_i=n^{0.8}}^{m_0} m_i^{-0.2})^2}{2m_0} \right\} = o(1). \end{aligned}$$

To obtain the exponential bound, we let $Z_j = \sum_{i=0}^{j} R_i \mathbb{I}(m_i > n^{0.8})$. We have $\mathbf{E} Z_j \leq \sum_{m_i=n^{0.8}}^{m_0} m_i^{-0.2} = O(n^{0.8})$ and then we can use the Chernoff bounds, since our bounds for $R_i = 1$ hold given the history of the process so far.

It follows that,

$$
\begin{aligned}
\sum_{j=0}^{m_0} \mathbf{Pr}\Bigg(\sum_{i=0}^{j} W_i \geq n^{0.4+\varepsilon/2} \Bigg) &= \sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} W_i R_i \geq n^{0.4+\varepsilon/2} \right) \\
&\leq \sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} R_i > n^{0.8+\varepsilon/3} \right) + \sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} W_i R_i \geq n^{0.4+\varepsilon/2} \,\middle|\, \sum_{i=0}^{j} R_i \leq n^{0.8+\varepsilon/3} \right) \\
&\leq o(1) + m_0 \max_{j \leq n^{0.8+\varepsilon/3}} \exp\left\{ -\frac{\left(n^{0.4+\varepsilon/2} - \sum_{m_i=0}^{m_0} m_i^{-1} \log^4 m_i\right)^2}{j \log^2 n} \right\} \\
&\leq o(1) + m_0 \max_{j \leq n^{0.8+\varepsilon/3}} \exp\left\{ -\frac{\left(n^{0.4+\varepsilon/2} - n^{o(1)}\right)^2}{j \log^2 n} \right\} = o(1). \qquad (22)
\end{aligned}
$$

To obtain the third line we use the fact that w.h.p. $|W_i| \leq \log n$, which follows from a high probability bound of $o(\log n)$ on the maximum degree of G.

Cases 3 & 4: $\neg \mathcal{A}_i$

Let $T_1 = \max\{i < \tau : \mathcal{A}_i \text{ occurs}\}$. At time T_1 we have $(z_{T_1} + y_{T_1})\lambda_{T_1} \geq m_{T_1} \log^3 n$ and hence the estimates (13), (14) hold. Thereafter $|z_{T_1+1} - z_{T_1}|, |y_{T_1+1} - y_{T_1}|, |m_{T_1+1} - m_{T_1}| = O(\Delta(G_{T_1-1}))$. The maximum degree of $\Delta(G_{T_1})$ is bounded w.h.p. by $\log n$. At time $T_1 + 1$ we have $(z_{T_1+1} + y_{T_1+1})\lambda_{T_1+1} < m_{T_1+1} \log^3 n$ hence $\lambda_{T_1} \leq \frac{2 \log^3 n}{m_{T_1}}$ and so subsequently for $i \geq T_1$ we have

$$
|Y_4|, |Z_3| = O(\log^3 n) \text{ and } Y_j = Z_{j-1} = \emptyset \text{ for } j \geq 5. \qquad (23)
$$

Case 3: $\neg \mathcal{A}_i \wedge B_i$

Given the above we replace (18) by

$$
1 = p_{2,i} + p_{3,i} + \frac{\zeta_i}{2m_i} + O\left(\frac{\log^3 n}{m_i} \right). \qquad (24)
$$

Following this we replace (19) by

$$
\mathbf{E}(X_i' \mid \mathcal{H}) \leq -\frac{5p_{3,i}^2}{9} + O\left(\frac{\log^3 n}{m_i} \right). \qquad (25)
$$

In the events $\neg \mathcal{A}_i \wedge \mathcal{B}_i$, $y_i \geq m_i^{0.8}$ and so $p_{3,i} \geq m_i^{-0.2}$. Therefore

$$
\mathbf{E}(X_i' | \mathcal{H}_i) \leq -c' m_i^{-0.4} \leq -c_5 n^{-0.4}
$$

for some constant $c_5 > 0$. Thus if Case 3 occurs we have by the Azuma inequality that

$$
\sum_{\ell \geq 0} \mathbf{Pr}\left(\sum_{i=0}^{j} X_i' \geq n^{0.4+\varepsilon/2} \right) \leq m_0 \max_{0 \leq j \leq m_0} \exp\left\{ -\frac{(n^{0.4+\varepsilon/2} + c_5 j n^{-0.4})^2}{j \log^2 n} \right\} + n^{-6} = o(1).
$$

The n^{-6} term accounts for the probability that the degree of G exceeds $\log n$. The maximum degree bounds $|\zeta_{i+1} - \zeta_i|$.

Case 4: $\neg\mathcal{A}_i \wedge \neg B_i$
As in Case 2 we have

$$\mathbf{E}(W_i' R_i|\mathcal{H}_i) \leq O(m_i^{-1} \log^4 n)$$

where R_i is defined exactly as in Case 3. Hence, just as in (22) we get

$$\sum_{j=0}^{m_0} \mathbf{Pr}\left(\sum_{i=0}^{j} W_i' \geq n^{0.4+\varepsilon/2}\right) = o(1).$$

The above analysis and equation (10) shows that w.h.p.

$$\min\{\zeta_i, n^{0.4+\varepsilon}\} \leq \log^2 n + 4n^{0.4+\varepsilon/2} < n^{0.4+0.9\varepsilon}.$$

Hence w.h.p. there does not exist $i < \tau$ such that $\zeta_i > n^{0.4+\varepsilon}$. And this therefore completes the proof of Lemma 5.1.

6 Conclusion

We have made significant progress in determining the number of random edges needed for Hamiltonicity when we condition on minimum degree at least three. Further progress will lie on improving the bound on the number of edges needed to apply Pósa's theorem that is given in [6]. This may not be so easy, as explained in Remark 4.1 of [6].

References

[1] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal on Combinatorics 1 (1980) 311-316.

[2] B.Bollobás and A.M.Frieze, On matchings and hamiltonian cycles in random graphs, *Annals of Discrete Mathematics* 28 (1985) 23-46.

[3] V. Chvátal, Almost all graphs with 1.44n edges are 3-colourable, *Random Structures & Algorithms* 2 (1991) 11-28.

[4] A.M. Frieze, On a Greedy 2-Matching Algorithm and Hamilton Cycles in Random Graphs with Minimum Degree at Least Three, *Random structures and Algorithms* 45 (2014) 443-497.

[5] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press. 2015.

[6] A.M. Frieze and B. Pittel, On a sparse random graph with minimum degree three: Likely Posa's sets are large, *Journal of Combinatorics* 4 (2013)123-156 *Random Structures and Algorithms* 43 (2013) 1-15.

[7] S. Janson and M. J. Luczak, Asymptotic normality of the k-core in random graphs, *The annals of applied probability* 18.3 (2008) 1085-1137.

[8] M. Krivelevich, E. Lubetzky and B. Sudakov, Cores of random graphs are born Hamiltonian, *Proceedings of the London Mathematical Society* 109 (2014), 161-188.

[9] B. McKay, Asymptotics for 0-1 matrices with prescribed line sums, in *Enumeration and Design*, (Academic Press, 1984) 225-238.

[10] L. Pósa, Hamiltonian circuits in random graphs, *Discrete Mathematics* 14 (1976) 359-364.

[11] R.W Robinson and N. C. Wormald, Almost all cubic graphs are Hamiltonian, *Random Structures & Algorithms* 3 (1992) 117-125.