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.

May 31, 2019

Abstract

In this paper we consider the existence of Hamilton cycles in the random graph $G = G_{n,m}^{\delta \geq 3}$. This a random graph chosen uniformly from $G_{n,m}^{\delta \geq 3}$, 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 expansion of so-called Posá sets.

1 Introduction

In this paper we consider the existence of Hamilton cycles in the random graph $G = G_{n,m}^{\delta \geq 3}$. This a random graph chosen uniformly from $G_{n,m}^{\delta \geq 3}$, 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 expansion of so-called Posá sets. In summary we prove,

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

*Research supported in part by NSF Grant DMS1363136
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 $\nu, \mu$ are the (random) number of vertices and edges in the 3-core. 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 about 1.8. In addition, we note that Krivelevich, Lubetzky and Sudakov [7] 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 $G = (V, E)$ and $E = E_1 \cup E_2$ where $E_2 = \{e_1, ..., e_a\} \subset (V_1) - E_1$. Let $G_1 = (V, E_1)$ and let $P$ be a set of vertex disjoint paths in $G_1$ that covers $V$. Suppose that for some $0 < \beta < 1$,

$P_1 \ |P| \leq \min\left\{ \frac{|E_2|}{n^{2-2\beta} \log^2 n}, \frac{n^\beta}{2 \log n} \right\}.$

$P_2 \ \text{Given} \ e_1, e_2, ..., e_i-1, e_i \ \text{is chosen uniformly from} \ \left(\binom{A_i}{2}\right) \ \text{where} \ A_i \ \text{is a set of size at least} \ n - 2i \ \log n.$

$P_3 \ X \subseteq V, \ |X| \leq n^\beta \ \text{implies that} \ |N(X)| \geq 2|X|.$

(Here $N(X) = \{y \in V \setminus X : \exists x \in X \ \text{such that} \ \{x, y\} \in E_1\}$.)

Let $G = (V, E)$, where $E = E_1 \cup E_2$. Then $G$ is Hamiltonian with probability $1 - o(n^{-3})$.

*Proof.* Let $P = \{P_1, P_2, ..., P_t\}$ be a minimum cardinality set of vertex disjoint paths in $G_1$ that covers $V$ (and satisfies $P_1$). Let the endpoints of $P_i$ be $v_{i,1}$ and $v_{i,2}$ for $i \in [t]$. Because $P$ 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, $H_0 = v_{(1,1)}, P_1, v_{(1,2)}v_{(2,1)}, P_2, v_{(2,2)}v_{(3,1)}P_3, ..., v_{(\ell,1)}P_\ell v_{(\ell,2)}v_{(1,1)}, v_{(1,1)}$ is a Hamilton cycle in the graph $G_2 = (V, E \cup R)$ 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 $\ell$ rounds of an extension-rotation procedure. Fix $i \geq 0$ and suppose 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 $F_i = \{e_{b+1}, ..., e_a\}$ where $e_1, e_2, ..., e_b$ are the edges of $E_2$ that have been used so far. We explain used momentarily.

We start round $i + 1$ by deleting an edge $e$ from $R_i$ to create a Hamilton path $Q_i$. We then use Posá rotations to try to find a Hamilton cycle in $\Gamma_i - e$. Given a path $P = (x_1, x_2, ..., x_s)$ and an edge $\{x_s, x_j\}$ where $1 < j < s - 1$, the path $(x_1, ..., x_i, x_s, x_{s-j}, ..., x_{i+1})$ 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 $F_2 = F_2(i) = (E_1 \cup R_i \cup F_1) \setminus \{e\}$ 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 $e \neq \{x_1, y\} \in F_2$ then this round is complete. We have a Hamilton cycle containing one less member of $R$.

In the event there is no such $y$, we proceed as follows: Let $END(Q_1, x_1) = \{x_2, x_3, \ldots, x_q\}$ and let $Q_j, i = 2, \ldots, q$ denote a path from $x_1$ to $x_q$ found by rotations. Then, for $2 \leq j \leq q$, we let $END(Q_j, x_j)$ denote the set of end vertices of paths obtainable from $Q_j$ by a sequence of rotations with $x_j$ fixed. If for some $j$ we find $y \in END(Q_j, x_j)$ such that $e \neq \{x_j, y\} \in F_2$ then, as before, this round is complete. We have a Hamilton cycle containing one less member of $R$.

Failing this, we will use the edges of $F_1$ to search for an edge of the form $\{x_j, y_j\}$ where $y_j \in END(Q_j, x_j)$. Posá’s lemma states that $|N(END(Q_j, x_j))| < 2|END(Q_j, x_j)|$ and so we can assume that $q > n^\beta$ and that $|END(Q_j, x_j)| > n^\beta$ for all $1 \leq j \leq q$.

For $1 \leq l \leq a = |E_2|$ let $Y_l$ be the indicator for the event that (i) we have completed all rotations for round $i$ and (ii) $e_l = \{x_j, y\} \in F_2(i)$ and (iii) $y \in END(Q_j, x_j)$. Let $Z = \sum_{l=1}^{|E_2|} Y_l$. From P3, we have,

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

for $j \leq n^\beta / 2 \log n$.

In the event that $G$ is not Hamiltonian, $Z \leq |P|$. But $Y_l, 1 \leq l \leq a$ dominates a $\text{Bernoulli}(n^{2\beta - 2}/5)$ random variable. This domination holds regardless of $Y_1, Y_2, \ldots, Y_{l-1}$. Hence, from P1, we have

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

\[\Box\]

### 2.2 Choice of $E_2$

Let

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

and let

$$\Omega = \left\{(H, Y) : H \in \mathcal{G}_{n, cm-s}, \ Y \subseteq \left(\frac{n}{2}\right), |Y| = s \text{ and } E(H) \cap Y = \emptyset\right\}$$
where $G_{n,m}^{3 \geq 3} = \{ G_{n,m}^{3 \geq 3} \}$.

We consider two ways of randomly choosing an element of $\Omega$.

(a) First choose $G$ uniformly from $G_{n,cn}^{3 \geq 3}$ 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 $\Omega$.

(b) Choose $H$ uniformly from $G_{n,cn-s}^{3 \geq 3}$ 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 $\Omega$.

The following lemma implies that as far as properties that happen $\text{whp}$ 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 $\text{P2}$.

3 P1 and P3

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. $|S| + |N(S)| \geq n^{1-o(1)}$. So, we see that we can take $\beta = 0.99$ in Lemma 2.1. This covers $\text{P3}$.

4 Random Sequence Model

We must now take some time to explain the model we use for $G_{n,m}^{3 \geq 3}$. 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 \( x \) is chosen randomly from \([N]^{2M}\) then \( G_x \) is close in distribution to \( G_{N,M} \). Indeed, conditional on being simple, \( G_x \) is distributed as \( G_{N,M} \). To see this, note that if \( G_x \) is simple then it has vertex set \([N]\) and \( M \) edges. Also, there are \( M!2^M \) distinct equally likely values of \( x \) which yield the same graph.

Our situation is complicated by there being lower bounds of 2, 3 respectively on the minimum degree in two disjoint sets \( J_2, J_3 \subseteq [N] \). Initially \( J_2 = J_3 = \emptyset \) but we will have to consider instances where they are non-empty, as our 2-matching algorithm progresses. (These sets are intrinsic to the algorithm \texttt{2GREEDY} described in the next section and a 2-matching is a graph of maximum degree at most 2.) The vertices in \( J_0 = [N] \setminus (J_2 \cup J_3) \) are of fixed bounded degree and the sum of their degrees is \( D = \sigma(N) \). So we let

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

Let \( G = G(N, M, J_2, J_3; D) \) be the multi-graph \( G_x \) for \( x \) chosen uniformly from \([N]^{2M}_{J_2,J_3;D}\). It is clear then that conditional on being simple, \( G(n, m, \emptyset, [n]; 0) \) has the same distribution as \( G_{n,m} \). It is important therefore to estimate the probability that this graph is simple. For this and other reasons, we need to have an understanding of the degree sequence \( d_x \) when \( x \) is drawn uniformly from \([N]^{2M}_{J_2,J_3;D}\). Let

\[
f_k(\lambda) = e^{\lambda} - \sum_{i=0}^{k-1} \frac{\lambda^i}{i!},
\]

for \( k \geq 0 \).

**Lemma 4.1.** Let \( x \) 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 \( P_i \), where

\[
\Pr(P_i = t) = \frac{\lambda^t}{t!f_i(\lambda)}, \quad t = i, i + 1, \ldots.
\]

Here \( \lambda \) 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 \( \{d_x(j)\}_{j \in [N]} \) is distributed as \( \{Z_j\}_{j \in [N]} \) conditional on \( Z = \sum_{j \in [n]} Z_j = 2M \).

**Proof** This is Lemma 3.1 of [4].

To use Lemma 4.1 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 \( \Pr(Z = 2M) \) and \( \Pr(Z - Z_1 = 2M - k) \), for \( k = o(N) \). These estimates are possible precisely because \( \mathbb{E}(Z) = 2M \). 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

\[
\sigma^2_\ell = \frac{f_\ell(\lambda)(\lambda^2 f_{\ell-1}(\lambda) + \lambda f_{\ell-1}(\lambda)) - \lambda^2 f_{\ell-1}(\lambda)^2}{f_\ell(\lambda)^2} \quad \text{and} \quad \sigma^2 = \frac{1}{N^*} \sum_{\ell=2}^{3} N_\ell \sigma^2_\ell, \tag{2}
\]

that if \( N^* \sigma^2 \to \infty \) and \( k = O(\sqrt{N^*}) \) then

\[
\Pr(Z = 2M - k) = 1 \sigma \sqrt{\frac{2}{\pi N^*}} \left( 1 + O \left( \frac{k^2 + 1}{N^* \sigma^2} \right) \right). \tag{3}
\]

Given (3) and \( \sigma^2_\ell = O(\lambda) \), \( \ell = 2, 3 \),

we obtain

**Lemma 4.2.** Let \( x \) 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 \( \ell \leq k \leq \log N^* \),

\[
\Pr(d_x(j) = k) = \frac{\lambda^k}{k! f_\ell(\lambda)} \left( 1 + O \left( \frac{k^2 + 1}{N^* \lambda} \right) \right). \tag{4}
\]

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 \( \ell_1 \leq k_1 \leq \log N^* \),

\[
\Pr(d_x(j_1) = k_1, d_x(j_2) = k_2) = \frac{\lambda_{k_1}}{k_1! f_{\ell_1}(\lambda)} \frac{\lambda_{k_2}}{k_2! f_{\ell_2}(\lambda)} \left( 1 + O \left( \frac{\log^2 N^*}{N^* \lambda} \right) \right). \tag{5}
\]

(b)

\[
d_x(j) \leq \frac{\log N}{(\log \log N)^{1/2}} \text{qs}^1 \tag{6}
\]

for all \( j \in J_2 \cup J_3 \).

**Proof**

This is Lemma 3.2 of [4]. \( \square \)

Let \( \nu^\ell_x(s) \) denote the number of vertices in \( J_\ell, \ell = 2, 3 \) of degree \( s \) in \( G_x \). Equation (3) and a standard tail estimate for the binomial distribution shows the following:

\[\text{An event } \mathcal{E} = \mathcal{E}(N^*) \text{ occurs quite surely (qs, in short) if } \Pr(\mathcal{E}) = 1 - O(N^{-a}) \text{ for any constant } a > 0.\]
Lemma 4.3. Suppose that \( \log N^* = O((N^* \lambda)^{1/2}) \) and \( N_t \to \infty \) with \( N \). Let \( x \) be chosen randomly from \([N]^{2M}_{j_2,j_3:D} \). Then \( qs \),

\[
D(x) = \left\{ \left| \nu_x(j) - \frac{N_t \lambda^j}{j! f(\lambda)} \right| \leq \left( 1 + \left( \frac{N_t \lambda^j}{j! f(\lambda)} \right)^{1/2} \right) \log^2 N, \ k \leq j \leq \log N \right\}.
\] (7)

We can now show \( G_x, x \in [n]^{2m}_{0,0} \) is a good model for \( G^\delta_{n,m} \geq 3 \). For this we only need to show now that

\[
\Pr(G_x \text{ is simple}) = \Omega(1).
\] (8)

For this we can use a result of McKay [8]. If we fix the degree sequence of \( x \) then \( x \) itself is just a random permutation of the multi-graph in which each \( j \in [n] \) appears \( d_x(j) \) times. This in fact is another way of looking at the configuration model of Bollobás [1]. The reference [8] shows that the probability \( G_x \) is simple is asymptotically equal to \( e^{-\left(1+o(1)\right)\rho(\rho+1)} \) where \( \rho = m_2/m \) and \( m_2 = \sum_{j \in [n]} d_x(j)(d_x(j) - 1) \). One consequence of the exponential tails in Lemma 4.3 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 \( G^\delta_{n,m} \geq 3 \).

All that is left now is to show that we can find a covering collection of paths that satisfy \( \mathbf{P1} \) e.g. \( |\mathcal{P}| \leq n^{0.48} \) will suffice. For this we need to analyse algorithm 2GREEDY of [4], which was described in Section 5.

5 Greedy Algorithm

We now describe the algorithm 2GREEDY of [4]. Our algorithm will be applied to the random graph \( G = G^\delta_{n,m} \geq 3 \) and analyzed in the context of \( G_x \). As the algorithm progresses, it makes changes to \( G \) and we let \( \Gamma \) 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

- \( \mu \) be the number of edges in \( \Gamma \),
- \( V_{0,j} = \{ v \in [n] : d_{\Gamma}(v) = 0, b(v) = j \}, j = 0,1, \)
- \( Y_k = \{ v \in [n] : d_{\Gamma}(v) = k \text{ and } b(v) = 0 \}, k = 1,2, \)
- \( Z_1 = \{ v \in [n] : d_{\Gamma}(v) = 1 \text{ and } b(v) = 1 \}, \)
- \( Y = \{ v \in [n] : d_{\Gamma}(v) \geq 3 \text{ and } b(v) = 0 \}, \quad \text{This is } J_3 \text{ of Section 4.} \)
\[ Z = \{ v \in [n] : d_\Gamma(v) \geq 2 \text{ and } b(v) = 1 \}, \quad \text{This is } J_2 \text{ of Section 4.} \]

- \( M \) 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 \( G_x \). This case is of course unnecessary when the input is simple i.e. for \( G_{n,m}^{s\geq k} \). Add \((v, w)\) to \( M \) and delete it from \( \Gamma \). 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 \( Z_1 \cup Y_1 \cup Y_2 \). Let \( w \) be a random neighbor of \( v \). Add \((v, w)\) to \( M \) and delete it to form \( \Gamma \). 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 there are at most \( n^{2/5} \) vertices left in \( \Gamma \). The output of 2GREEDY is set of edges in \( M \).

### 6 Analysis of 2GREEDY

We will use the following additional notation to that given in Section 5:

- \( m_i \): number of edges at time \( i \).
- \( Z_j, j \geq 2 \) and \( Y_j, j \geq 3 \) resp. are the subsets of \( Z \) and \( Y \) respectively constisting of vertices of degree \( j \).
- \( y_i = |Y|, z_i = |Z| \) at time \( i \).
- \( \zeta_i = |Y_1| + 2|Y_2| + |Z_1| \).

\[
\begin{align*}
p_{2,i} &= \frac{2|Z_2|}{2m_i} \quad \text{and} \quad p_{3,i} = \frac{3|Y_3|}{2m_i}.
\end{align*}
\]

We will show that w.h.p.

up until \( m_i \leq n^{0.42} \) we have \( \zeta_i \leq n^{0.41} = o(m_i) \).

Every component in \( M \) defines a path and the union of the vertices of these paths is \( V \).

The number \( \kappa \) of components of the 2-matching \( M \) output by 2GREEDY can be bounded as
follows. \( \kappa \) can be bounded by the number \( \kappa_1 \) of vertices of degree one or zero in \( M \) plus \( \kappa_2 \), the number of cycles. For every vertex \( v \in V \) that contributes to \( \kappa_1 \) 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 \( \Gamma \) or (ii) \( v \notin Z_1 \cup Y_1 \cup Y_2 \), at least 1 neighbor of \( v \) is removed from \( \Gamma \) 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 \( \kappa_1 \).

For the number of cycles spanned by \( M \), observe that at step \( i, \kappa_2 \) 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 \( \kappa_2 \).

Since w.h.p the maximum degree of \( G_0 \), and hence of \( \Gamma \), 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 \( (2 \log n)\zeta_i / 2m_i + O(1/m_i) \). Let \( \epsilon = 10^{-4} \). If \( \kappa_1 + \kappa_2 \) reaches \( n^{0.42} \) then there are at least \( \epsilon n^{0.42} / 2 \log n \) steps with \( m_i \in [n^{0.42 + (r-1)\epsilon}, n^{0.42 + r\epsilon}] \) for some integer \( 1 \leq r \leq 1/\epsilon \) that witness an increase of \( \kappa_1 + \kappa_2 \). The probability that this occurs for a fixed \( r \), while \( \zeta_i \leq n^{0.41} \), is bounded by

\[
\left( n^{0.42 + r\epsilon} / \epsilon n^{0.42} / 2 \log n \right) \leq \left( n^{0.42 + r\epsilon} / 2 \epsilon n^{0.42} \right) n^{0.41} \leq n^{-5}.
\]

Hence w.h.p. the total increase in \( \kappa_1 + \kappa_2 \) up until \( m_i \leq n^{0.42} \) or \( \zeta_i > n^{0.41} \), is bounded by \( n^{0.42} \). Once \( m_i \leq n^{0.42} \), at most \( n^{0.42} \) more components can be created, yielding in total at most \( 2n^{0.42} \) components.

We define the events

\( \mathcal{A}_i = \{(z_j + y_j)\lambda_j \geq \log^3 n \text{ for } j \leq i\} \) and \( \mathcal{B}_i = \{(\lambda_i \geq m_i^{-0.2}) \lor (y_i \geq m_i^{0.8})\} \).

We define the following random variables:

\[
\begin{align*}
X_i &= (\zeta_{i+1} - \zeta_i) \mathbb{I}(\mathcal{A}_i, \mathcal{B}_i, m_i \geq n^{0.42}, 0 < \zeta_i < n^{0.41}). \\
Y_i &= (\zeta_{i+1} - \zeta) \mathbb{I}(\mathcal{A}_i, \neg \mathcal{B}_i, m_i \geq n^{0.42}, 0 < \zeta_i < n^{0.41}). \\
X'_i &= (\zeta_{i+1} - \zeta) \mathbb{I}(\neg \mathcal{A}_i, m_i \geq n^{0.42}, 0 < \zeta_i < n^{0.41}). \\
Y'_i &= (\zeta_{i+1} - \zeta) \mathbb{I}(\neg \mathcal{A}_i, \neg \mathcal{B}_i, m_i \geq n^{0.42}).
\end{align*}
\]

For \( i > 0 \) we have that while \( m_i \geq n^{0.42} \), w.h.p.

\[
\min\{\zeta_i, n^{0.41}\} \leq M + \sum_{j=0}^{i-1} (X_i + Y_i + X'_i + Y'_i)
\]

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

We now prove high probability upper bounds on the random variables in (10) and only consider \( i \) such that

\[
m_i \geq n^{0.42}.
\]
We use the inequality \( m_i \geq n^{0.42} \) to impose that if \( \zeta_i \leq n^{0.41} \) then almost all of the vertices belong to \( Y \cup Z \). We will see from the analysis below that w.h.p.

\[
m_i \geq n^{0.42} \text{ implies } \zeta_i \leq n^{0.41}.
\]

Equation (80) of \[4\] states that if \( H_i \) denotes the history of the process up to the end of iteration \( i \), assuming the event \( A_i \) occurs, then

\[
\zeta_i > 0 \text{ implies } \mathbb{E}(\zeta_{i+1} - \zeta_i \mid H_i) \leq -\Omega(\min \{1, \lambda_i\}^2) + O\left(\frac{\log^2 m_i}{\lambda_i m_i}\right).
\]

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

**Case 1: \( A_i \wedge B_i \)**

**Case 1a**

If \( \lambda_i \geq m_i^{-0.2} \) we have from (13) that

\[
\mathbb{E}(X_i \mid H_i) \leq -c\lambda_i^2 \leq -cn^{0.4}
\]

for some constant \( c > 0 \).

**Case 1b:**

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

\[
\frac{|Z_3|}{|Z_2|} = \frac{\lambda_i}{3} \left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right),
\]

\[
\frac{|Z_4|}{|Z_2|} = \frac{\lambda_i^2}{12} \left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right),
\]

\[
\sum_{i \geq 5} |Z_i| \leq |Z_2|\lambda_i^3.
\]

Similarly

\[
\frac{|Y_4|}{|Y_3|} = \frac{\lambda_i}{4} \left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right),
\]

\[
\frac{|Y_5|}{|Y_3|} = \frac{\lambda_i^2}{20} \left(1 + O(m_i^{1/2}\lambda_i \log^2 m_i)\right),
\]

\[
\sum_{i \geq 6} |Y_i| \leq |Y_3|\lambda_i^3.
\]

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 \( \zeta_i \) 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:** $w \in Y_3 \cup Y_1 \cup Z_1$ 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,

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

(16)

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

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

(17)

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

$$
E[-1 + (d(w) - 1)(d(w, Z_2) + 2d(w, Y_3)) + O(f(w)))\mathbb{I}(w \in Z \setminus Z_2)|\mathcal{H}_i]
$$

$$
= -\Pr(w \in Z \setminus Z_2) + (3 - 1)\frac{3|Z_3|}{2m_i} + (4 - 1)\frac{4|Z_4|}{2m_i}(p_{2,i} + 2p_{3,i}) + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right).
$$

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

(18)

The -1 in the $d(w) - 1$ expression accounts for the edge $\{v, w\}$. Then the next term accounts for the other $d(w) - 1$ neighbors of $w$ and the possibility that they belong to either $Z_2$ or $Y_3$. To go from the second to the third line we used (14).

Finally observe that (14), (15) imply that

$$
1 = \frac{2|Z_2| + 3|Z_3| + 4|Z_4|}{2m_i} + \frac{3|Y_3| + 4|Y_4| + 5|Y_5|}{2m_i} + \frac{\zeta_i}{2m_i} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right)
$$

$$
= p_{2,i}\left(1 + \frac{\lambda_i}{2} + \frac{\lambda_i^2}{6}\right) + p_{3,i}\left(1 + \frac{\lambda_i}{3} + \frac{\lambda_i^2}{12}\right) + \frac{\zeta_i}{2m_i} + O\left(\frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3\right).
$$

(19)
Thus if Case 1 occurs we have by the Azuma inequality that

\[
\mathbb{E}(X_i | \mathcal{H}_i) \leq \left( -\frac{2\zeta_i}{2m_i} - \Pr(w \in Y) + [- \Pr(w \in Z_2) + p_{2,i}(p_{2,i} + 2p_{3,i})] \right) \left( 1 + O \left( \frac{1}{m_i} \right) \right)
\]

\[
\quad + \left( - \Pr(w \in Z \setminus Z_2) + p_{2,i} \left( \lambda_i + \frac{\lambda_i^2}{2} \right) (p_{2,i} + 2p_{3,i}) \right) + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right)
\]

\[
= -1 - \frac{\zeta_i}{2m_i} + p_{2,i} \left( 1 + \lambda_i + \frac{\lambda_i^2}{2} \right) (p_{2,i} + 2p_{3,i}) + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right) .
\]

Note that \( \lambda_i \leq \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} \).

Now use (19) to replace -1 by the squared expression to obtain

\[
\leq - \left[ p_{2,i} \left( 1 + \frac{\lambda_i}{2} + \frac{\lambda_i^2}{6} \right) + p_{3,i} \left( 1 + \frac{\lambda_i}{3} + \frac{\lambda_i^2}{12} \right) + \frac{\zeta_i}{2m_i} \right]^2
\]

\[
\quad + p_{2,i} \left( 1 + \lambda_i + \frac{\lambda_i^2}{2} \right) (p_{2,i} + 2p_{3,i}) - \frac{\zeta_i}{2m_i} + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right)
\]

\[
= -\frac{\lambda_i^2 p_{2,i}^2}{12} + 2p_{2,i} p_{3,i} \left( \frac{\lambda_i}{6} + \frac{\lambda_i^2}{12} \right) - p_{3,i}^2 \left( 1 + \frac{2\lambda_i}{3} + \frac{5\lambda_i^2}{18} \right) - \frac{\zeta_i}{2m_i}
\]

\[
\quad + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right)
\]

\[
= -\left( \frac{\lambda_i p_{2,i}}{4} - p_{3,i} \left( \frac{2}{3} + \frac{\lambda_i}{3} \right) \right)^2 - \frac{\lambda_i^2 p_{2,i}^2}{48} - p_{3,i}^2 \left( \frac{5}{9} + \frac{2\lambda_i}{9} + \frac{\lambda_i^2}{6} \right) - \frac{3\zeta_i}{2m_i}
\]

\[
\quad + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right)
\]

\[
\leq -\frac{\lambda_i^2 p_{2,i}^2}{48} - \frac{5p_{3,i}^2}{9} - \frac{\zeta_i}{2m_i} + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^3 \right) .
\]

(20)

In Case 1b we have that the events \( \mathcal{A}_i \cap \mathcal{B}_i, \lambda_i \leq m_i^{-0.2}, m_i \geq n^{0.42} \) occur. \( \mathcal{A}_i \) and \( \lambda_i \leq m_i^{-0.2} \) implies 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

\[
\mathbb{E}(X_i | \mathcal{H}_i) \leq -c'm_i^{-0.4} \leq -cn^{-0.4} .
\]

Thus if Case 1 occurs we have by the Azuma inequality that

\[
\sum_{i=0}^{j} \Pr \left( \sum_{i=0}^{j} X_i \geq n^{0.405} \right) \leq m_0 \max_{0 \leq j \leq m_0} \exp \left\{ - \left( \frac{n^{0.405} + c'jn^{-0.4}}{j \log^2 n} \right)^2 \right\} + n^{-6} = o(1) .
\]

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: \( A_i \land \neg B_i \)

To bound \( \sum_{i \in I}^j Y_i \), let \( R_i \) be the indicator of the event that \( \neg B_i \land \{ \zeta_i \leq n^{0.41} \} \) plus one of the cases (a),(b),(d),(e) and(f) from (16) occurs. Then, just as in Case 1, since the contribution of Case c to \( \mathbb{E}(X_i | H_i) \) is 0 and \( Y_i = 0 \) if \( \zeta_i \geq n^{0.41} \), we have that

\[
\mathbb{E}(Y_i R_i | H_i) \leq - \frac{\lambda_i^2 p_{2,i}^2}{48} - \frac{5p_{3,i}^2}{9} - \frac{\zeta_i}{2m_i} + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^2 \right)
\]

\[\leq - \frac{\lambda_i^2 p_{2,i}^2}{48} + O \left( \frac{\lambda_i \log^2 m_i}{m_i^{1/2}} + \lambda_i^2 \right)
\]

\[\leq O(m_i^{-1} \log^4 m_i). \tag{21}\]

For the last inequality we used that in the event \( A_i \land \neg B_i \) (14), (15) and (19) imply that \( p_{2,i} = 1 - o(1) \). In addition

\[
\mathbb{P}(R_i = 1) \leq \mathbb{P}(\text{Case(a)}) + \mathbb{P}(\text{Case(b)}) + \mathbb{P}(\text{Case(d)}) + \mathbb{P}(\text{Case(e)}) + \mathbb{P}(\text{Case(f)})
\]

\[= O \left( \frac{\zeta_i}{2m_i} + p_{3,i} + p_{2,i} p_{3,i} + p_{2,i} (1 - p_{3,i} - p_{2,i}) + \lambda_i \right) = O \left( \frac{\zeta_i}{2m_i} + p_{3,i} + \lambda_i \right). \tag{22}\]

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

In the event \( \neg B_i \) 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 \( \zeta_i \leq n^{0.41} \) then \( \mathbb{P}(R_i = 1) \leq m_i^{-0.2} \). Let \( \bar{R}_i = \bar{R}_i(\zeta_i < n^{0.41}) \), then,

\[
\sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} \bar{R}_i > n^{0.803} \right) \leq \sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} \bar{R}_i(\zeta_i > n^{0.8}) > n^{0.803} - n^{0.8} \right)
\]

\[\leq m_0 \exp \left\{ - \frac{(n^{0.803} - n^{0.8} - \sum_{i=1}^{m_0} m_i^{-0.2})^2}{2m_0} \right\} = o(1). \]

To obtain the exponential bound, we let \( Z_j = \sum_{i=0}^{j} \bar{R}_i(\zeta_i > n^{0.8}) \). We have

\[\mathbb{E} Z_j \leq \sum_{i=0}^{m_0} m_i^{-0.2} = O(n^{0.8}) \] and then we can use the Chernoff bounds, since our bounds for \( \bar{R}_i = 1 \) hold given the history of the process so far.

It follows that,

\[
\sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} Y_i \geq n^{0.405} \right) = \sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} Y_i \bar{R}_i \geq n^{0.405} \right)
\]

\[\leq \sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} \bar{R}_i > n^{0.803} \right) + \sum_{j=0}^{m_0} \mathbb{P} \left( \sum_{i=0}^{j} Y_i \bar{R}_i \geq n^{0.405} | \sum_{i=0}^{j} \bar{R}_i \leq n^{0.803} \right)
\]

\[\leq o(1) + m_0 \max_{j \leq n^{0.003}} \exp \left\{ - \frac{(n^{0.405} - \sum_{i=0}^{m_0} m_i^{-1} \log^3 m_i)^2}{j \log^2 n} \right\}
\]

\[\leq o(1) + m_0 \max_{j \leq n^{0.003}} \exp \left\{ - \frac{(n^{0.405} - o(1))^2}{j \log^2 n} \right\} = o(1). \tag{23}\]
To obtain the third line we use the fact that w.h.p. $|Y_i| \leq \log n$, which follows from a high probability bound of $o(\log n)$ on the maximum degree of $G$.

**Cases 3 & 4:** $\neg A_i$

Let $T_1 = \max \{i : (A_i \text{ occurs}) \land (m_i \geq n^{0.42})\}$. 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 (14), (15) 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

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

(24)

**Case 3:** $\neg A_i \land B_i$

Given the above we replace (19) by

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

(25)

Following this we replace (20) by

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

(26)

In the events $\neg A_i \land B_i$, $y_i \geq m_i^{0.8}$ and so $p_{3,i} \geq m_i^{-0.2}$. Therefore

$$\mathbb{E}(X_i' \mid \mathcal{H}_i) \leq -c' m_i^{-0.4} \leq -cn^{-0.4}.$$  

Thus if Case 3 occurs we have by the Azuma inequality that

$$\sum_{i=0}^{\infty} \Pr\left(\sum_{i=0}^{\infty} X_i' \geq n^{0.405}\right) \leq m_0 \max_{0 \leq j \leq m_0} \exp\left\{-\frac{(n^{0.405} + cjn^{-0.4})^2}{j \log^3 n}\right\} + n^{-6} = o(1).$$

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:** $\neg A_i \land \neg B_i$

As in Case 2 we have

$$\mathbb{E}(Y_i'R_i \mid \mathcal{H}_i) \leq O(m_i^{-1}\log^4 n)$$

where $R_i$ (and subsequentially $\bar{R}_i$) is defined exactly as in Case 3. Hence, just as in (23) we get

$$\sum_{j=0}^{m_0} \Pr\left(\sum_{i=0}^{j} Y_i' \geq n^{0.405}\right) = o(1).$$

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

$$\min\{\zeta_i, n^{0.41}\} \leq \log^2 n + 4n^{0.405} < n^{0.409}.$$

Hence w.h.p. there does not exist $i$ such that $m_i \geq n^{0.42}$ and $\zeta_i > n^{0.41}$. And this therefore completes the proof that w.h.p. $\zeta_i \leq n^{0.41}$ up to the point where $m_i \leq n^{0.42}$, verifying (9).
7 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 Posá’s theorem that is given in [6]. This may not be so easy, as explained in Remark 4.1 of [6].

References


