How to avoid using the Regularity Lemma: Pósa's conjecture revisited

Ian Levitt [a], Gábor N. Sárközy [b,d], Endre Szemerédi [c,e,f]

[a] *Mathematics Department, Rutgers University, New Brunswick, NJ 08903, USA*

[b] *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA*

[c] *Computer Science Department, Rutgers University, New Brunswick, NJ 08903, USA*

[d] *Computer and Automation Research Institute, Hungarian Academy of Sciences, P.O. Box 63, Budapest H-1518, Hungary*

[e] *Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary*

[f] *Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

ARTICLE INFO

ABSTRACT

In this paper we investigate how the use of the Regularity Lemma and the Blow-up Lemma can be avoided in certain extremal problems of dense graphs. We present the ideas for the following well-known Pósa conjecture on the square of a Hamiltonian cycle. In 1962 Pósa conjectured that any graph G of order n and minimum degree at least $\frac{2}{3}n$ contains the square of a Hamiltonian cycle. In an earlier paper we proved this conjecture with the use of the Regularity Lemma–Blow-up Lemma method for $n \geq n_0$ where n_0 is very large. Here we present another proof (and a general method) that avoids the use of the Regularity Lemma and thus the resulting n_0 is much smaller.

Published by Elsevier B.V.

1. Introduction

1.1. Notations and definitions

$V(G)$ and $E(G)$ denote the vertex-set and the edge-set of the graph G. (A, B, E) denotes a bipartite graph $G = (V, E)$, where $V = A + B$, and $E \subset A \times B$. For a graph G and a subset U of its vertices, $G|_U$ is the restriction of G to U. $N(v)$ is the set of neighbors of $v \in V$, and $N_S(v)$ is the set of neighbors of $v \in V \cap S$. Hence the size of $N(v)$ is $|N(v)| = deg(v) = deg_G(v)$, the degree of v. $\delta(G)$ stands for the minimum, $\Delta(G)$ for the maximum and $\overline{d}(G) = \frac{1}{|V(G)|}\sum_{v \in V(G)} deg(v)$ for the average degree in G. $K_r(t)$ is the balanced complete r-partite graph with color classes of size t. We write $N(p_1, p_2, \ldots) = \cap_i N(p_i)$, the set of common neighbors. When A, B are subsets of $V(G)$, we denote by $e(A, B)$ the number of edges of G with one endpoint in A and the other in B. In particular, we write $deg(v, U) = e(\{v\}, U)$ for the number of edges from v to U. For non-empty A and B,

$$d(A, B) = \frac{e(A, B)}{|A||B|}$$

is the *density* of the graph between A and B. In particular, we write $d(A) = d(A, A) = 2|E(G|_A)|/|A|^2$.

1.2. Powers of cycles

The kth *power* of C is the graph obtained from C by joining every pair of vertices at a distance at most k in C. Let G be a graph on $n \geq 3$ vertices. A classical result of Dirac [1] asserts that if $\delta(G) \geq n/2$, then G contains a Hamiltonian cycle. As a natural generalization of Dirac's theorem, in 1962 Pósa (see Erdős [2]) conjectured the following:

Conjecture 1 (*Pósa*). *Let G be a graph on n vertices. If $\delta(G) \geq \frac{2}{3}n$, then G contains the square of a Hamiltonian cycle.*

Later in 1974 Seymour [20] generalized this conjecture:

Conjecture 2 (*Seymour*). *Let G be a graph on n vertices. If $\delta(G) \geq \frac{k}{k+1}n$, then G contains the kth power of a Hamiltonian cycle.*

Seymour indicated the difficulty of the conjecture by observing that the truth of this conjecture would imply the remarkably difficult Hajnal–Szemerédi Theorem [11], namely that if $\Delta(G) < r$, then G is r colorable such that the sizes of the color classes are all $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$.

These problems received significant attention. In the direction of Conjecture 1, first Jacobson (unpublished) showed that if $\delta(G) \geq 5n/6$, then the conclusion of the conjecture holds. Faudree, Gould, Jacobson and Schelp [9] confirmed the conclusion with $\delta(G) \geq (3/4 + \varepsilon)n + C(\varepsilon)$. Later the same authors improved this to $\delta(G) \geq 3n/4$. By using a result in [10], Häggkvist (unpublished) gave a very simple proof in the case $\delta(G) \geq 1 + 3n/4$ and $n \equiv 0 \pmod 4$. Fan and Häggkvist in [3] lowered the bound to $\delta(G) \geq 5n/7$. Fan and Kierstead improved this further to $\delta(G) \geq (17n + 9)/24$ in [4], and Faudree, Gould and Jacobson [8] to $\delta(G) \geq 7n/10$. Then Fan and Kierstead [5] improved the condition to the almost optimal $\delta(G) \geq \left(\frac{2}{3} + \varepsilon\right)n + C(\varepsilon)$. They also proved [6] that already $\delta(G) \geq (2n - 1)/3$ is sufficient for the existence of the square of a Hamiltonian *path*. Furthermore, they also proved [7] that if $\delta(G) \geq 2n/3$ and G contains the square of a cycle with length greater than $2n/3$, then G contains the square of a Hamiltonian cycle. Finally, Kierstead and Quintana [13] proved that if $\delta(G) \geq 2n/3$ and G contains a maximal 4-clique, then G contains the square of a Hamiltonian cycle.

For Conjecture 2, in the above mentioned paper of Faudree et al. in [9], it is proved that for any $\varepsilon > 0$ and positive integer k there is a C such that if an n-graph G satisfies

$$\delta(G) \geq \left(\frac{2k - 1}{2k} + \varepsilon\right)n + C,$$

then G contains the kth power of a Hamiltonian cycle.

Using the Regularity Lemma–Blow-up Lemma method first in [16] we proved Conjecture 2 in asymptotic form, then in [14,17] we proved both conjectures for $n \geq n_0$. The proofs used the Regularity Lemma [21], the Blow-up Lemma [15,18] and the Hajnal–Szemerédi Theorem [11]. Since the proofs used the Regularity Lemma the resulting n_0 is very large (it involves a tower function). Apart from the use of the Extremal Condition, the n_0 that is demonstrated in this paper is polynomial in the parameter α (see below). It is still quite large however, on the order of 10^{24}. While it is possible to optimize n_0 and decrease its value significantly with a more careful argument and treatment of the Extremal Condition, the purpose of this paper is to present another proof (and a general proof technique) for $k = 2$ that avoids the use of the Regularity Lemma.

Theorem 1. *There exists a natural number n_0 such that if a graph G has order n with $n \geq n_0$ and $\delta(G) \geq \frac{2}{3}n$, then G contains the square of a Hamiltonian cycle.*

2. Outline of the proof

We will follow the same rough outline (connecting-absorbing-reservoir) as in [19]; however, the main ingredient there, the Regularity Lemma, will be replaced with more elementary arguments here.

We will use the following main parameter

$$\alpha = \frac{1}{10}. \tag{1}$$

We assume throughout the paper that n is sufficiently large.

Let us consider a graph G of order n with

$$\delta(G) \geq \frac{2}{3}n. \tag{2}$$

We must show that G contains the square of a Hamiltonian cycle.

In [14] our proof was divided into two main cases, the extremal case when G satisfies the following so-called extremal condition and the non-extremal case when this condition is not satisfied.

Extremal Condition (EC) with parameter α: *There exist (not necessarily disjoint) $A, B \subset V(G)$ such that*

- $\left(\frac{1}{3} - \alpha\right)n \leq |A|, |B| \leq \frac{1}{3}n$, *and*
- $d(A, B) < \alpha$.

In the extremal case in [14] our proof did not use the Regularity Lemma, thus we can use that part of the proof here again.

Lemma 2 (*Lemma 12 in [14]*). *There exists a natural number n_1 such that if a graph G has order n with $n \geq n_1$, $\delta(G) \geq \frac{2}{3}n$ and G satisfies the extremal condition EC with parameter α, then G contains the square of a Hamiltonian cycle.*

Hence we may assume that our graph G does not satisfy the extremal condition EC with parameter α. In this case our proof technique will follow the same outline (and notation) as in [19]. First in Section 3 we will prove the auxiliary Connecting Lemma (stated below as Lemma 5) that claims that any two disjoint ordered pairs of vertices can be connected by a short square-path. Then using the Connecting Lemma and the probabilistic method in Section 4 we will construct a "not too long" *absorbing* square-path P_A that will have the remarkable property that *every* "not too large" subset of vertices can be absorbed into this square-path. Thus if this P_A will be a part of a square-cycle C that contains "most" of the vertices already, then immediately P_A (and thus C) absorbs the leftover vertices and we have a Hamiltonian square-cycle. This is a significant simplification of our proof technique from [14], where the corresponding step in Section 6 was quite complicated.

Thus our goal is to construct a square-cycle C that contains the absorbing path P_A and most of the other vertices. For this purpose we will need another technical lemma in Section 5, the Reservoir Lemma, which allows us to use the Connecting Lemma (through the reservoir) even if some of the vertices are already occupied by the square-cycle we are building. Finally in the main part of the proof in Section 6 we will show that unless C contains most of the vertices already, we can extend it by using leftover vertices that are not from the reservoir or the absorbing path. This is where, in contrary to the proof in [19], we are able to achieve this goal without the use of the Regularity Lemma, but using more elementary arguments. Thus the main point of the present paper is that the proof method of [19] can be adapted into a method that avoids the use of the Regularity Lemma. We believe that this new approach (although some of the arguments are problem-specific) could be successful for other well-known extremal problems where the Regularity Lemma–Blow-up Lemma method has been used (e.g. Conjecture 2 for $k > 2$ or the main Dirac-type result of [19] itself).

3. Connecting

A k-square-path (or simply a k-path) in G is a sequence of vertices $\{v_1, v_2, \ldots, v_k\}$ such that $\{v_i, v_{i+1}\} \in E(G)$ for each $1 \le i \le k-1$ and $\{v_i, v_{i+2}\} \in E(G)$ for each $1 \le i \le k-2$. We say that P connects the ordered pairs (v_1, v_2) and (v_k, v_{k-1}) and these will be called the endpairs of P. Thus an endpair (a, b) is an ordered pair, a is the first (or the last) vertex on the path, and b is the second (or the second-to-last) vertex on the path. We will often call a square-path simply a path.

For two paths P and Q, let (a, b) be an endpair of P and (b, a) be an endpair of Q, and assume that $V(P) \cap V(Q) = \{a, b\}$. By $P \circ Q$ we denote the path obtained (in a unique way) as a concatenation of P and Q. We can extend this definition to more than two paths. The Connecting Lemma claims that two disjoint ordered pairs can be connected by a short path.

Lemma 3 (*Connecting Lemma*). *For every two disjoint ordered edges of G, (a, b) and (c, d), there is a k-path, $k \le \frac{10}{\alpha^4}$, which connects (a, b) and (c, d). Furthermore, this statement remains true even if at most $\alpha^9 n$ vertices are forbidden to be used on this connecting path.*

Proof. We will build a similar cascade structure as in the proof of the Connecting Lemma in [19]. We construct sets $A_0, A_1, A_2, \ldots$ and bipartite graphs $G_1, G_2, \ldots$, where $V(G_i) = A_{i-1} \cup A_i$, as follows. Let $A_0 = \{b\}$ and $A_1 = \{x | (a, x), (b, x) \in E(G)\}$ and let G_1 be the star with b as the center and A_1 as the set of its leaves. Note that $|A_1| \ge n/3$. Further, let

$$A_2' = \{y | \exists x \in A_1 \text{ such that } (b, y), (x, y) \in E(G)\}$$

and

$$G_2' = \{(x, y) | x \in A_1, y \in A_2', (b, y), (x, y) \in E(G)\}.$$

Then for every edge $(x, y) \in G_2'$ that is disjoint from (a, b) the vertices (a, b, x, y) form a 4-path in G. Furthermore, for each $x \in A_1$, we have $deg_{G_2'}(x) \ge n/3$. Let

$$A_2^0 = \{y \in A_2' | deg_{G_2'}(y) < \alpha^4 n\}, \qquad A_2 = A_2' \setminus A_2^0 \quad \text{and} \quad G_2 = G_2'[A_1 \cup A_2].$$

Assume that we have constructed $A_0, A_1, \ldots, A_j$ and $G_1, \ldots, G_j, j \ge 2$ already. To construct A_{j+1} and G_{j+1} we do the following. First for every $y \in A_j$ we consider the auxiliary bipartite graph B_y^j between $N_{G_j}(y)$ and $V(G)$, where a pair $(x, z) \in E(B_y^j)$ for $x \in N_{G_j}(y)$, $z \in V(G)$ if $(x, z), (y, z) \in E(G)$. Define

$$A_{j+1}' = \{z | \exists y \in A_j \text{ such that } deg_{B_y^j}(z, N_{G_j}(y)) \ge \alpha^8 n\}$$

and

$$G_{j+1}' = \{(y, z) | y \in A_j, deg_{B_y^j}(z, N_{G_j}(y)) \ge \alpha^8 n\}.$$

Finally, let

$$A_{j+1}^0 = \{z \in A_{j+1}' | deg_{G_{j+1}'}(z) < \alpha^4 n\},$$

$$A_{j+1} = A_{j+1}' \setminus A_{j+1}^0 \quad \text{and} \quad G_{j+1} = G_{j+1}'[A_j \cup A_{j+1}].$$

We call the entire structure $A_0, A_1, A_2, \ldots$ along with the bipartite graphs $G_1, G_2, \ldots$ an (a, b)-cascade. Notice that some of the sets A_j may intersect. For the sake of the construction we treat them as disjoint. Note also that we had to change the

construction slightly for $j \geq 3$ and require $deg_{B_y^j}(z) \geq \alpha^8 n$ to make sure that we can always return from any edge of G_j back to (a, b) by a legitimate square-path on which all the vertices are distinct, even if at most $\alpha^9 n$ vertices are forbidden.

A vertex $y \in A_j$ is called *heavy* if $deg_{G_j}(y) \geq (1/3 + \alpha^4)n$. $\quad \square$

Claim 4. *There exists an index $j \leq j_0 = \lceil \frac{4}{\alpha^4} \rceil + 2$ such that A_j contains at least $\alpha^4 n$ heavy vertices.*

For the proof of this claim, first we prove that for every $j \geq 2$ and for every $y \in A_j$ we have

$$deg_{G'_{j+1}}(y) \geq (1/3 - \alpha^4)n. \tag{3}$$

Indeed, let s be the number of vertices $z \in V(G)$ with $deg_{B_y^j}(z) < \alpha^8 n$. Then

$$s\alpha^8 n + (n - s)|N_{G_j}(y)| \geq |E(B_y^j)| \geq |N_{G_j}(y)|n/3.$$

From this using $|N_{G_j}(y)| = deg_{G_j}(y) \geq \alpha^4 n$ and $s \leq n$, we get

$$n - s \geq n/3 - \frac{s\alpha^8 n}{|N_{G_j}(y)|} \geq n/3 - \alpha^4 n,$$

proving (3). Note also that the total number of edges of G'_{j+1} incident to the exceptional vertices in A^0_{j+1} is smaller than $\alpha^4 n^2$.

Let us look at the sequence of sets $A_1, A_2, \ldots$, where we have $|A_1| \geq n/3$. Clearly we must have a $j \leq \lceil \frac{4}{\alpha^4} \rceil$ for which

$$|A_{j+1}|, |A_{j+2}| \leq (1 + \alpha^4)|A_j|. \tag{4}$$

Indeed, if $j = 1$ does not satisfy (4), then either A_2 or A_3 (say A_3) has size at least $(1 + \alpha^4)|A_1| \geq (1 + \alpha^4)n/3$. If $j = 3$ does not satisfy (4), then either A_4 or A_5 (say A_5) has size at least $(1 + \alpha^4)|A_3| \geq (1 + \alpha^4)n/3$. Continuing in this fashion, in each step we add at least $\alpha^4 n/3$ new vertices to A_1, so in at most $\lceil \frac{2n/3}{\alpha^4 n/3} \rceil = \lceil \frac{2}{\alpha^4} \rceil$ steps we get a set A_j with more than n vertices, a contradiction.

Furthermore, we may assume that for this j in addition to (4) the following holds as well

$$|A_{j+1}| \geq (1 - \alpha^3)|A_j|. \tag{5}$$

Otherwise A_{j+1} would contain at least $\alpha^4 n$ heavy vertices and we would be finished with the proof of Claim 4. Indeed, suppose not. On the one hand from the above we have

$$|E(G_{j+1})| \geq |A_j|n/3 - 2\alpha^4 n^2, \tag{6}$$

but on the other hand using (5) we would get

$$\begin{aligned}|E(G_{j+1})| &\leq \alpha^4 n|A_j| + |A_{j+1}|(1/3 + \alpha^4)n \\ &\leq \alpha^4 n|A_j| + (1 - \alpha^3)(1/3 + \alpha^4)|A_j|n \leq |A_j|n/3 - \alpha^3|A_j|n/3 + 2\alpha^4|A_j|n,\end{aligned}$$

a contradiction (using (1)).

Thus we may assume that there is a $j \leq \lceil \frac{4}{\alpha^4} \rceil$ for which both (4) and (5) hold. We fix this j. We will show that A_{j+2} contains at least $\alpha^4 n$ heavy vertices as desired in the claim. For this purpose first we show that A_{j+1} contains at least $\alpha n/2$ vertices z for which

$$deg_{G_{j+1}}(z) \geq (1/3 - \alpha)n. \tag{7}$$

Otherwise, similarly as above using (4) we would get

$$\begin{aligned}|E(G_{j+1})| &\leq \alpha n|A_j|/2 + |A_{j+1}|(1/3 - \alpha)n \\ &\leq \alpha n|A_j|/2 + (1 + \alpha^4)(1/3 - \alpha)|A_j|n \leq |A_j|n/3 - \alpha|A_j|n/2 + \alpha^4|A_j|n/3,\end{aligned}$$

a contradiction with (6) (using (1)).

Consider a vertex $z \in A_{j+1}$ satisfying (7). Next we show that

$$deg_{G'_{j+2}}(z) \geq (1/3 + \alpha)n. \tag{8}$$

Indeed, otherwise let $A \subset N_{G_{j+1}}(z), B \subset N_G(z) \setminus N_{G'_{j+2}}(z)$ be arbitrary subsets with sizes $(1/3 - \alpha)n \leq |A|, |B| \leq n/3$ (this is possible as both of these sets have size at least $(1/3 - \alpha)n$). Since G does not satisfy EC with parameter α, we have $d(A, B) \geq \alpha$. In particular, we can pick a vertex $w \in B$ with $deg_G(w, A) \geq \alpha|A| \gg \alpha^8 n$, a contradiction, since in this case w would belong to $N_{G'_{j+2}}(z)$ by definition.

Thus we get from (3) and (8)

$$|E(G_{j+2})| \geq |A_{j+1}|n/3 + \alpha^2 n^{2/2} - 2\alpha^4 n^2. \tag{9}$$

However, this implies that in A_{j+2} we must have at least $\alpha^4 n$ heavy vertices, and thus proving the claim. Indeed, otherwise using (1), (4) and (5) we would get

$$
\begin{aligned}
|E(G_{j+2})| &\le \alpha^4 n |A_{j+1}| + |A_{j+2}|(1/3 + \alpha^4)n \\
&\le \alpha^4 n |A_{j+1}| + (1 + 2\alpha^3)(1 + \alpha^4)(1/3 + \alpha^4)|A_{j+1}|n \le |A_{j+1}|n/3 + \alpha^3 |A_{j+1}|n,
\end{aligned}
$$

a contradiction with (9).

Now to finish the proof of the Connecting Lemma, given two disjoint ordered edges of G, (a, b) and (c, d), we consider the (a, b)-cascade $(A_j^{(1)}, G_j^{(1)})$ and the (c, d)-cascade $(A_j^{(2)}, G_j^{(2)})$. For $i = 1, 2$, let $A_{j(i)}^{(i)}$ be the set that contains many $(\ge \alpha^4 n)$ heavy vertices as guaranteed by Claim 4. An easy averaging argument shows that there must be many $(\ge \alpha n)$ vertices $u \in V(G)$ such that u has many $(\ge \alpha^5 n)$ heavy neighbors in both $A_{j(i)}^{(i)}$, $i = 1, 2$. Consider one such a u, a heavy neighbor $h^{(1)}$ of u in $A_{j(1)}^{(1)}$ and a heavy neighbor $h^{(2)}$ of u in $A_{j(2)}^{(2)}$. It is easy to see from the definition that we have

$$
deg_{G'_{j(1)+1}}(h^{(1)}), \; deg_{G'_{j(2)+1}}(h^{(2)}) \ge 2n/3,
$$

since $h^{(1)}$ and $h^{(2)}$ are heavy vertices.

Let $A \subset N_G(u) \cap N_{G'_{j(1)+1}}(h^{(1)})$, $B \subset N_G(u) \cap N_{G'_{j(2)+1}}(h^{(2)})$ be arbitrary subsets with sizes $\lfloor n/3 \rfloor$ (this is possible as both of these sets have size at least $n/3$). Since G does not satisfy EC with parameter α, we have $d(A, B) \ge \alpha$. In particular, we can pick an edge $(v^{(1)}, v^{(2)})$ with $v^{(1)} \in A$ and $v^{(2)} \in B$. By the definition of the (a, b)-cascade, there is a $(j(1) + 3)$-path $P^{(1)}$ connecting (a, b) and $(v^{(1)}, h^{(1)})$ and by the definition of the (c, d)-cascade, there is a $(j(2) + 3)$-path $P^{(2)}$ connecting (c, d) and $(v^{(2)}, h^{(2)})$. By putting together $P^{(1)}$ and $P^{(2)}$ and including u in the middle we get a k-path connecting (a, b) and (c, d) with

$$
k = (j(1) + 3) + (j(2) + 3) + 2 \le 2(j_0 + 4) \le \frac{10}{\alpha^4}.
$$

Furthermore, the condition $deg_{B_y^j}(z) \ge \alpha^8 n$ guarantees that the proof goes through (and we can find a connecting path) even if we have a set of at most $\alpha^9 n$ forbidden vertices.

4. Absorbing

Again we follow the method of [19], we just have to adapt the ideas to square-cycles in graphs instead of tight cycles in hypergraphs. For the sake of completeness we give the details here again. As we sketched above the Absorbing Lemma claims that we can construct a "not too long" *absorbing* path A that absorbs *every* "not too large" subset of vertices.

Lemma 5 (*Absorbing Lemma*). *There is an l-path P_A in G with $l \le \alpha^9 n$, such that for every subset $U \subset V(G) \setminus V(P_A)$ of size at most $\alpha^{20} n$ there is a path P_{A_U} in G with $V(P_{A_U}) = V(P_A) \cup U$ and such that P_{A_U} has the same endpairs as P_A.*

Proof. Given a vertex $v \in V(G)$ we say that an (ordered) 5-tuple of vertices (x, a, b, c, d) absorbs v if these 5 vertices are all neighbors of v, the vertices $\{a, b, c, d\}$ are all neighbors of x and (a, b, c, d) is a (simple) path in G. Indeed, in this case the (square-)path (a, b, x, c, d) *may* absorb v to get the extended path (a, b, v, x, c, d). Note that both paths have the same endpairs. First we show that for every $v \in V(G)$ there are many 5-tuples absorbing v. □

Claim 6. *For every $v \in V(G)$ there are at least $\frac{1}{2(6^4)}\alpha^4 n^5$ 5-tuples absorbing v.*

Indeed, let us consider an arbitrary $v \in V(G)$. We can choose x as an arbitrary neighbor of v. Since G satisfies (2) we can choose x in at least $\frac{2}{3}n$ different ways. Consider the common neighbors of v and x, $N(v, x)$. We know from (2) that $|N(v, x)| \ge \frac{n}{3}$. Keep a subset $N \subseteq N(v, x)$ with $|N| = \lfloor \frac{n}{3} \rfloor$. Since G does not satisfy the extremal condition EC with parameter α, we know that $d(N) = d(N, N) \ge \alpha$. Then we have $\overline{d}(G|_N) \ge \alpha|N|$, and so we can choose a subgraph H of $G|_N$ with $\delta(H) > \frac{\alpha}{2}|N|$. In particular, we also have $|V(H)| \ge \frac{\alpha}{2}|N|$. Let a be an arbitrary vertex of H (at least $\frac{\alpha}{2}|N|$ different choices), let b be an arbitrary neighbor of a in H (at least $\frac{\alpha}{2}|N|$ different choices), let c be an arbitrary neighbor of b in H that is different from a (at least $\frac{\alpha}{2}|N| - 1$ different choices), and finally let d be an arbitrary neighbor of c in H that is different from a and b (at least $\frac{\alpha}{2}|N| - 2$ different choices). Then (x, a, b, c, d) is a good 5-tuple that absorbs v. The number of ways we can select (x, a, b, c, d) from the above is at least

$$
\frac{2}{3}n \frac{\alpha}{2}|N| \frac{\alpha}{2}|N| \left(\frac{\alpha}{2}|N| - 1\right)\left(\frac{\alpha}{2}N - 2\right) \ge \frac{1}{2(6^4)}\alpha^4 n^5
$$

(for sufficiently large n), finishing the proof of the claim.

For each $v \in V(G)$, let $\mathcal{A}_v$ be the family of all 5-tuples absorbing v. The next claim can be proved by an application of the probabilistic method.

Claim 7. *There exists a family $\mathcal{F}$ of at most $2\alpha^{14}n$ disjoint, absorbing 5-tuples of vertices of G such that for every $v \in V(G)$ we have $|\mathcal{A}_v \cap \mathcal{F}| > \alpha^{20}n$.*

For this purpose we first select a family $\mathcal{F}'$ of 5-tuples at random by including each of $n(n-1)(n-2)(n-3)(n-4) \sim n^5$ of them independently with probability $\alpha^{14} n^{-4}$ (some of the selected 5-tuples may not be absorbing at all). Using Chernoff's inequality (see, e.g. [12]) with probability $1 - o(1)$, as $n \to \infty$, we have

- $|\mathcal{F}'| < 2\alpha^{14} n$,
- for each $v \in V(G)$, $|\mathcal{A}_v \cap \mathcal{F}'| \geq \frac{1}{3(6^4)} \alpha^{18} n$.

Furthermore, the expected number of intersecting pairs of 5-tuples in $\mathcal{F}'$ is at most

$$n^5 \times 5 \times 5 \times n^4 \times (\alpha^{14} n^{-4})^2 = 25\alpha^{28} n,$$

and thus, by Markov's inequality, with probability at least $1/26$,

- there are at most $26\alpha^{28} n$ pairs of intersecting 5-tuples in $\mathcal{F}'$.

Thus with positive probability, a random family $\mathcal{F}'$ satisfies all the three properties above. Thus there exists one such family, for simplicity, we also denote this family by $\mathcal{F}'$. From $\mathcal{F}'$ we delete all 5-tuples that intersect other 5-tuples and all 5-tuples that are not absorbing at all. Let us denote by $\mathcal{F}$ the remaining subfamily. Then $\mathcal{F}$ consists of disjoint, absorbing 5-tuples such that for each $v \in V(G)$ we have using (1)

$$|\mathcal{A}_v \cap \mathcal{F}| \geq \frac{1}{3(6^4)} \alpha^{18} n - 52\alpha^{28} n > \frac{1}{4(6^4)} \alpha^{18} n > \alpha^{20} n,$$

proving Claim 7.

Let $f = |\mathcal{F}|$, let $F_1, \ldots, F_f$ be the 5-tuples in $\mathcal{F}$ and let $F = \cup_{i=1}^{f} F_i$. Since for each $i = 1, \ldots, f$, F_i is absorbing for at least one vertex $v \in V(G)$, F_i spans a 5-path. Our next task is to connect all these 5-paths into one, not too long absorbing path P_A. For this purpose, we will apply the Connecting Lemma (Lemma 3) repeatedly, and for each $i = 1, \ldots, f - 1$ we will connect the endpairs of F_i and F_{i+1} by a short path. Thus we get the following claim.

Claim 8. *There exists a path P_A in G of the form*

$$P_A = F_1 \circ C_1 \circ \ldots \circ F_{f-1} \circ C_{f-1} \circ F_f,$$

where the paths $C_1, \ldots, C_{f-1}$ each have at most $\frac{10}{\alpha^4}$ vertices.

Indeed, we apply Lemma 3 to connect F_1 and F_2, we apply Lemma 3 again to connect F_2 and F_3, etc. finally we apply Lemma 3 to connect F_{f-1} and F_f. Note that Lemma 3 can always be applied as the set of forbidden vertices (vertices on the part of P_A that is constructed already and vertices in F) has size at most

$$f\left(\frac{10}{\alpha^4} + 5\right) \leq 2\alpha^{14} \frac{11}{\alpha^4} n \leq \alpha^9 n.$$

Thus we connected all paths in $\mathcal{F}$ into one path of length at most $\alpha^9 n$. It remains to show that P_A has the absorbing property. Let $U \subset V \setminus V(P_A)$ of size at most $\alpha^{20} n$. Since for every $v \in U$ we have $|\mathcal{A}_v \cap \mathcal{F}| > \alpha^{20} n$, we can insert all vertices of U into P_A one by one, each time using a new absorbing 5-tuple.

5. The reservoir

In the Reservoir Lemma we will set aside some vertices that we can always use for connecting even if the other vertices are occupied already.

Lemma 9 (*Reservoir Lemma*). *For every subset $W \subset V(G)$, $|W| \leq \alpha^9 n$, there exists a subset $R \subset V(G) \setminus W$ (a reservoir) such that $|R| = \lfloor \alpha^{20} n/2 \rfloor$ and*

$$deg_G(x, R) \geq (2/3 - \alpha^{10})|R| \quad \text{for every } x \in V(G). \tag{10}$$

Proof. Set $r = \lfloor \alpha^{20} n/2 \rfloor$. We choose R randomly out of all $\binom{n - |W|}{r}$ possibilities and apply the probabilistic method again. By Chernoff's bound again, for sufficiently large n, (10) will be true for R with high probability. Then we can fix a choice of R for which (10) is true. $\square$

Then indeed, we can connect through the reservoir.

Lemma 10 (*Reservoir-Connecting Lemma*). *For every two disjoint ordered edges of G, (a, b) and (c, d), there is a k-path in $R \cup \{a, b, c, d\}$, $k \leq \frac{10}{\alpha^4}$, which connects (a, b) and (c, d). Furthermore, this statement remains true even if at most $\alpha^9|R|$ vertices of R are forbidden to be used on this connecting path.*

Proof. Indeed, since by (10) inside R we have almost the same degree condition as in G, the proof of the Connecting Lemma goes through inside R, the slight loss in the minimum degree is not going to create any problems. Note also that we may assume that $G|_R$ does not satisfy the EC with parameter α as this is true with high probability. $\square$

6. The proof of Theorem 1

We start with the outline of the proof.

Step 1: By applying the Absorbing Lemma (Lemma 5), we find an absorbing path P_A with $|P_A| \leq \alpha^9 n$.

Step 2: By applying the Reservoir Lemma (Lemma 9), we set aside a reservoir $R \subset V(G) \setminus V(P_A)$ with $|R| = \lfloor \alpha^{20} n/2 \rfloor$.

Step 3: We find a (square-)cycle C in G that contains P_A as a subpath, all but at most $\alpha^{20} n/2$ vertices of $V(G) \setminus (V(P_A) \cup R)$ (denote the set of these missing vertices by T) and some vertices of R (denote the set of remaining vertices in R by R'). Note that $|R' \cup T| \leq \alpha^{20} n$.

Step 4: Using the absorbing property of P_A, insert $R' \cup T$ into C, resulting in a Hamiltonian cycle of G.

It remains to explain Step 3 in the outline above. The rest of the paper contains the construction of this C. We start with an arbitrary path P in G that starts with P_A as a subpath. Then we will gradually extend this P (sometimes with the use of the Reservoir-Connecting Lemma, so using vertices from the reservoir R) until it contains all but at most $\alpha^{20} n/2$ vertices of $V(G) \setminus (V(A) \cup R)$. We connect the two endpairs of P through the reservoir by applying the Reservoir-Connecting Lemma one more time to get the cycle C that is desired in Step 3. Thus we only have to show how to extend P until it contains all but at most $\alpha^{20} n/2$ vertices of $V(G) \setminus (V(P_A) \cup R)$.

Denote by P' the square-path without the absorbing path, $V(P') = V(P) \setminus V(P_A)$, and set $m := |V(P')|$. Throughout the paper, we will represent the neighborhood on P' of a vertex $a \in T$ by a bitstring of length $|P'|$, indexed by the vertices of P' in their order along the path. For $a \in T$, denote this bitstring by I_a, and write $I_a(S)$ for the substring on the vertices $S \subset P$, retaining the original order. In the case that $S = \{v\}$ we will write simply $I_a(v)$. For $v \in P'$, $I_a(v)$ is a one iff $v \in N_{P'}(a)$.

A general observation is that for any $a \in T$, there can be no run of ones longer than 3 in I_a, otherwise we could easily extend P by inserting a between the vertices of any run of length 4. Call a zero followed by a maximal run of ones a 3-, 2-, or 1-block, depending on the length of the maximal run. Call a zero that is followed by another zero a 0-block. Thus I_a is comprised of disjoint 3-, 2-, 1-, and 0-blocks. We note that only the 3-blocks have a density of ones that is greater than $2/3$.

For any given $a \in T$, we will often make use of a partition of I_a into substrings (and thus, a partition of P' into subpaths) according to the 3-blocks. We denote the substrings, which we refer to as intervals, by $I_a^0, I_a^1, I_a^2 \ldots, I_a^l$. The interval I_a^j is defined to begin *after* the jth 3-block and end with the $(j + 1)$th 3-block. Of course, I_a^l may not end with a 3-block. Our first case is when T is large compared to the absorbing path.

6.1. T is larger than $\alpha^8 n$

An interval I_a^j comprised of a (possibly empty) run of 2-blocks followed by a 3-block is called a "heavy" interval. Note that only the heavy intervals have a density of ones greater than $2/3$.

We begin by defining an operation, `HEAVY SWAP`, which exchanges vertices of T with vertices of P' in such a way as to extend P'. We will identify the conditions necessary for the operation to take place.

1. There exists a vertex $x_1 \in P'$ such that $H = \{a \in T | a$ has a heavy interval beginning at $x_1\}$ is non-empty.
2. The minimum length of the heavy intervals beginning at x_1 is less than $3|H| - 2$.

With these conditions in place, we define the operation. Let $a \in H$ be a vertex whose heavy interval beginning at x_1 is of minimum length, say $3k + 1$. Define the subpath $Q \subset P'$ of length $3k + 4$ comprised of the 3 vertices preceding x_1 and the vertices along the heavy interval in I_a,

$$Q = (o_1, o_2, o_3, x_1, o_4, o_5, x_2, o_6, o_7, \ldots, x_{k-1}, o_{2k}, o_{2k+1}, x_k, o_{2k+2}, o_{2k+3}, o_{2k+4}).$$

We have $o_i \in N_{P'}(a)$ and $x_i \notin N_{P'}(a)$, and the substring

$$I_a(Q) = (1, 1, 1, 0, 1, 1, 0, 1, 1, \ldots, 0, 1, 1, 0, 1, 1, 1).$$

In fact, by the minimality of a, for every $b \in H$,

$$I_b(Q \setminus \{o_{2k+4}\}) = I_a(Q \setminus \{o_{2k+4}\}).$$

The conditions ensure that $|H| \geq k + 1$, and so we can find k vertices $b_1, b_2, \ldots, b_k$ from $H \setminus \{a\}$. The path

$$Q' = (o_1, o_2, b_1, o_3, o_4, b_2, o_5, o_6, b_3, o_7, \ldots, o_{2k}, b_k, o_{2k+1}, o_{2k+2}, a, o_{2k+3}, o_{2k+4})$$

is a legitimate square-path, with which we replace Q in P'. This defines the operation, which extends P' by one.

Claim 11. *If there exists a subset $H_0 \subset T$ of size $n^{3/4}$ such that for all $a \in H_0$, I_a contains at least $3n^{3/4}$ heavy intervals, then we can extend P.*

We will demonstrate that the conditions necessary for `HEAVY SWAP` are satisfied. Call a heavy interval "short" if it is of length less than $n^{1/4}$. Then for each $a \in H_0$, there are at least $2n^{3/4}$ *short* heavy intervals in I_a. Indeed, otherwise we get for the size of the union of the long heavy intervals strictly more than $n^{3/4} n^{1/4} = n$ vertices, a contradiction. By the pigeonhole principle, there is a vertex $x_1 \in P'$ where at least

$$\frac{2n^{3/4} n^{3/4}}{m} \geq 2n^{1/2}$$

short heavy intervals begin. Let $H \subset H_0$ be those vertices which all have a short heavy interval beginning at the vertex x_1. As $|H| \geq 2n^{1/2}$ and for every $a \in H$ the length of the heavy interval of I_a beginning at x_1 is less than $n^{1/4} \ll 2n^{1/2}$, we can perform the operation HEAVY SWAP.

A simple calculation shows that if $deg(a, P') \geq \frac{2}{3}m + n^{3/4}$, then I_a contains at least $3n^{3/4}$ heavy intervals. Otherwise, with s the size of the union of the heavy intervals, recalling the observation on the density of the heavy intervals,

$$deg(a, P') < \frac{2}{3}(s - 3n^{3/4}) + 3n^{3/4} + \frac{2}{3}(m - s) = \frac{2}{3}m + n^{3/4},$$

a contradiction. Setting $T' = \{a \in T | deg(a, P') \geq \frac{2}{3}m + n^{3/4}\}$, assuming that the premise of Claim 11 fails, it follows that $|T'| \leq n^{3/4}$. Letting $|T| = t$, we have in this case that for every $a \in T \setminus T'$,

$$deg_T(a) \geq \frac{1}{2}t + \frac{1}{2}(\alpha t + n^{3/4}). \tag{11}$$

Indeed, for $a \in T \setminus T'$, using $deg(a, P') \leq \frac{2}{3}m + n^{3/4}$, $|R| \leq \alpha^{20}n$, $|P| \leq \alpha^9 n$,

$$deg_T(a) \geq \frac{2}{3}n - \left(\frac{2}{3}m + n^{2/4}\right) - \alpha^9 n - \alpha^{20}n \geq \frac{2}{3}t - \frac{\alpha^9 n + \alpha^{20}n}{3} - n^{3/4}$$

$$\geq \frac{1}{2}t + \frac{\alpha t + n^{3/4}}{2}$$

if $t \geq \frac{1}{1-\alpha}(2\alpha^9 n + 2\alpha^{20}n + 7n^{3/4})$, which is true for large enough n when $t \geq \alpha^8 n$.

With degree condition (11), we are guaranteed to find a square-path in $T \setminus T'$ of length at least αt. Indeed, as any two vertices in $T \setminus T'$ have degree in T of $\frac{1}{2}t + \frac{1}{2}(\alpha t + n^{3/4})$, they have a common neighborhood in T of size at least $\alpha t + n^{3/4}$. Hence, the greedy algorithm is guaranteed to be able to extend any square-path of length less than αt in T by a vertex not in T' and not on the path being extended. In this case, we can extend P by connecting a square-path of length αt through the reservoir.

Thus, we may assume that T is close to the size of the absorbing path.

6.2. T is smaller than $\alpha^8 n$

The outline of the proof in this case is as follows: assuming that the premise of Claim 11 does not hold, we find a large matching in P' that can be moved out of the path by exchanging with vertices of T without disturbing P. Then we can either extend P or there is a large set of disjoint triangles in P' that can be moved out without disturbing P. If we still are unable to extend P then we can find a long square-path in P' which we move out and then connect through Lemma 9.

We call an interval I_a^j "even" if it contains no 0-blocks and exactly one 1-block. Note that the even intervals have a density of ones exactly $2/3$.

As in the case of the heavy intervals, we will define the operation EVEN SWAP that exchanges vertices of T with vertices of P', taking advantage of a certain alignment of even intervals. In this case we will not be able to extend P', but rather we will identify vertices of P' and vertices of T which can be exchanged for the purpose of guaranteeing a certain structure in T. We first identify the conditions necessary to perform the operation EVEN SWAP:

1. There exists a vertex $x_1 \in P'$ such that $D = \{a \in T | a$ has an even interval whose 1-block begins at $x_1\}$ is non-empty
2. The minimum length of the even intervals whose 1-blocks begin at x_1 is less than $3|D|$.

With these conditions in place, we define the operation. Let $a \in D$ be a vertex whose even interval aligned with x_1 is of minimum length, say $3k$. Define the subpath $Q \subset P'$ of length $3k + 2$ comprised of the 2 vertices preceding x_1 and the vertices of the even interval of I_a containing the position of x_1,

$$Q = (o_1, o_2, x_1, o_3, x_2, o_4, o_5, \ldots, x_{k-1}, o_{2k-2}, o_{2k-1}, x_k, o_{2k}, o_{2k+1}, o_{2k+2}).$$

We have $o_i \in N_{P'}(a)$ and $x_i \notin N_{P'}(a)$, and the substring

$$I_a(Q) = (1, 1, 0, 1, 0, 1, 1, \ldots, 0, 1, 1, 0, 1, 1, 1).$$

As $|D| \geq k$ we can find distinct vertices $b_1, b_2, \ldots, b_{k-1}$ from $D \setminus \{a\}$, and the path

$$Q' = (o_1, o_2, b_1, o_3, o_4, b_2, o_5, \ldots, o_{2k-2}, b_{k-1}, o_{2k-1}, o_{2k}, a, o_{2k+1}, o_{2k+2})$$

is a legitimate subpath of P' of length $3k + 2$ with which Q can be replaced. This defines the operation.

For $a \in T$ and I_a^j an even interval for a, referring to the notation defined above, we consider x_1 and x_2 to be "swappable" with a via the operation EVEN SWAP. Unfortunately, in order to bring either x_1 or x_2 from the path into T, we are forced to swap every x_i. For this reason, we will only consider the zero vertices surrounding the 1-block of an even interval swappable if the interval is of length less than $\frac{1}{\alpha^2}$. We define a "short" even interval to be one whose length is less than $\frac{1}{\alpha^2}$.

There is one other class of vertex which is swappable with a. Let x be any vertex whose position is a zero in I_a. If the two positions preceding x and the two positions succeeding x are all ones in I_a, then a and x can be exchanged. More precisely, if the subpath

$$Q = (o_1, o_2, x, o_3, o_4)$$

is such that $o_i \in N_{P'}(a)$ and $x \notin N_{P'}(a)$, then

$$Q' = (o_1, o_2, a, o_3, o_4)$$

is a legitimate subpath of P'.

We are now prepared to define S_a, the set of swappable vertices for a. For $a \in T$, we define $v \in S_a$ iff $I_a(v) = 0$ and v is either

(a) preceded and succeeded by two ones in I_a, or
(b) the zero of a 1-block in a short even interval for I_a, or
(c) the zero following a 1-block in a short even interval for I_a.

We observe that in the case of small T, since $|R| \leq \alpha^{20}n$, $|P| \leq \alpha^9 n$, $|T| \leq \alpha^8 n$, for every $a \in T$,

$$deg(a, P') \geq \frac{2}{3}m - 2\alpha^8 n.$$

Again, define $H_0 = \{a \in T | I_a \text{ contains at least } 3n^{3/4} \text{ heavy intervals}\}$. Claim 11 holds regardless of the size of T, and so we may assume $|H_0| \leq n^{3/4}$. It follows by the remark following Claim 11 that for $a \in T \setminus H_0$, $deg(a, P') \leq \frac{2}{3}m + n^{3/4}$.

For $a \in T \setminus H_0$, we provide a lower bound on the size of S_a, the main consequence of which is that S_a is α-dense for every $a \in T \setminus H_0$.

Claim 12. *For $a \in T \setminus H_0$, $|S_a| > (\frac{1}{3} - 3\alpha^2)n$.*

By the last two remarks, as $a \in T \setminus H_0$,

$$\frac{2}{3}n - 2\alpha^8 n \leq deg(a, P') \leq \frac{2}{3}m + n^{3/4}. \tag{12}$$

Define a counting function on bitstrings as follows: to each zero assign a value of -2 and to each one a $+1$ and sum the values over the length of the bitstring. Thus a bitstring with a density of ones exactly $2/3$ has a count of 0. By (12) the number of ones in I_a is at least $(\frac{2}{3} - 2\alpha^8)n$, and thus the number of zeros is at most $m - (\frac{2}{3} - 2\alpha^8)n$. Hence, the count for I_a is at least

$$\frac{2}{3}n - 2\alpha^8 n - 2\left(m - \frac{2}{3}n + 2\alpha^8 n\right) = 2n - 2m - 6\alpha^8 n \geq -6\alpha^8 n.$$

Consider the intervals I_a^j. At most $3n^{3/4}$ are heavy, each with a count of $+1$, and thus the total contribution to the count from the heavy intervals is at most $3n^{3/4}$. The rest of the I_a^j include at least a 1-block or a 0-block. For each I_a^j containing at least one 1-block, distinguish an arbitrary 1-block of I_a^j, and denote by b_1^j the number of undistinguished 1-blocks.. For any I_a^j, denote by b_0^j the number of 0-blocks in I_a^j, if any. If I_a^j contains at least one 1-block, the count on I_a^j is $-b_1^j - 2b_0^j$. If I_a^j contains only 0-blocks, the count is $-2b_0^j + 1 < b_0^j$. Let $b_0 = \sum_j b_0^j$ and $b_1 = \sum_j b_1^j$. Summing over the intervals, the count on I_a is at most $-b_1 - b_0 + 3n^{3/4}$. It follows that $b_0 + b_1 \leq 6\alpha^8 n + 3n^{3/4} \leq 7\alpha^8 n$.

As $m \geq n - 2\alpha^8 n$, using (12), the number of zeros in I_a is at least

$$m - \left(\frac{2}{3}n + n^{3/4}\right) \geq \frac{1}{3}n - 2\alpha^8 n - n^{3/4}.$$

For each 0-block or undistinguished 1-block in the interval I_a^j, the zero belonging to the block, the zero following the block, and the 2 zeros surrounding the distinguished 1-block (if it exists), are not in positions corresponding to vertices of S_a. There are at most $2\alpha^2 n$ zeros surrounding the 1-blocks of long even intervals. The rest of the zeros correspond to positions of vertices in S_a. Thus, the total number of vertices of P' in S_a is at least

$$\frac{1}{3}n - 2\alpha^8 n - n^{3/4} - 4(b_0 + b_1) - 2\alpha^2 n \geq \frac{1}{3}n - 3\alpha^2 n,$$

proving the claim.

Let $|T \setminus H_0| = t$. From Claim 12, an easy calculation shows that there is a set $S \subset \bigcup_a S_a \subset P'$ such that $|S| \geq (\frac{1}{3} - 4\alpha^2)n$ and $\forall v \in S$, we have $v \in S_a$ for at least $\alpha^2 t$ vertices $a \in T \setminus H_0$.

Let $S' \subset S$ be such that

1. $|S'| \leq \alpha^4 t$, and
2. for any $u, v \in S'$, the distance from u to v along the path P' is at least $\frac{1}{\alpha^2}$.

Then we can move S' from the path to T by exchanging the vertices of S' with vertices of $T \setminus H_0$. To see this, let $S' \subset S$ be as described. For any $v \in S'$, and for each a for which $v \in S_a$, it is either of type (a), (b), or (c) by definition of S_a. By the

pigeonhole principle there is a set of at least $\alpha^2 t/3$ vertices $a \in T \setminus H_0$ for which v is of the same type in S_a. If this is type (a) then v can be exchanged directly. If it is type (b) or (c), then we have short even intervals for at least $\alpha^2 t/3 \gg 1/\alpha^2$ vertices $a \in T \setminus H_0$, all of whose 1-blocks are aligned. The conditions for EVEN SWAP are satisfied and we can exchange v with a vertex of $T \setminus H_0$, but we may have to perform as many as $1/3\alpha^2$ exchanges of other vertices in the short even interval. The distance condition on $u, v \in S'$ precludes the possibility that exchanging vertices within an interval of length $\frac{1}{\alpha^2}$ in order to move u from P' to T destroys the conditions necessary to move v out—it ensures that exchanging u does not diminish our ability to exchange v apart from simply using up vertices of T. From these observations, each exchange from S' to T may use up to $\frac{1}{3\alpha^2}$ vertices of T. When trying to exchange $u \in S'$, as long as there are $\frac{1}{\alpha^2}$ vertices $a \in T$ for which $u \in S_a$, we are able to perform the operation. With fewer than

$$\frac{\alpha^2 t/3}{1/3\alpha^2} = \alpha^4 t$$

exchanges, this condition is guaranteed.

By the extremal condition, as $|S| \geq (\frac{1}{3} - 4\alpha^2)n$, S has density α. We can easily find a matching M with $\alpha^4 t/2$ edges such that every two vertices of $V(M)$ are separated by constant distance on P'. By the previous remark, we can exchange these vertices of M without disturbing P'. Therefore, we assume that $V(M) \subset T$. We define H_0 as before, and let M' be the set of edges of M that are disjoint from H_0. Then we have $|M'| \geq \frac{\alpha^4 t}{2} - n^{3/4} \geq \frac{\alpha^4 t}{3}$.

For $(a, b) \in M'$, define the "overlap" of (a, b) to be $S_{a,b} = S_a \cap S_b$. We have the following claim:

Claim 13. *If $|S_{a,b}| > n^{1/2}$ for at least $3n^{1/2}$ edges $(a, b) \in M'$, then we can extend P.*

For any edge $(a, b) \in M$ and $x \in S_{a,b}$, we classify x as being of one of two types with respect to (a, b). If the two vertices following x on P' are both in $N_{P'}(a)$ and the two vertices preceding x are both in $N_{P'}(b)$, (or vice versa) we say that it is of type (1). Otherwise it is of type (2).

If the premise of the claim holds, then by the pigeonhole principle there is a vertex $x \in P'$ such that for at least $n^{1/2}3n^{1/2}/m > 3$ edges (a, b), (c, d), and (e, f), x is in the overlap for all three edges. We study two cases. In Case 1, x is of type (1) with respect to one of these three edges. In Case 2, x is of type (2) for all three edges.

The vertices along P' around x are relevant. Specifically, we focus on the subpath of length 10 on P',

$$Q = (u_0, u_1, u_2, u_3, u_4, u_5, u_6, x, u_7, u_8).$$

When specifying a substring we will let a "$*$" denote that the value of the substring in that position is unrestricted. Of course, the value in position x is a zero for all strings.

Case 1: In this case, x is of type (1) with respect to one of the edges, say (a, b). Without loss of generality, the substrings on Q for a and b are

$$I_a(Q) = (*, *, *, *, *, 1, 1, 0, 1, *),$$

and

$$I_b(Q) = (*, *, *, *, *, *, 1, 0, 1, 1).$$

In this case we define the subpath

$$Q' = (u_0, u_1, u_2, u_3, u_4, u_5, u_6, a, b, u_7, u_8).$$

Replacing Q by Q' results in a legitimate square-path, where we have replaced the subpath Q of length 10 by Q' of length 11, and thus extended P.

Case 2: In this case x of type (2) with respect to all three edges. By definition of S_a, whenever $x \in S_a$ either the two vertices which follow x on the path or the two vertices which precede x on the path are in $N_{P'}(a)$. It follows without loss of generality that there are two edges, say (a, b) and (c, d), for which the two vertices following x are in the neighborhoods of a, b, c, and d. Because x is not of type (1) for any edge, and because x is in an even interval for every vertex a, b, and c, their substrings on Q are all of the form

$$I_a(Q) = I_b(Q) = I_c(Q) = (*, *, *, 1, 1, 0, 1, 0, 1, 1).$$

If u_2 is not in the neighborhood of a or b, then the substrings on Q for a and b are both exactly

$$I_a(Q) = I_b(Q) = (1, 1, 0, 1, 1, 0, 1, 0, 1, 1),$$

in which case we are back to Case 1, with u_2 taking the place of x. Otherwise u_2 is in the neighborhood of, say, a, and so the substrings of a and b on Q are

$$I_a(Q) = (*, 0, 1, 1, 1, 0, 1, 0, 1, 1),$$

and

$$I_b(Q) = (*, *, *, 1, 1, 0, 1, 0, 1, 1).$$

In this case, we define the subpath

$$Q' = (u_0, u_1, u_2, u_3, a, b, u_4, u_6, c, u_7, u_8).$$

Replacing the subpath Q of length 10 by Q' of length 12 results in a legitimate square-path, extending P' by 2. These cases are exhaustive, and thus the claim is established.

Thus we may assume that there exists a set $M'' \subset M'$ of at least $\alpha^4 t/3 - 3n^{1/2} \geq \alpha^4 t/4$ edges such that each edge of M'' has an overlap of less than $n^{1/2}$. From Claim 12 it follows that for each edge $(a, b) \in M''$, we have

$$|S_a \cup S_b| \geq 2\left(\frac{1}{3} - 4\alpha^2\right)n - n^{1/2} \geq \left(\frac{2}{3} - 9\alpha^2\right)n.$$

Again, we conclude that there is a set $S \subset \bigcup_a S_a \subset P$ such that $|S| \geq (\frac{2}{3} - 10\alpha^2)n$ and $\forall v \in S$ we have $v \in S_a \cup S_b$ for at least $\alpha^2|M''| \geq \alpha^6 t/4$ edges (a, b) of M''. The reader may check that any set $S' \subset S$ satisfying

1. $|S'| \leq \alpha^8 t/4$, and
2. $\forall u, v, \in S'$, the distance from u to v along P' is less than $\frac{1}{\alpha^2}$

can be exchanged for vertices from distinct edges of M''.

By (2), for any $v \in S$,

$$N_S(v) \geq \frac{2}{3}n - \left(n - \left(\frac{2}{3} - 10\alpha^2\right)n\right) = \left(\frac{1}{3} - 10\alpha^2\right)n.$$

As G is not extremal, the density of $N_S(v)$ is at least α. It follows that every vertex in S is contained in many triangles within S. Let $Z \subset S$ be a set of $\alpha^8 t/12$ vertex-disjoint triangles whose vertex set $V(Z)$ satisfies the above two conditions. By the observation we can exchange $V(Z)$ with vertices of M'' without disturbing P, and so we assume that there is a set of $\alpha^8 t/12$ vertex-disjoint triangles $Z \subset T$. As before, we let $Z' = Z \setminus H_0$ be the set of at least $\alpha^8 t/12 - n^{3/4} \geq \alpha^8 t/24$ triangles all of whose vertices satisfy (12).

As Claim 13 applies to any set of disjoint edges from Z', we can assume that there is a set Z'' containing at least $\alpha^8 t/24 - 3n^{1/2} \geq \alpha^8 t/48$ vertex-disjoint triangles such that any edge (a, b) of any triangle of Z'' has $|S_{a,b}| < n^{1/2}$. By this bound and Claim 12, it follows that for every triangle $(a, b, c) \in Z''$, $|S_a \cup S_b \cup S_c| \geq (1 - \alpha)n$. Again, we can find a set $S \subset \bigcup_{a \in V(Z'')} S_a$ such that $|S| \geq (1-2\alpha)n$ and for every $v \in S$ there are at least $\alpha|U''| \geq \alpha^9 t/48$ triangles $(a, b, c) \in Z''$ such that $v \in S_a \cup S_b \cup S_c$. For any set $S' \subset S$ of size $\alpha^{11} t/48$ all of whose vertices are distance $\frac{1}{\alpha^2}$ apart on P', we can exchange at once each vertex of S' with vertices of $V(Z'')$. With $deg_S(v) \geq (2/3 - 2\alpha)n$, we can easily find a square-path of length $\alpha^{13} t < \alpha^{11} t/48$ satisfying the distance condition. Exchanging this path into T and connecting through the reservoir extends P.

In every case we either extend P directly or find a square-path of length at least $\alpha^{13} t \geq \alpha^{24} n$ which we connect through the reservoir. The allowance for forbidden vertices in Lemma 9 ensures that we can continue to perform connections until T is small enough that $R \cup T$ can be absorbed by P_A. As in the program outlined at the beginning of the section, at this point we connect the endpairs of P through the reservoir to form a cycle containing P_A, and finally absorb $R \cup T$.

7. Conclusion

In order to solve the Pósa problem for every n in the case of $k = 2$, we plan to replace the costly Connecting Lemma with an alternative which we are developing. It is our hope that we will be able to push down n_0 to around 100 at which point we will be able to develop a computer program, taking advantage of much of the structure identified in this paper to solve the conjecture for every n. This is a work in progress.

Acknowledgements

References

[1] G.A. Dirac, Some theorems on abstract graphs, Proceedings of the London Mathematical Society 2 (1952) 68–81.
[2] P. Erdős, Problem 9, in: M. Fieldler (Ed.), Theory of Graphs and its Applications, Czech. Acad. Sci. Publ., Prague, 1964, p. 159.
[3] G. Fan, R. Häggkvist, The square of a hamiltonian cycle, SIAM Journal on Discrete Mathematics (1994) 203–212.
[4] G. Fan, H.A. Kierstead, The square of paths and cycles, Manuscript.
[5] G. Fan, H.A. Kierstead, The square of paths and cycles, Journal of Combinatorial Theory, Series B 63 (1995) 55–64.
[6] G. Fan, H.A. Kierstead, Hamiltonian square-paths, Journal of Combinatorial Theory, Series B 67 (1996) 167–182.
[7] G. Fan, H.A. Kierstead, Partitioning a graph into two square-cycles, Journal of Graph Theory 23 (1996) 241–256.
[8] R.J. Faudree, R.J. Gould, M. Jacobson, On a problem of Pósa and Seymour, Manuscript.
[9] R.J. Faudree, R.J. Gould, M.S. Jacobson, R.H. Schelp, On a problem of Paul Seymour, in: V.R. Kulli (Ed.), Recent Advances in Graph Theory, Vishwa International Publication, 1991, pp. 197–215.

[10] R. Häggkvist, On F-hamiltonian graphs, in: J.A. Bondy, U.S.R. Murty (Eds.), Graph Theory and Related Topics, Academic Press, New York, 1979, pp. 219–231.

[11] A. Hajnal, E. Szemerédi, Proof of a conjecture of Erdős, in: P. Erdős, A. Rényi, V.T. Sós (Eds.), Combinatorial Theory and its Applications, vol. II, in: Colloq. Math. Soc. J. Bolyai, vol. 4, North-Holland, Amsterdam, 1970, pp. 601–623.

[12] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley, New York, 2000.

[13] H.A. Kierstead, J. Quintana, Square Hamiltonian cycles in graphs with maximal 4-cliques, Discrete Mathematics 178 (1998) 81–92.

[14] J. Komlós, G.N. Sárközy, E. Szemerédi, On the square of a Hamiltonian cycle in dense graphs, Random Structures and Algorithms 9 (1996) 193–211.

[15] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1) (1997) 109–123.

[16] J. Komlós, G.N. Sárközy, E. Szemerédi, On the Pósa-Seymour conjecture, Journal of Graph Theory 29 (1998) 167–176.

[17] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Seymour conjecture for large graphs, Annals of Combinatorics 2 (1998) 43–60.

[18] J. Komlós, G.N. Sárközy, E. Szemerédi, An algorithmic version of the Blow-up Lemma, Random Structures and Algorithms 12 (1998) 297–312.

[19] V. Rödl, A. Ruciński, E. Szemerédi, A Dirac-type theorem for 3-uniform hypergraphs, Combinatorics, Probability and Computing 15 (2006) 229–251.

[20] P. Seymour, Problem section, in: T.P. McDonough, V.C. Mavron (Eds.), Combinatorics: Proceedings of the British Combinatorial Conference 1973, Cambridge University Press, 1974, pp. 201–202.

[21] E. Szemerédi, Regular partitions of graphs, in: Colloques Internationaux C.N.R.S. No. 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay, 1976, pp. 399–401.