On random k-out subgraphs of large graphs

Alan Frieze[*] Tony Johansson
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213
U.S.A.
Email: alan@random.math.cmu.edu; tjohanss@andrew.cmu.edu

July 11, 2015

Abstract

We consider random subgraphs of a fixed graph $G = (V, E)$ with large minimum degree. We fix a positive integer k and let G_k be the random subgraph where each $v \in V$ independently chooses k random neighbors, making kn edges in all. When the minimum degree $\delta(G) \geq (\frac{1}{2} + \varepsilon)n$, $n = |V|$ then G_k is k-connected w.h.p. for $k = O(1)$; Hamiltonian for k sufficiently large. When $\delta(G) \geq m$, then G_k has a cycle of length $(1-\varepsilon)m$ for $k \geq k_\varepsilon$. By w.h.p. we mean that the probability of non-occurrence can be bounded by a function $\phi(n)$ (or $\phi(m)$) where $\lim_{n\to\infty} \phi(n) = 0$.

1 Introduction

The study of random graphs since the seminal paper of Erdős and Rényi [2] has by and large been restricted to analysing random subgraphs of the complete graph. This is not of course completely true. There has been a lot of research on random subgraphs of the hypercube and grids (percolation). There has been less research on random subgraphs of arbitrary graphs G, perhaps with some simple properties.

In this vain, the recent result of Krivelevich, Lee and Sudakov [8] brings a refreshing new dimension. They start with an arbitrary graph G which they assume has minimum degree at least k. For $0 \leq p \leq 1$ we let G_p be the random subgraph of G obtained by independently keeping each edge of G with probability p. Their main result is that if $p = \omega/k$ then G_p has a cycle of length $(1 - o_k(1))k$ with probability $1 - o_k(1)$. Here $o_k(1)$ is a function of k that tends to zero as $k \to \infty$. Riordan [11] gave a much simpler proof of this result. Krivelevich and Samotij [10] proved the existence of long cycles for the case where $p \geq \frac{1+\varepsilon}{k}$ and G is $\mathcal{H}$-free for some fixed set of graphs $\mathcal{H}$. Frieze and Krivelevich [6] showed that G_p is non-planar with probability $1 - o_k(1)$ when $p \geq \frac{1+\varepsilon}{k}$ and G has minimum degree at least k. In related works, Krivelevich, Lee and Sudakov [9] considered a random subgraph of a "Dirac Graph" i.e. a graph with n vertices and

[*]Research supported in part by NSF grant ccf1013110

minimum degree at least $n/2$. They showed that if $p \geq \frac{C \log n}{n}$ for sufffficently large n then G_p is Hamiltonian with probability $1 - o_n(1)$.

The results cited above can be considered to be generalisations of classical results on the random graph $G_{n,p}$, which in the above notation would be $(K_n)_p$. In this paper we will consider generalising another model of a random graph that we will call $K_n(k-out)$. This has vertex set $V = [n] = \{1, 2, \ldots, n\}$ and each $v \in V$ independently chooses k random vertices as neighbors. Thus this graph has kn edges and average degree $2k$. This model in a bipartite form where the two parts of the partition restricted their choices to the opposing half was first considered by Walkup [13] in the context of perfect matchings. He showed that $k \geq 2$ was sufficient for bipartite $K_{n,n}(k - out)$ to contain a perfect matching. Matchings in $K_n(k - out)$ were considered by Shamir and Upfal [12] who showed that $K_n(5 - out)$ has a perfect matching w.h.p., i.e. with probability $1 - o(1)$ as $n \to \infty$. Later, Frieze [4] showed that $K_n(2 - out)$ has a perfect matching w.h.p. Fenner and Frieze [5] had earlier shown that $K_n(k - out)$ is k-connected w.h.p. for $k \geq 2$. After several weaker results, Bohman and Frieze [1] proved that $K_n(3 - out)$ is Hamiltonian w.h.p. To generalise these results and replace K_n by an arbitrary graph G we will define $G(k - out)$ as follows: We have a fixed graph $G = (V, E)$ and each $v \in V$ independently chooses k random neighbors, from its neighbors in G. It will be convenient to assume that each v makes its choices with replacement. To avoid cumbersome notation, we will from now on assume that G has n vertices and we will refer to $G(k - out)$ as G_k. We implicitly consider G to be one of a sequence of larger and larger graphs with $n \to \infty$. We will say that events occur w.h.p. if their probability of non-occurrence can be bounded by a function that tends to zero as $n \to \infty$.

For a vertex $v \in V$ we let $d_G(v)$ denotes its degree in G. Then we let $\delta(G) = \min_{v \in V} d_G(v)$. We will first consider what we call Strong Dirac Graphs (SDG) viz graphs with $\delta(G) \geq \left(\frac{1}{2} + \varepsilon\right) n$ where ε is an arbitrary positive constant.

Theorem 1. *Suppose that G is an SDG. Suppose that the k neighbors of each vertex are chosen without replacement. Then w.h.p. G_k is k-connected for $2 \leq k = o(\log^{1/2} n)$.*

If the k neighbors of each vertex are chosen with replacement then there is a probability, bounded above by $1 - e^{-k^2}$ that G_k will have minimum degree $k - 1$, in which case we can only claim that G_k will be $(k - 1)$-connected.

Theorem 2. *Suppose that G is an SDG. Then w.h.p. there exists a constant k_ε such that if $k \geq k_\varepsilon$ then G_k is Hamiltonian.*

We get essentially the same result if the k neighbors of each vertex are chosen with replacement.

Note that we need $\varepsilon > 0$ in order to prove these results. Consider for example the case where G consists of two copies of $K_{n/2}$ plus a perfect matching M between the copies. In this case there is a probability greater than or equal to $\left(1 - \frac{2k}{n}\right)^{n/2} \sim e^{-k}$ that no edge of M will occur in G_k.

We note the following easy corollary of Theorem 2.

Corollary 3. *Let k_ε be as in Theorem 2. Suppose that G is an SDG and we give each edge of G a random independent uniform $[0, 1]$ edge weight. Let Z denote the length of the shortest travelling salesperson tour of G. Then $\mathbf{E}(Z) \leq \frac{2(k_\varepsilon+1)}{1+2\varepsilon}$.*

We will next turn to graphs with large minimum degree, but not necessarily SDG's. Our proofs use Depth First Search (DFS). The idea of using DFS comes from Krivelevich, Lee and Sudakov [8].

Theorem 4. *Suppose that G has minimum degree m where $m \to \infty$ with n. For every $\varepsilon > 0$ there exists a constant k_ε such that if $k \geq k_\varepsilon$ then w.h.p. G_k contains a path of length $(1-\varepsilon)m$.*

Using this theorem as a basis, we strengthen it and prove the existence of long cycles.

Theorem 5. *Suppose that G has minimum degree m where $m \to \infty$ with n. For every $\varepsilon > 0$ there exists a constant k_ε such that if $k \geq k_\varepsilon$ then w.h.p. G_k contains a cycle of length $(1-\varepsilon)m$.*

We finally note that in a recent paper, Frieze, Goyal, Rademacher and Vempala [3] have shown that G_k is useful in the construction of sparse subgraphs with expansion properties that mirror those of the host graph G.

2 Connectivity: Proof of Theorem 1

In this section we will assume that each vertex makes its choices without replacement. Let $G = (V, E)$ be an SDG. Let $c = 1/(8e)$. We need the following lemma.

Lemma 6. *Let G be an SDG and let $C = 48/\varepsilon$. Then w.h.p. there exists a set $L \subseteq V$, where $|L| \leq C \log n$, such that each pair of vertices $u, v \in V \setminus L$ have at least $12 \log n$ common neighbors in L.*

Proof. Define $L_p \subseteq V$ by including each $v \in V$ in L_p with probability $p = C \log n/2n$. Since $\delta(G) \geq (1/2+\varepsilon)n$, each pair of vertices in G has at least $2\varepsilon n$ common neighbors in G. Hence, the number of common neighbors in L_p for any pair of vertices in $V \setminus L_p$ is bounded from below by a $\mathrm{Bin}(2\varepsilon n, p)$ random variable.

$$
\begin{aligned}
& \mathbf{Pr}\left\{\exists u, v \in V \setminus L_p \text{ with less than } 12 \log n \text{ common neighbors in } L\right\} \\
\leq\ & n^2 \mathbf{Pr}\left\{\mathrm{Bin}(2\varepsilon n, p) \leq 12 \log n\right\} \\
=\ & n^2 \mathbf{Pr}\left\{\mathrm{Bin}(2\varepsilon n, p) \leq \varepsilon n p\right\} \\
\leq\ & n^2 e^{-\varepsilon n p/8} \\
=\ & o(1).
\end{aligned}
$$

The expected size of L_p is $\frac{1}{2} C \log n$ and so the Chernoff bounds imply that w.h.p. $|L_p| \leq C \log n$. Thus there exists a set L, $|L| \leq C \log n$, with the desired property. $\square$

Let L be a set as provided by the previous lemma, and let G'_k denote the subgraph of G_k induced by $V \setminus L$.

Lemma 7. *Let $c = 1/(8e)$. If $k \geq 2$ then w.h.p. all components of G'_k are of size at least cn. Furthermore, removing any set of $k-1$ vertices from G'_k produces a graph consisting entirely of components of size at least cn, and isolated vertices.*

Proof. We first show that w.h.p. G'_k contains no isolated vertex. The probability of G'_k containing an isolated vertex is bounded by

$$\mathbf{Pr}\left\{\exists v \in V \setminus L \text{ which chooses neighbors in } L \text{ only}\right\} \leq n\left[\frac{C\log n}{\frac{1}{2}n}\right]^k = o(1),$$

where L and C are as in Lemma 6.

We now consider the existence of small non-trivial components S after the removal of at most $k-1$ vertices A. Then,

$$\begin{aligned}
&\mathbf{Pr}\left\{\exists S, A,\, 2 \leq |S| \leq cn,\, |A| = k-1, \text{ such that } S \text{ only chooses neighbors in } S \cup L \cup A\right\} \\
&\leq \sum_{l=2}^{cn}\sum_{|S|=l}\sum_{|A|=k-1}\left[\frac{l+k-2+C\log n}{\left(\frac{1}{2}+\varepsilon\right)n}\right]^{lk} \\
&\leq \sum_{l=2}^{cn}\binom{n}{l}\binom{n-l}{k-1}\left[\frac{l+C\log n}{\frac{1}{2}n}\right]^{lk} \\
&\leq \sum_{l=2}^{cn}\left(\frac{ne}{l}\right)^l n^{k-1}\left[\frac{l+C\log n}{\frac{1}{2}n}\right]^{lk} \\
&= 2^k e\sum_{l=2}^{cn}\left[\frac{2^k e(l+C\log n)^k}{n^{k-1}l}\right]^{l-1}\frac{(l+C\log n)^k}{l}.
\end{aligned}$$

Now when $2 \leq l \leq \log^2 n$ we have

$$2^k e(l+C\log n)^k \leq \log^{3k} n \text{ and } \frac{(l+C\log n)^k}{l} \leq \log^{3k} n.$$

And when $\log^2 n \leq l \leq cn$ we have

$$2^k e(l+C\log n)^k \leq (2+o(1))^k e l^k \text{ and } \frac{(l+C\log n)^k}{l} = (1+o(1))l^{k-1},$$

which implies that

$$\begin{aligned}
\left[\frac{2^k e(l+C\log n)^k}{n^{k-1}l}\right]^{l-1}\frac{(l+C\log n)^k}{l} &\leq \frac{((2+o(1))^k e)^{l-1} l^{l(k-1)}}{n^{(k-1)(l-1)}} \leq \\
&((2+o(1))^k e)^{l-1} c^{l(k-1)} n^{k-1} = ((2+o(1))^k e)^{l-1} c^{l(k-1)},
\end{aligned}$$

since $n^{k-1} = (n^{(k-1)/(l-1)})^{l-1} = (1+o(1))^{l-1}$.

Continuing, we get a bound of

$$2^k e\left(\sum_{l=2}^{\log^2 n}\left[\frac{\log^{6k} n}{n^{k-1}}\right]^{l-1} + \sum_{l=\log^2 n}^{cn}((2+o(1))^k e c^{k-1})^{l-1}\right) = o(1).$$

$\square$

This proves that w.h.p. G'_k consists of $r \leq 1/c$ components $J_1, J_2, ..., J_r$ and that removing any $k-1$ vertices will only leave isolated vertices and components of size at least cn.

Lemma 8. *W.h.p., for any $i \neq j$, there exist k vertex-disjoint paths (of length 2) from J_i to J_j in G_k.*

Proof. Let X be the number of vertices in L which pick at least one neighbor in J_1 and at least one in J_2. Furthermore, let X_{uvw} be the indicator variable for $w \in L$ picking $u \in J_1$ and $v \in J_2$ as its neighbors. Note that these variables are independent of G'_k. Let $c = 1/(8e)$ as in Lemma 7 and let $C = 24/\varepsilon$ as in Lemma 6. For $w \in L$ we let

$$X_w = \sum_{\substack{(u,v)\in J_1\times J_2 \\ w\in N_G(J_1)\cap N_G(J_2)}} X_{uvw}.$$

These are independent random variables with values in $\{0, 1, \ldots, k\}$ and $X = \sum_{w\in L} X_w$. Then,

$$\begin{aligned}
\mathbf{E}\, X &= \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \mathbf{E}\, X_{uvw} \\
&= \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \frac{\binom{d_G(w)}{k-2}}{\binom{d_G(w)}{k}} \\
&\geq \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \frac{1}{n^2} \\
&\geq \frac{24(cn)^2\log n}{n^2} \\
&= 24c^2\log n.
\end{aligned}$$

We apply the following inequality, Theorem 1 of Hoeffding [7]: Let $Z_1, Z_2, \ldots, Z_M$ be independent and satisfy $0 \leq Z_i \leq 1$ for $i = 1, 2, \ldots, M$. If $Z = Z_1 + Z_2 + \cdots + Z_M$ then for all $t \geq 0$,

$$\mathbf{Pr}\left\{|Z - \mathbf{E}\, Z| \geq t\right\} \leq e^{-2t^2/M}. \tag{1}$$

Putting $Z_w = X_w/k$ for $w \in L$ and $Z = X/k$ and applying (1), we get

$$\begin{aligned}
\mathbf{Pr}\left\{X \leq k\right\} = \mathbf{Pr}\left\{Z \leq 1\right\} \leq \mathbf{Pr}\left\{Z \leq \frac{\mathbf{E}\, Z}{2}\right\} &\leq \exp\left\{-\frac{(\mathbf{E}\, Z)^2}{2|L|}\right\} \\
&= \exp\left\{-\frac{(\mathbf{E}\, X)^2}{2k^2|L|}\right\} = o(1). \qquad (2)
\end{aligned}$$

Now for $w_1 \neq w_2 \in L$ let $\mathcal{E}(w_1, w_2)$ be the event that w_1, w_2 make a common choice. Then

$$\mathbf{Pr}\left\{\exists w_1, w_2 : \mathcal{E}(w_1, w_2)\right\} = O\left[\frac{k^2\log^2 n}{n}\right] = o(1). \tag{3}$$

To see this, observe that for a fixed w_1, w_2 and a choice of w_2, the probability this choice is also one of w_1's is at most $\frac{k}{n/2}$. Now multiply by the number k of choices for w_2. Finally multiply by $|L|^2$ to account for the number of possible pairs w_1, w_2.

Equations (2) and (3) together show that w.h.p., there are k node-disjoint paths from J_1 to J_2. Since the number of linear size components is bounded by a constant, this is true for all pairs J_i, J_j w.h.p. $\square$

We can complete the proof of Theorem 1. Suppose we remove l vertices from L and $k-1-l$ vertices from the remainder of G. We know from Lemma 6 that $V \setminus L$ induces components $C_1, C_2, \ldots, C_r$ of size at least cn. There cannot be any isolated vertices in $V \setminus L$ as G_k has minimum degree at least k. Recall that each vertex makes k choices without replacement. Lemmas 6, 7 and 8 imply that $r = 1$ and that every vertex in L is adjacent to C_1. $\square$

3 Hamilton cycles: Proof of Theorem 2

Let G be a graph with $\delta(G) \geq (1/2+\varepsilon)n$, and let k be a positive integer.

Let $\mathcal{D}(k,n) = \{D_1, D_2, ..., D_M\}$ be the $M = \prod_{v \in V} \binom{d_G(v)}{k} \leq \binom{n-1}{k}^n$ directed graphs obtained by letting each vertex x of G choose k G-neighbors $y_1, ..., y_k$, and including in D_i the k arcs (x, y_i). Define $\vec{N}_i(x) = \{y_1, ..., y_k\}$ and for $S \subseteq V$ let $\vec{N}_i(S) = \bigcup_{x \in S} \vec{N}_i(x) \setminus S$. For a digraph D we let $G(D)$ denote the graph obtained from D by ignoring orientation and coalescing multiple edges, if necessary. We let $\Gamma_i = G(D_i)$ for $i = 1, 2, \ldots, M$. Let $\mathcal{G}(k,n) = \{\Gamma_1, \Gamma_2, ..., \Gamma_M\}$ be the set of k-out graphs on G. Below, when we say that D_i is Hamiltonian we actually mean that Γ_i is Hamiltonian. (It will occasionally enable more succint statements).

For each D_i, let $D_{i1}, D_{i2}, ..., D_{i\kappa}$ be the $\kappa = k^n$ different edge-colorings of D_i in which each vertex has $k-1$ outgoing green edges and one outgoing blue edge. Define Γ_{ij} to be the colored (multi)graph obtained by ignoring the orientation of edges in D_{ij}. Let Γ_{ij}^g be the subgraph induced by green edges.

$\vec{N}(S)$ refers to $\vec{N}_i(S)$ when i is chosen uniformly from $[M]$, as it will be for G_k.

Lemma 9. *Let $k \geq 5$. There exists an $\alpha > 0$ such that the following holds w.h.p.: for any set $S \subseteq V$ of size $|S| \leq \alpha n$, $|\vec{N}(S)| \geq 3|S|$.*

Proof. The claim fails if there exists an S with $|S| \leq \alpha n$ such that there exists a T, $|T| = 3|S|-1$ such that $\vec{N}(S) \subseteq T$. The probability of this is bounded from above by

$$\begin{aligned}
& \sum_{l=1}^{\alpha n} \binom{n}{l} \binom{n-l}{3l-1} \prod_{v \in S} \left[\binom{4l-2}{k} \Big/ \binom{d_G(v)}{k} \right] \\
\leq \quad & \sum_{l=1}^{\alpha n} \left(\frac{ne}{l} \right)^l \left(\frac{ne}{3l-1} \right)^{3l-1} \left[\frac{4le}{n/2} \right]^{kl} \\
\leq \quad & \sum_{l=1}^{\alpha n} \left[e^4 (8e)^k \left(\frac{l}{n} \right)^{k-4} \right]^l \\
= \quad & o(1),
\end{aligned}$$

for $\alpha = 2^{-16} e^{-9}$. $\square$

We say that a digraph D_i *expands* if $|\vec{N}_i(S)| \geq 3|S|$ whenever $|S| \leq \alpha n$, $\alpha = 2^{-16}e^{-9}$. Since almost all D_i expand, we need only prove that an expanding D_i almost always gives rise to a Hamiltonian Γ_i. Write $\mathcal{D}'(k,n)$ for the set of expanding digraphs in $\mathcal{D}(k,n)$ and let $\mathcal{G}'(k,n) = \{\Gamma_i : D_i \in \mathcal{D}'(k,n)\}$.

Let H be any graph, and suppose $P = (v_1, ..., v_k)$ is a longest path in H. If $t \neq 1, k-1$ and $\{v_k, v_t\} \in E(H)$, then $P' = (v_1, ..., v_t, v_k, v_{k-1}, ..., v_{t+1})$ is also a longest path of H. Repeating this rotation for P and all paths created in the process, keeping the endpoint v_1 *fixed*, we obtain a set $EP(v_1)$ of other endpoints.

For $S \subseteq V(H)$ we let $N_H(S) = \{w \notin S : \exists v \in S \ s.t. \ vw \in E(H)\}$.

Lemma 10 (Pósa)**.** *For any endpoint x of any longest path in any graph H, $|N_H(EP(x))| \leq 2|EP(x)| - 1$.*

We say that an undirected graph expands if $|N_H(S)| \geq 2|S|$ whenever $|S| \leq \alpha n$, assuming $|V(H)| = n$. Note that the definition of expanding slightly differs from the digraph case.

Lemma 11. *Consider a green subgraph Γ_{ij}^g. W.h.p., there exists an $\alpha > 0$ such that for every longest path P in Γ_{ij}^g and endpoint x of P, $|EP(x)| > \alpha n$.*

Proof. Let $H = \Gamma_{ij}^g$. We argue that if D_i expands then so does H. If $|\vec{N}_i(S)| \geq 3|S|$, then $|N_H(S)| \geq 2|S|$, since each vertex of S picks at most one blue edge outside of S. Thus H expands. In particular, Lemma 9 implies that if $|S| \leq \alpha n$, then $|\vec{N}(S)| \geq 3|S|$ and hence $|N_H(S)| \geq 2|S|$. By Lemma 10, this implies that $|EP(x)| > \alpha n$ for any longest path P and endpoint x. □

Define a_{ij} to be 1 if $G(\Gamma_{i,j})$ is connected and Γ_{ij}^g contains a longest path of Γ_{ij}, $1 \leq i \leq M_1$ (i.e. Γ_{ij} is not Hamiltonian), and 0 otherwise.

Let M_1 be the number of expanding digraphs D_i among $D_1, ..., D_M$ for which $G(D_i)$ is connected and Γ_i is not Hamiltonian. We aim to show that $M_1/M \to 0$ as n tends to infinity. W.l.o.g. suppose $\mathcal{N}(k,n) = \{D_1, ..., D_{M_1}\}$ are the connected expanding digraphs which are not Hamiltonian.

Lemma 12. *For $1 \leq i \leq M_1$, we have $\sum_{j=1}^{\kappa} a_{ij} \geq (k-1)^n$.*

Proof. Fix $1 \leq i \leq M_1$ and a longest path P_i of Γ_i. Uniformly picking one of $D_{i1}, ..., D_{i\kappa}$, we have

$$\begin{aligned} \mathbf{Pr}\{a_{ij} = 1\} &\geq \mathbf{Pr}\left\{E(P_i) \subseteq E(\Gamma_{ij}^g)\right\} \\ &\geq \left(1 - \frac{1}{k}\right)^{|E(P_i)|} \\ &\geq \left(1 - \frac{1}{k}\right)^n \end{aligned}$$

The lemma follows from the fact that there are k^n colorings of D_i. □

Let $\Delta \in \mathcal{D}(k-1,n)$ be expanding and non-Hamiltonian and for the purposes of exposition consider its edges to be colored green. Let $D \in \mathcal{D}(k,n)$ be the random digraph obtained by letting each vertex of Δ randomly choose another edge, which will be colored blue. Let $\overline{B_\Delta}$ be the event (in the probability space of randomly chosen blue edges to be added to Δ):

D has an edge between the endpoints of a longest path of $G(\Delta)$, or

D has an edge from an endpoint of a longest path of Δ to the complement of the path.

Note that the occurrence of $\overline{B_\Delta}$ implies that the corresponding $a_{ij} = 0$. If $a_{ij} = 1$ then the connectivity of Γ_{ij} imlies that $G(D)$ has a longer path than $G(\Delta)$. Let B_Δ be the complement of $\overline{B_\Delta}$ and for Hamiltionian Δ let $B_\Delta = \emptyset$.

Let N_Δ be the number of i,j such that $\Gamma_{ij}^g = \Delta$. We have

$$\sum_{i,j:\Gamma_{ij}^g=\Delta} a_{ij} = N_\Delta \mathbf{Pr}\{B_\Delta\} \tag{4}$$

The number of non-Hamiltonian graphs is bounded by

$$\begin{aligned} M_1 &\leq \sum_{i=1}^{M}\sum_{j=1}^{\kappa}\frac{a_{ij}}{(k-1)^n} \\ &\leq \frac{\sum_\Delta N_\Delta \mathbf{Pr}\{B_\Delta\}}{(k-1)^n} \\ &\leq \frac{Mk^n \max_\Delta \mathbf{Pr}\{B_\Delta\}}{(k-1)^n} \\ &= M\frac{\max_\Delta \mathbf{Pr}\{B_\Delta\}}{(1-1/k)^n} \end{aligned} \tag{5}$$

Fix a $\Delta \in \mathcal{N}(k-1,n)$ and a longest path P_Δ of $G(\Delta)$. Let EP be the set of vertices which are endpoints of a longest path of $G(\Delta)$ that is obtainable from P_Δ by rotations. For $x \in EP$, say x is of Type I if x has at least $\varepsilon n/2$ neighbors outside P_Δ, and Type II otherwise. Let E_1 be the set of Type I endpoints, and E_2 the set of Type II endpoints.

Partition the set of expanding green graphs by

$$\mathcal{D}'(k-1,n) = \mathcal{H}(k-1,n) \cup \mathcal{N}_1(k-1,n) \cup \mathcal{N}_2(k-1,n) \tag{6}$$

where $\mathcal{H}(k-1,n)$ is the set of Hamiltonian graphs, $\mathcal{N}_1(k-1,n)$ the set of non-Hamiltonian graphs with $|E_1| \geq \alpha n/2$ and $\mathcal{N}_2(k-1,n)$ the set of non-Hamiltonian graphs with $|E_1| < \alpha n/2$. Here $\alpha > 0$ is provided by Lemma 11.

Lemma 13. *For $\Delta \in \mathcal{N}_1(k-1,n)$, $\mathbf{Pr}\{B_\Delta\} \leq e^{-\varepsilon\alpha n/4}$.*

Proof. Let each $x \in E_1$ choose a neighbor $y(x)$. The event B_Δ is included in the event $\{\forall x \in E_1 : y(x) \in P_\Delta\}$. We have

$$\begin{aligned} \mathbf{Pr}\{B_\Delta\} &\leq \mathbf{Pr}\{\forall x \in E_1 : y(x) \in P_\Delta\} \\ &= \prod_{x\in E_1}\frac{d_{P_\Delta}(x)}{d_G(x)} \\ &\leq \left(1-\frac{\varepsilon}{2}\right)^{\alpha n/2} \end{aligned}$$

where $d_{P_\Delta}(x)$ denotes the number of neighbors of x inside P_Δ. $\square$

Lemma 14. *For $\Delta \in \mathcal{N}_2(k-1,n)$, $\mathbf{Pr}\{B_\Delta\} \le e^{-\varepsilon\alpha^2 n/129}$.*

Proof. Let $X \subseteq E_2$ be a set of $\alpha n/4$ Type II endpoints. X exists because $|EP| \ge \alpha n$ and at most $\alpha n/2$ vertices in EP are of type I. For each $x \in X$, let P_x be a path obtained from P_Δ by rotations that has x as an endpoint. Let $A(x)$ be the set of Type II vertices $y \notin X$ such that a path from x to y in Δ can be obtained from P_x by a sequence of rotations with x fixed. By Lemma 11 we have $|A(x)| \ge \alpha n/4$ for each x, since $A(x) = EP(x) \setminus (E_1 \cup X)$.

Let $P_{x,y}$ be a path with endpoints $x \in X, y \in A(x)$ obtained from P_x by rotations with x fixed, and label the vertices on $P_{x,y}$ by $x = z_0, z_1, ..., z_l = y$. Suppose y chooses some z_i on the path with its blue edge. If $\{z_{i+1}, x\} \in E(G)$, let $B_y(x) = \{z_{i+1}\}$. Write $v(y)$ for z_{i+1}. If $\{z_{i+1}, x\} \notin E(G)$, or if y chooses a vertex outside P, let $B_y(x) = \emptyset$.

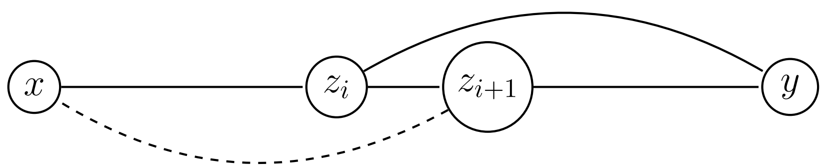

Figure 1: Suppose y chooses z_i. The vertex z_{i+1} is included in $B(x)$ if and only if $\{x, z_{i+1}\} \in E(G)$.

There will be at least $2\left(\frac{1}{2} + \frac{\varepsilon}{2}\right)n - n = \varepsilon n$ choices for i for which $\{x, z_{i+1}\} \in E(G)$. Let Y_x be the number of $y \in A(x)$ such that $B_y(x)$ is nonempty. This variable is bounded stochastically from below by a binomial $\mathrm{Bin}(\alpha n/4, \varepsilon)$ variable, and by a Chernoff bound we have that

$$\mathbf{Pr}\left\{\exists x : Y_x \le \frac{\varepsilon\alpha n}{8}\right\} \le n \exp\left\{-\frac{\varepsilon\alpha n}{32}\right\} \tag{7}$$

Define $B(x) = \bigcup_{y \in A(x)} B_y(x)$. If x chooses a vertex in $B(x)$ then $\overline{B_\Delta}$ occurs. Conditional on $Y_x \ge \varepsilon\alpha n/8$ for all $x \in X$, let $y_1, y_2, ..., y_r$ be $r = \varepsilon\alpha n/8$ vertices whose choice produces a nonempty $B_y(x)$. Let $Z_x = |B(x)|$, and for $i = 1, ..., r$ define Z_i to be 1 if $v(y_i)$ is distinct from $v(y_1), ..., v(y_{i-1})$ and 0 otherwise. We have $Z_x = \sum_{i=1}^r Z_i$, and each Z_i is bounded from below by a Bernoulli variable with parameter $1 - \alpha/8$. To see this, note that y_i has at least εn choices resulting in a nonempty $B_{y_i}(x)$ since x and y_i are of Type II, so

$$\mathbf{Pr}\{\exists j < i : v(y_j) = v(y_i)\} \le \frac{i-1}{\varepsilon n} \le \frac{\varepsilon\alpha n/8}{\varepsilon n} = \frac{\alpha}{8} \tag{8}$$

Since $\alpha/8 < 1/2$, Z_x is bounded stochastically from below by a binomial $\mathrm{Bin}(\varepsilon\alpha n/8, 1/2)$ variable, and so

$$\mathbf{Pr}\left\{\exists x : Z_x < \frac{\varepsilon\alpha n}{32}\right\} \le n \exp\left\{-\frac{\varepsilon\alpha n}{128}\right\} \tag{9}$$

Each x for which $Z_x \ge \varepsilon\alpha n/32$ will choose a vertex in $B(x)$ with probability

$$\frac{|B(x)|}{d_G(x)} \ge \frac{\varepsilon\alpha n/32}{n} = \frac{\varepsilon\alpha}{32} \tag{10}$$

Hence we have

$$\mathbf{Pr}\left\{B_\Delta\right\} \le \left(1-\frac{\varepsilon\alpha}{32}\right)^{\alpha n/4} + n\exp\left\{-\frac{\varepsilon\alpha n}{32}\right\} + n\exp\left\{-\frac{\varepsilon\alpha n}{128}\right\} \le e^{-\varepsilon\alpha^2 n/129}. \tag{11}$$

□

We can now complete the proof of Theorem 2. From Lemmas 13 and 14 we have

$$\mathbf{Pr}\left\{B_\Delta\right\} \le \max\left\{e^{-\varepsilon\alpha n/4}, e^{-\varepsilon\alpha^2 n/129}\right\}.$$

Going back to (5) with $k = C/\varepsilon$ we have

$$\begin{aligned}
\mathbf{Pr}\left\{G_k \text{ is non-Hamiltonian}\right\} &= o(1) + \frac{M_1}{M} \\
&\le o(1) + \frac{\max_\Delta \mathbf{Pr}\left\{B_\Delta\right\}}{(1-1/k)^n} \\
&= o(1) + \left[\frac{e^{-\varepsilon\alpha^2/129}}{1-\varepsilon/C}\right]^n \\
&\le o(1) + \exp\left\{-\varepsilon\left(\frac{\alpha^2}{129} - \frac{2}{C}\right)n\right\} \\
&= o(1),
\end{aligned}$$

for $C = 259/\alpha^2$. □

We can now prove Corollary 3. We follow an argument based on Walkup [14]. If X_e is the *length* of edge $e = uv$ of G then we can write $X_e = \min\left\{Y_{uv}, Y_{vu}\right\}$ where Y_{uv}, Y_{vu} are independent copies of the random variable Y where $\mathbf{Pr}\left\{Y \ge y\right\} = (1-y)^{1/2}$. The density of Y is close to $y/2$ for y close to zero. Now consider G_{k_ε} where the choices $\{v_1, v_2, \ldots, v_{k_\varepsilon}\}$ of vertex u are the k_ε edges of lowest weight Y_{uv} among $uv \in E(G)$. Now consider the total weight of the Hamilton cycle H posited by Theorem 2. The expected weight of an edge of H is at most $2 \times \frac{k_\varepsilon+1}{2\left(\frac{1}{2}+\varepsilon\right)n}$ and the corollary follows.

4 Long Paths: Proof of Theorem 4

Let D_k denote the directed graph with out-degree k defined by the vertex choices. Consider a Depth First Search (DFS) of D_k where we construct D_k as we go. At all times we keep a stack U of vertices which have been visited, but for which we have chosen fewer than k out-edges. T denotes the set of vertices that have not been visited by DFS. Each step of the algorithm begins with the top vertex u of U choosing one new out-edge. If the other end of the edge v lies in T (we call this a *hit*), we move v from T to the top of U.

When DFS returns to $v \in U$ and at this time v has chosen all of its k out-edges, we move v from U to S. In this way we partition V into

S - Vertices that have chosen all k of its out-edges.

U - Vertices that have been visited but have chosen fewer than k edges.

T - Unvisited vertices.

Key facts: Let h denote the number of hits at any time and let κ denote the number of times we have re-started the search i.e. selected a vertex in T after the stack S empties.

P1 $|S \cup U|$ increases by 1 for each hit, so $|S \cup U| \geq h$.

P2 More specifically, $|S \cup U| = h + \kappa - 1$.

P3 At all times $S \cup U$ contains a path which contains all of U.

The goal will be to prove that $|U| \geq (1 - 2\varepsilon)m$ at some point of the search, where ε is some arbitrarily small positive constant.

Lemma 15. *After εkm steps, i.e. after εkm edges have been chosen in total, the number of hits $h \geq (1 - \varepsilon)m$ w.h.p.*

Proof. Since $\delta(G_k) \geq k$, each tree component of G_k has at least k vertices, and at least k^2 edges must be chosen in order to complete the search of the component. Hence, after εkm edges have been chosen, at most $\varepsilon km/k^2 \leq \varepsilon m/2$ tree components have been found. This means that if $h \leq (1 - \varepsilon)m$ after εkm edges have been sent out, then **P2** implies that $|S \cup U| \leq (1 - \varepsilon/2)m$.

So if $h \leq (1 - \varepsilon)m$ each edge chosen by the top vertex u has probability at least $\frac{d(u)-|S\cup U|}{d(u)} \geq \varepsilon/2$ of making a hit. Hence,

$$\mathbf{Pr}\left\{h \leq (1 - \varepsilon)m \text{ after } \varepsilon km \text{ steps}\right\} \leq \mathbf{Pr}\left\{\text{Bin}(\varepsilon km, \varepsilon/2) \leq (1 - \varepsilon)m\right\} = o(1), \quad (12)$$

for $k \geq 2/\varepsilon^2$, by the Chernoff bounds. $\square$

We can now complete the proof of Theorem 4. By Lemma 15, after εkm edges have been chosen we have $|S \cup U| \geq (1 - \varepsilon)m$ w.h.p. For a vertex to be included in S, it must have chosen all of its edges. Hence, $|S| \leq \varepsilon km/k = \varepsilon m$, and we have $|U| \geq (1 - 2\varepsilon)m$. Finally observe that U is the set of vertices of a path of G_k. $\square$

5 Long Cycles: Proof of Theorem 5

Suppose now that we consider $G_{4k} = LR_k \cup DR_k \cup LB_k \cup DB_k$ where each for each vertex v and for each $c \in \{$"light red", "dark red", "light blue","dark blue"$\}$ the vertex v makes k choices of neighbor $N_c(v)$, distinct from any previous choices for this vertex. The edges $\{v, w\}, w \in N_c(v)$ are given the color c. Let LR_k, DR_k, LB_k, DB_k respectively be the graphs induced by the differently colored edges. We have by Theorem 4 that w.h.p. there is a path P of length at least $(1 - \varepsilon)m$ in the light red graph LR_k. At this point we start using a modification of DFS (denoted by $\Delta\Phi\Sigma$) and the differently colored choices to create a cycle.

We divide the steps into epochs $T_0, T_{00}, T_{01}, \ldots$, indexed by binary strings. We stop the search immediately if there is a high chance of finding a cycle of length at least

$(1-19\varepsilon)m$. If executed, epoch $T_{\boldsymbol{\iota}}, \boldsymbol{\iota} = 0***$ will extend the exploration tree by at least $(1-5\varepsilon)m$ vertices, unless an unlikely failure occurs. Theorem 4 provides T_0. In the remainder, we will assume $\boldsymbol{\iota} \neq 0$.

Epoch $T_{\boldsymbol{\iota}}$ will use light red colors if i has odd length and ends in a 0, dark red if i has even length and ends in a 0, light blue if i has odd length and ends in a 1, and dark blue if i has even length and ends in a 1. Epochs $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ (where $\boldsymbol{\iota}j$ denotes the string obtained by appending j to the end of $\boldsymbol{\iota}$) both start where $T_{\boldsymbol{\iota}}$ ends, and this coloring ensures that every vertex discovered in an epoch will initially have no adjacent edges in the color of the epoch.

During epoch $T_{\boldsymbol{\iota}}$ we maintain a stack of vertices $S_{\boldsymbol{\iota}}$. When discovered, a vertex is placed in one of the three sets $A_{\boldsymbol{\iota}}, B_{\boldsymbol{\iota}}, C_{\boldsymbol{\iota}}$, and simultaneously placed in $S_{\boldsymbol{\iota}}$ if it is placed in $A_{\boldsymbol{\iota}}$. Once placed, the vertex remains in its designated set even if it is removed from $S_{\boldsymbol{\iota}}$. Let $d_T(v,w)$ be the length of the unique path in the exploration tree T from v to w. We designate the set for v as follows.

$A_{\boldsymbol{\iota}}$ - v has less than $(1-2\varepsilon)d(v)$ G-neighbors in T.

$B_{\boldsymbol{\iota}}$ - v has at least $(1-2\varepsilon)d(v)$ G-neighbors in T, but less than $\varepsilon d(v)$ G-neighbors w such that $d_T(v,w) \geq (1-19\varepsilon)m$.

$C_{\boldsymbol{\iota}}$ - v has at least $(1-2\varepsilon)d(v)$ G-neighbors in T, and at least $\varepsilon d(v)$ G-neighbors w such that $d_T(v,w) \geq (1-19\varepsilon)m$.

At the initiation of epoch $T_{\boldsymbol{\iota}}$, a previous epoch will provide a set $T_{\boldsymbol{\iota}}^0$ of $3\varepsilon m$ vertices, as described below. Starting with $A_{\boldsymbol{\iota}} = B_{\boldsymbol{\iota}} = C_{\boldsymbol{\iota}} = \emptyset$, each vertex of $T_{\boldsymbol{\iota}}^0$ is placed in $A_{\boldsymbol{\iota}}, B_{\boldsymbol{\iota}}$ or $C_{\boldsymbol{\iota}}$ according to the rules above. Let $S_{\boldsymbol{\iota}} = A_{\boldsymbol{\iota}}$, ordered with the latest discovered vertex on top.

If at any point during $T_{\boldsymbol{\iota}}$ we have $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$, we immediately interrupt $\Delta\Phi\Sigma$ and use the vertices of $B_{\boldsymbol{\iota}}$ or $C_{\boldsymbol{\iota}}$ to find a cycle, as described below.

An epoch $T_{\boldsymbol{\iota}}$ consists of up to εkm steps, and each step begins with a $v \in A_{\boldsymbol{\iota}}$ at the top of the stack $S_{\boldsymbol{\iota}}$. This vertex is called *active*. If v has chosen k neighbors, remove v from the stack and perform the next step. Otherwise, let v randomly pick one neighbor w from $N_G(v)$. If $w \notin T$, then w is assigned to $A_{\boldsymbol{\iota}}, B_{\boldsymbol{\iota}}$ or $C_{\boldsymbol{\iota}}$ as described above. If $w \in A_{\boldsymbol{\iota}}$, perform the next step with w at the top of $S_{\boldsymbol{\iota}}$. If $w \in B_{\boldsymbol{\iota}} \cup C_{\boldsymbol{\iota}}$ perform the next step with the same v. If $w \in T$, perform the next step without placing w in $S_{\boldsymbol{\iota}}$.

The exploration tree T is built by adding to it any vertex found during $\Delta\Phi\Sigma$, along with the edge used to discover the vertex.

Note that unless $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$, we initially have $|A_{\boldsymbol{\iota}}| \geq \varepsilon m$, guaranteeing that εkm steps may be executed. Epoch $T_{\boldsymbol{\iota}}$ *succeeds* and is ended (possibly after fewer than εkm steps) if at some point we have $|A_{\boldsymbol{\iota}}| = (1-2\varepsilon)m$. If all εkm steps are executed and $|A_{\boldsymbol{\iota}}| < (1-2\varepsilon)m$, the epoch fails.

Lemma 16. *Epoch $T_{\boldsymbol{\iota}}$ succeeds with probability at least $1-e^{-\varepsilon^2 m/8}$, unless $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$ is reached.*

Proof. An epoch fails if less than $(1-3\varepsilon)m$ steps result in the active vertex choosing a neighbor outside T. Since the active vertex is always in $A_{\boldsymbol{\iota}}$, we have

$$\mathbf{Pr}\left\{T_{\boldsymbol{\iota}} \text{ finishes with } |A_{\boldsymbol{\iota}}| < (1-2\varepsilon)m\right\} \leq \mathbf{Pr}\left\{\mathrm{Bin}(\varepsilon km, 2\varepsilon) < (1-3\varepsilon)m\right\} \leq e^{-\varepsilon^2 m/8}$$

for $k \geq 1/2\varepsilon^2$, by Hoeffding's inequality. This proves the lemma. □

Ignoring the colors of the edges, an epoch produces a tree which is a subtree of T. Let $P_{\boldsymbol{\iota}}$ be the longest path of vertices in $A_{\boldsymbol{\iota}}$, and let $R_{\boldsymbol{\iota}}$ be the set of vertices discovered during $T_{\boldsymbol{\iota}}$ which are not in $P_{\boldsymbol{\iota}}$. If the epoch succeeds, $P_{\boldsymbol{\iota}}$ has length at least $(1-6\varepsilon)m$, and at most $3\varepsilon m$ vertices discovered during $T_{\boldsymbol{\iota}}$ are not on the path. Indeed, a vertex of $A_{\boldsymbol{\iota}}$ is outside $P_{\boldsymbol{\iota}}$ if and only if it has chosen all its k neighbors. Thus, the number of vertices not on the path is bounded by

$$|R_{\boldsymbol{\iota}}| \leq \frac{\varepsilon km}{k} + |B_{\boldsymbol{\iota}}| + |C_{\boldsymbol{\iota}}| < 3\varepsilon m.$$

If the epoch fails, the path $P_{\boldsymbol{\iota}}$ may be shorter, but $|R_{\boldsymbol{\iota}}|$ is still bounded by $3\varepsilon m$.

If $T_{\boldsymbol{\iota}}$ succeeds, the epochs $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ will be initiated at the end of $T_{\boldsymbol{\iota}}$, by letting $T^0_{\boldsymbol{\iota}0}$ and $T^0_{\boldsymbol{\iota}1}$ be the last $3\varepsilon m$ vertices discovered during $T_{\boldsymbol{\iota}}$. If $T_{\boldsymbol{\iota}}$ fails, $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ will not be initiated. The exploration tree T will resemble an unbalanced binary tree, in which each successful epoch gives rise to up to two new epochs. Epochs are ordered and $T_{\boldsymbol{\iota}_1}$ is initiated before $T_{\boldsymbol{\iota}_2}$ if and only if $\boldsymbol{\iota}_1 < \boldsymbol{\iota}_2$. Here let $\boldsymbol{\iota}_i = \boldsymbol{x}\boldsymbol{y}_i, i = 1,2$ where $\boldsymbol{x}$ is the longest common substring of $\boldsymbol{\iota}_1, \boldsymbol{\iota}_2$. We will have $\boldsymbol{\iota}_1 < \boldsymbol{\iota}_2$ if either $\boldsymbol{y_1}$ is the empty string or if $\boldsymbol{y_1}$ starts with 0 and $\boldsymbol{y_2}$ starts with 1.

Lemma 17. *W.h.p., $\Delta\Phi\Sigma$ will discover an epoch $T_{\boldsymbol{\iota}}$ having $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$.*

Proof. Suppose that no epoch ends with $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$. Under this assumption, each successful epoch $T_{\boldsymbol{\iota}}$ gives rise to $X'_{\boldsymbol{\iota}}$ new epochs. By Lemma 16, $X'_{\boldsymbol{\iota}}$ can be stochastically bounded from below by $X_{\boldsymbol{\iota}}$, where for some $c > 0$, $X_{\boldsymbol{\iota}} = 0$ with probability e^{-2cm}, $X_{\boldsymbol{\iota}} = 1$ with probability $2e^{-cm}(1-e^{-cm})$ and $X_{\boldsymbol{\iota}} = 2$ with probability $(1-e^{-cm})^2$. The number of successful epochs is then bounded from below by the total number of offspring in a Galton-Watson branching process with offspring distribution described by $X_{\boldsymbol{\iota}}$. The offspring distribution for this lower bound has generating function

$$G_m(s) = e^{-2cm} + 2se^{-cm}(1-e^{-cm}) + s^2(1-e^{-cm})^2.$$

Let s_m be the smallest fixed point $G_m(s_m) = s_m$. We have, with $\xi = e^{-cm}$,

$$s_m = \frac{1-2\xi(1-\xi) - [(1-2\xi(1-\xi))^2 - 4(1-\xi)^2\xi^2]^{1/2}}{2(1-\xi)^2} \to 0, \text{ as } m \to \infty.$$

Hence, the probability that the branching process never expires is at least $1-s_m$, which tends to 1.

The number of epochs is bounded by a finite number. Hence, the branching process cannot be infinite. This contradiction finishes the proof. □

We may now finish the proof of the theorem. Condition first on $\Delta\Phi\Sigma$ being stopped by an epoch $T_{\boldsymbol{\iota}}$ having $|C_{\boldsymbol{\iota}}| = \varepsilon m$. In this case, let each $v \in C_{\boldsymbol{\iota}}$ choose k neighbors using edges with the epoch's color. Each choice has probability at least ε of finding a cycle of length at least $(1-19\varepsilon)m$, by choosing a neighbor w such that $d_T(v,w) \geq (1-19\varepsilon)m$. The probability of not finding a cycle of length at least $(1-19\varepsilon)m$ is bounded by

$$(1-\varepsilon)^{\varepsilon k m} \to 0.$$

Now condition on $\Delta\Phi\Sigma$ being stopped by an epoch $T_{\boldsymbol{\iota}}$ having $|B_{\boldsymbol{\iota}}| = \varepsilon m$. Note that we must have $\boldsymbol{\iota} = \boldsymbol{\iota}'1$ for some $\boldsymbol{\iota}'$. Indeed, if $\boldsymbol{\iota} = \boldsymbol{\iota}'0$, then any v discovered in $\boldsymbol{\iota}$ must have at least $11\varepsilon d(v)$ G-neighbors at distance at least $(1-19\varepsilon)m$, at its time of discovery. If not, and $v \notin A_{\boldsymbol{\iota}}$ then it has at most $2\varepsilon d(v)$ G-neighbors outside T, at most $3\varepsilon d(v) + 3\varepsilon d(v)$ G-neighbors in $R_{\boldsymbol{\iota}} \cup R_{\boldsymbol{\iota}'}$. There are at most $(1-19\varepsilon)d(v)$ G-neighbors in $T \setminus (R_{\boldsymbol{\iota}} \cup R_{\boldsymbol{\iota}'})$ at distance less than $(1-19\varepsilon)d(v)$ and so there are at least $11\varepsilon d(v)$ G-neighbors in T at distance at least $(1-19\varepsilon)d(v)$ from v, which implies that $v \in C_{\boldsymbol{\iota}}$, contradiction.

Since the epoch produces a tree with at most m vertices, using the pigeonhole principle we can choose a $W \subseteq B_{\boldsymbol{\iota}}$ such that $|W| = \varepsilon^2 m$ and $d_T(v,w) \leq \varepsilon m$ for any $v, w \in W$.

Note also that $d(v) \leq 2m$ for any $v \in B_{\boldsymbol{\iota}}$. This can be seen as follows: For any $v \in W$ let $\rho_v \in T^0_{\boldsymbol{\iota}}$ be the vertex which minimizes $d_T(v, \rho_v)$. Note that we may have $\rho_v = v$. There are at most $|Q|$ G-neighbors of v on the path Q from v to ρ_v. Then note that there are at most $2((1-19\varepsilon)m - |Q|)$ G-neighbors of v on $T \setminus (Q \cup R_{\boldsymbol{\iota}} \cup R_{\boldsymbol{\iota}'} \cup R_{\boldsymbol{\iota}'0})$ that are within $(1-19\varepsilon)m$ of v. Here the factor 2 comes from counting G-neighbors in $T_{\boldsymbol{\iota}}$ and $T_{\boldsymbol{i}'0}$. So the maximum number of $w \in N_G(v) \cap T$ such that $d_T(v,w) \leq (1-19\varepsilon)m$ is bounded by

$$|Q| + 2((1-19\varepsilon)m - |Q|) + |R_{\boldsymbol{\iota}}| + |R_{\boldsymbol{\iota}'}| + |R_{\boldsymbol{\iota}'0}| \leq (2-29\varepsilon)m \tag{13}$$

Equation (13) then implies that $d(v) \leq (2-29\varepsilon)m + 3\varepsilon d(v)$.

Define an ordering on T by saying that $t_1 \leq t_2$ if t_1 was discovered before t_2 during $\Delta\Phi\Sigma$, or if $t_1 = t_2$. If $S \subseteq T'$, and $t \leq s$ for all $s \in S$, write $t \leq S$. Similarly define $\geq, >$ and $<$.

Let each $v \in W$ choose k neighbors in the color of epoch $T_{\boldsymbol{\iota}}$. We say that v is *good* if it chooses $v_1, v_2 \in P_{\boldsymbol{\iota}'}$ and $v_3 \in P_{\boldsymbol{\iota}'0}$ such that

$$d_T(v_1, v_2) + d_T(v_3, T^0_{\boldsymbol{\iota}}) + d_T(\rho_v, v) \geq (1-17\varepsilon)m$$

where $d_T(v_3, S) = \min_{s \in S} d_T(v_3, s)$. For each $v \in W$ define $n_0(v) = |N_G(v) \cap P_{\boldsymbol{\iota}} \setminus T^0_{\boldsymbol{\iota}}|$, $n_1(v) = |N_G(v) \cap P_{\boldsymbol{\iota}'} \setminus T^0_{\boldsymbol{\iota}}|$ and $n_2(v) = |N_G(v) \cap P_{\boldsymbol{\iota}'0} \setminus T^0_{\boldsymbol{\iota}}|$. Since $v \in B_{\boldsymbol{\iota}}$ we have

$$n_0(v) + n_1(v) + n_2(v) = |(N_G(v) \cap T) \setminus (R_{\boldsymbol{\iota}'} \cup R_{\boldsymbol{\iota}'0} \cup R_{\boldsymbol{\iota}} \cup T^0_{\boldsymbol{\iota}})| \geq (1-14\varepsilon)m.$$

Since the $n_0(v) + n_1(v)$ vertices of $N_G(v) \cup P_{\boldsymbol{\iota}} \cup P_{\boldsymbol{\iota}'} \setminus T^0_{\boldsymbol{\iota}}$ are on a path, we must have $n_0(v) + n_1(v) \leq (1-16\varepsilon)m$, otherwise v has $2\varepsilon m \geq \varepsilon d(v)$ neighbors at distance at least $(1-18\varepsilon)m$, contradicting $v \in B_{\boldsymbol{\iota}}$. This implies $n_2(v) \geq 2\varepsilon m$. Similarly, $n_1(v) \geq 2\varepsilon m$.

Fix a vertex $v \in W$ and define $V_1, V_2 \subseteq (N_G(v) \cap P_{\boldsymbol{\iota}'}) \setminus T^0_{\boldsymbol{\iota}}$ and $V_3 \subseteq (N_G(v) \cap P_{\boldsymbol{\iota}'0}) \setminus T^0_{\boldsymbol{\iota}}$, $|V_1| = |V_2| = |V_3| = \varepsilon m$ as follows. V_1 is the set of the first εm vertices of

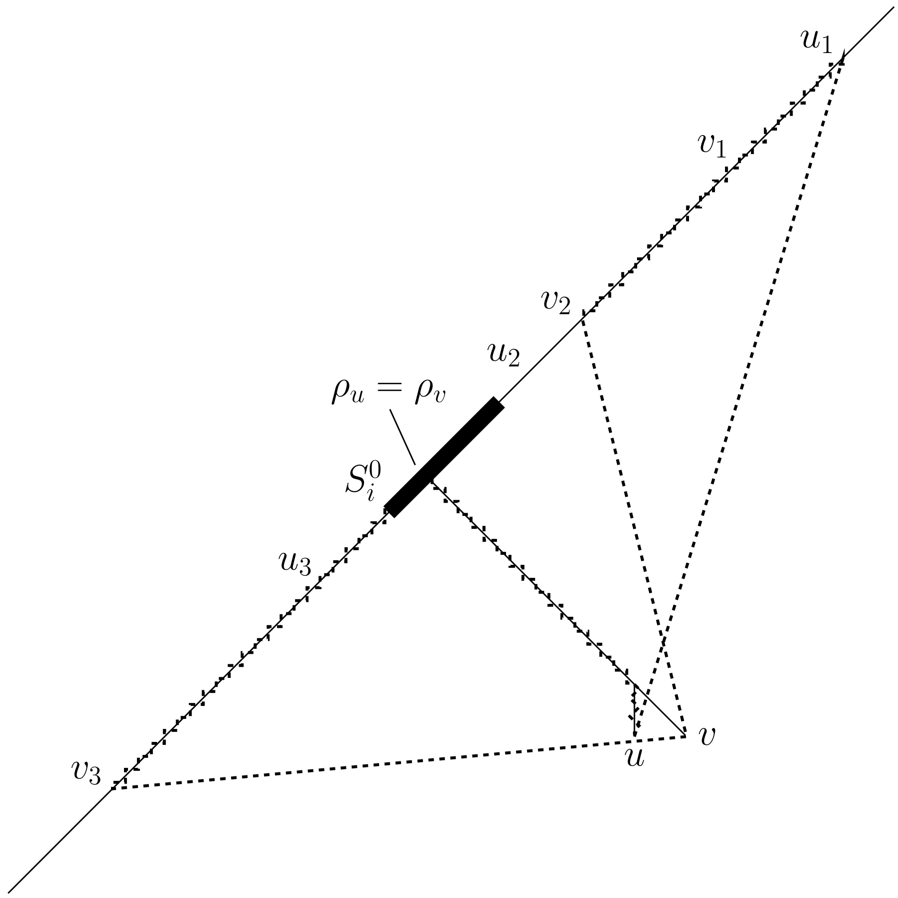

Figure 2: Example depiction of cycle found when $|B_{\iota}| = \varepsilon m$.

$N_G(v) \cap P_{\iota'}$ discovered during $\Delta\Phi\Sigma$. V_2 is the set of the last εm vertices of $N_G(v) \cap P_{\iota'}$ discovered before any vertex of T_{ι}^0. Lastly, V_3 consists of the εm last vertices discovered in $N_G(v) \cap P_{\iota'0}$. Since $n_1(v) \geq 2\varepsilon m$ and $n_2(v) \geq 2\varepsilon m$, the sets V_1, V_2, V_3 exist and are disjoint.

Since $d(v) \leq 2m$, the probability that v chooses $v_1 \in V_1, v_2 \in V_2$ and $v_3 \in V_3$ is at least $(\varepsilon/2)^3$. If this happens, we have

$$d_T(v_1, v_2) + d_T(v_3, T_{\iota}^0) + d_T(\rho_v, v) \geq n_1(v) - 2\varepsilon m + n_2(v) - \varepsilon m + n_3(v) \geq (1 - 17\varepsilon)m.$$

In other words, $v \in W$ is good with probability at least $(\varepsilon/2)^3$. Since $|W| = \varepsilon^2 m$, w.h.p. there exist two good vertices $u, v \in W$. Since $u, v \notin P_{\iota}$, the shortest path from ρ_v to v does not contain u, and the shortest path from ρ_u to u does not contain v. Also, by choice of W we have $d_T(\rho_u, u) \geq d_T(\rho_v, v) - 2\varepsilon m$. Suppose u and v pick $u_1 \leq u_2 \leq u_3$ and $v_1 \leq v_2 \leq v_3$, and w.l.o.g. suppose $d_T(u_1, v_2) \geq d_T(v_1, v_2)$. The cycle $(u, u_1, ..., v_2, v, v_3, ..., \rho_u, ..., u)$ has length

$$\begin{aligned} & 1 + d_T(u_1, v_2) + 1 + 1 + d_T(v_3, \rho_u) + d_T(\rho_u, u) \\ \geq \;& d_T(v_1, v_2) + d_T(v_3, T_{\iota}^0) + d_T(\rho_v, v) - 2\varepsilon m \\ \geq \;& (1 - 19\varepsilon)m. \end{aligned}$$

References

[1] T. Bohman and A.M. Frieze, Hamilton cycles in 3-out *Random Structures and Algorithms* 35 (2010) 393-417.

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

[3] A.M. Frieze, N. Goyal, L. Rademacher and S. Vempala, Expanders via Random Spanning Trees.

[4] A.M. Frieze, Maximum matchings in a class of random graphs, *Journal of Combinatorial Theory B* 40 (1986) 196-212.

[5] T. Fenner and A.M. Frieze, On the connectivity of random m-orientable graphs and digraphs, *Combinatorica* 2 (1982) 347-359.

[6] A. Frieze and M. Krivelevich, On the non-planarity of a random subgraph. *Combinatorics, Probability and Computing* 22 (2013) 722-732.

[7] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963) 13-30.

[8] M. Krivelevich, C. Lee and B. Sudakov, Long paths and cycles in random subgraphs of graphs with large minimum degree.

[9] M. Krivelevich, C. Lee and B. Sudakov, Robust Hamiltonicity of Dirac graphs, *Transactions of the American Mathematical Society* 366 (2014) 3095-3130.

[10] M. Krivelevich and W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, *SIAM Journal on Discrete Mathematics* 26 (2012) 964-982.

[11] O. Riordan, Long cycles in random subgraphs of graphs with large minimum degree.

[12] E. Shamir and E. Upfal, One-factor in random graphs based on vertex choice, *Discrete Mathematics* 41 (1982) 281-286.

[13] D.W. Walkup, Matchings in random regular bipartite graphs, *Discrete Mathematics* 31 (1980) 59-64.

[14] D.W. Walkup, On the expected value of a random assignment problem, *SIAM Journal on Computing* 8 (1979) 440-442.