On the Turán Properties of Infinite Graphs

Andrzej Dudek and Vojtěch Rödl

Department of Mathematics and Computer Science
Emory University, Atlanta, USA
{adudek,rodl}@mathcs.emory.edu

Submitted: Dec 9, 2006; Accepted: Mar 16, 2008; Published: Mar 20, 2008
Mathematics Subject Classifications: 05C35, 05C38

Abstract

Let $G^{(\infty)}$ be an infinite graph with the vertex set corresponding to the set of positive integers $\mathbb{N}$. Denote by $G^{(l)}$ a subgraph of $G^{(\infty)}$ which is spanned by the vertices $\{1, \ldots, l\}$. As a possible extension of Turán's theorem to infinite graphs, in this paper we will examine how large $\liminf_{l\to\infty} \frac{|E(G^{(l)})|}{l^2}$ can be for an infinite graph $G^{(\infty)}$, which does not contain an increasing path I_k with $k+1$ vertices. We will show that for sufficiently large k there are I_k–free infinite graphs with $\frac{1}{4} + \frac{1}{200} < \liminf_{l\to\infty} \frac{|E(G^{(l)})|}{l^2}$. This disproves a conjecture of J. Czipszer, P. Erdős and A. Hajnal. On the other hand, we will show that $\liminf_{l\to\infty} \frac{|E(G^{(l)})|}{l^2} \leq \frac{1}{3}$ for any k and such $G^{(\infty)}$.

1 Introduction

1.1 Preliminaries

Let $G^{(\infty)} = \big(V(G^{(\infty)}), E(G^{(\infty)})\big)$ be an infinite graph with the vertex set corresponding to the set of natural numbers, *i.e.*, $V(G^{(\infty)}) = \mathbb{N}$, and the set of edges $E(G^{(\infty)})$. Denote by $G^{(l)}$ the subgraph of $G^{(\infty)}$ induced on the set $\{1, \ldots, l\}$. Let $G^{(\infty)}$ be a K_{k+1}–free graph. Then, by Turán's theorem for finite graphs we get that $\liminf_{l\to\infty} \frac{|E(G^{(l)})|}{l^2} \leq \limsup_{l\to\infty} \frac{|E(G^{(l)})|}{l^2} \leq \frac{1}{2}\big(1 - \frac{1}{k}\big)$. On the other hand, a K_{k+1}–free graph $G^{(\infty)}$ with edges $\{i, j\} \in E(G^{(\infty)})$ if $j - i \neq 0 \mod k$, achieves this bound. Hence, the Turán density for finite and infinite K_{k+1}–free graphs is the same.

In this paper we study the edge density of graphs without an increasing path of length k. We say that $I_k = i_1 i_2 \ldots i_{k+1}$ is an increasing path of $G^{(\infty)}$ if $i_1 < i_2 < \cdots < i_{k+1}$ and $\{i_j, i_{j+1}\} \in E(G^{(\infty)})$. One can easily see that for any fixed l there exists a graph $G^{(l)}$ not containing I_k such that $|E(G^{(l)})|$ equals to the Turán number for K_{k+1}–free graphs. Hence, for finite graphs forbidding I_k leads to the same restriction on number of edges as forbidding K_{k+1}. While the maximum value $\limsup_{l\to\infty} \frac{|E(G^{(l)})|}{l^2}$ can achieve over all I_k-free infinite graphs G is $\frac{1}{2}\left(1-\frac{1}{k}\right)$, the corresponding value for the limit inferior is harder to find. Set

$$p(G) = \liminf_{l\to\infty} \frac{|E(G^{(l)})|}{l^2}.$$

Furthermore, let the *path Turán number* be defined as

$$p(k) = \sup\{p(G) \mid G \text{ is } I_k\text{-free}\}.$$

J. Czipszer, P. Erdős and A. Hajnal were the first ones who examined these numbers. In [1], they showed that $p(2) = \frac{1}{8}$ and $p(3) = \frac{1}{6}$. The following was stated in [1] as a question and in [2, 3] as a conjecture.

Conjecture 1.1 ([1, 2, 3]) *For any $k \geq 2$ the following holds*

$$p(k) = \frac{1}{4}\left(1 - \frac{1}{k}\right). \tag{1}$$

In this paper we will show that in general this fails to be true. In fact, for sufficiently large k the value of $p(k)$ exceeds $\frac{1}{4}$.

Theorem 1.2 *For any $k \geq 162$ the path Turán number satisfies*

$$p(k) > \frac{1}{4} + \frac{1}{200}.$$

We were unable to decide if (1) holds for $k = 4$. Here we will show that (1) fails for $k = 16$.

Theorem 1.3 *The path Turán number $p(16)$ satisfies*

$$p(16) > \frac{1}{4}\left(1 - \frac{1}{16}\right).$$

Moreover, complementing Theorems 1.2 and 1.3 we will show the following upper bound, confirming that the Turán number for I_k–free infinite graphs differs significantly from those for finite graphs.

Theorem 1.4 *For any $k \geq 2$ the path Turán number satisfies*

$$p(k) \leq \frac{1}{3}.$$

1.2 Reformulation

In order to prove Theorems 1.2, 1.3 and 1.4 we will work with infinite sequences of k symbols rather than with infinite graphs. Let $\mathcal{C} = \{c_n\}_{n=1}^{\infty}$ be a sequence of integers with $c_n \in \{1, 2, \ldots, k\}$, and

$$S_{\mathcal{C}}(k, l) = \left|\{(i, j) \mid 1 \leq i < j \leq l \text{ and } c_i < c_j\}\right|.$$

Furthermore, let

$$s_{\mathcal{C}}(k) = \liminf_{l\to\infty} \frac{|S_{\mathcal{C}}(k,l)|}{l^2},$$

and

$$s(k) = \sup_{\mathcal{C}} s_{\mathcal{C}}(k).$$

The following statement shows the equivalence between path Turán numbers and the numbers $s(k)$ for a fixed k.

Lemma 1.5 *Let $k \geq 2$. Then, $p(k) = s(k)$.*

Proof. For a given sequence $\mathcal{C} = \{c_n\}_{n=1}^{\infty}$ of k symbols, let $G^{(\infty)}$ be the infinite graph which corresponds to this sequence, *i.e.*, $V(G^{(\infty)}) = \mathbb{N}$ and $E(G^{(\infty)}) = \big\{\{i,j\} \mid i < j \text{ and } c_i < c_j\big\}$. Note, $G^{(\infty)}$ is an I_k–free. Hence, $|E(G^{(l)})| = S_{\mathcal{C}}(k,l)$, and consequently $p(k) \geq s(k)$. Conversely, let $G^{(\infty)}$ be an I_k–free infinite graph. Then, $G^{(\infty)}$ defines a partition of $\mathbb{N}$

$$\mathbb{N} = \bigcup_{j=1}^{k} N_j(G^{(\infty)}),$$

where

$$N_1(G^{(\infty)}) = \Big\{\alpha \in \mathbb{N} \mid \forall \beta \in \mathbb{N} : \{\alpha,\beta\} \in E(G^{(\infty)}) \Rightarrow \alpha < \beta\Big\},$$

and

$$N_i(G^{(\infty)}) = \Big\{\alpha \in \mathbb{N} \setminus \bigcup_{j=1}^{i-1} N_j(G^{(\infty)}) \mid \forall \beta \in \mathbb{N} : \{\alpha,\beta\} \in E(G^{(\infty)}) \Rightarrow$$
$$\alpha < \beta \text{ or } \beta \in \bigcup_{j=1}^{i-1} N_j(G^{(\infty)})\Big\},$$

for $i \in \{2,\ldots,k\}$. Let $\mathcal{C} = \{c_n\}_{n=1}^{\infty}$ be the sequence which corresponds to the above partition, *i.e.*, $c_n = i$ if $n \in N_i(G^{(\infty)})$. Note, $|E(G^{(l)})| \leq S_{\mathcal{C}}(k,l)$, and consequently, $p(k) \leq s(k)$. □

2 Auxiliary sequences

2.1 Sequence $\mathcal{A} = \{a_n\}_{n=1}^{\infty}$

The sequence $\mathcal{A} = \{a_n\}_{n=1}^{\infty}$ on the symbols $\{1,\ldots,k\}$, which we define below, will consists of infinitely many blocks. For $j \in \mathbb{N}$, the j-th block is a subsequence of $k2^j$ consecutive symbols, which consists of 2^j one's followed by 2^j two's, etc. We abbreviate such block of length $k2^j$ by $2^j \otimes \{1,2,\ldots,k\}$. Below are the first three blocks of the sequence $\mathcal{A}$:

$$1|1|2|2|\ldots|k|k\,\Big|\,1|1|1|1|2|2|2|2|\ldots|k|k|k|k\,\Big|\,1|1|1|1|1|1|1|1|2|2|2|2|2|2|2|2|\ldots|k|k|k|k|k|k|k|k\,\Big|$$

Formally, for a given k, $\mathcal{A} = \{a_n\}_{n=1}^{\infty}$ is a sequence of integers with $a_n \in \{1, \dots, k\}$ defined as follows:

(i) for any $n \leq 2k$, $a_n = i$ if and only if $2(i-1) < n \leq 2i$, otherwise

(ii) for any $n > 2k$, $a_n = i$ if and only if there exists an integer number $m \in \mathbb{N} \cup \{0\}$ such that

$$k(2 + 2^2 + \cdots + 2^m) + (i-1)2^{m+1} < n \leq k(2 + 2^2 + \cdots + 2^m) + i2^{m+1}.$$

For $i \in \{0, 1, \dots, k\}$, we identify the indices $n_i(m)$ for which value of the sequence changes from i to $i+1$, *i.e.*, $n_i(m) = k(2 + 2^2 + \cdots + 2^m) + i2^{m+1}$. Note, $n_k(m) = n_0(m+1)$ and

$$n_i(m) = (2k + 2i)2^m + o(2^m). \tag{2}$$

Proposition 2.1 *For any $i \in \{0, 1, \dots, k-1\}$ we have*

$$S_{\mathcal{A}}(k, n_i(m)) = \left(\frac{4}{3}k(k-1) + 4i(i-1)\right)4^m + o(4^m).$$

Proof. First, we will find a formula for $s_{\mathcal{A}}(k, n_0(m))$. Note that setting $S_{\mathcal{A}}(k, n_0(0)) = 0$ we obtain that for $m \geq 1$,

$$\begin{aligned} S_{\mathcal{A}}(k, n_0(m)) &= S_{\mathcal{A}}(k, n_0(m-1)) + \sum_{j=1}^{k-1} j(2 + 2^2 + \cdots + 2^m)2^m \\ &= S_{\mathcal{A}}(k, n_0(m-1)) + k(k-1)2^m(2^m - 1), \end{aligned}$$

and hence by induction,

$$S_{\mathcal{A}}(k, n_0(m)) = k(k-1)\sum_{j=1}^{m} 2^j(2^j - 1) = \frac{4}{3}k(k-1)4^m + o(4^m).$$

Similarly, for $i \geq 1$,

$$\begin{aligned} S_{\mathcal{A}}(k, n_i(m)) &= S_{\mathcal{A}}(k, n_0(m)) + \sum_{j=1}^{i-1} j(2 + 2^2 + \cdots + 2^{m+1})2^{m+1} \\ &= S_{\mathcal{A}}(k, n_0(m)) + i(i-1)2^{m+1}(2^{m+1} - 1) \\ &= \left(\frac{4}{3}k(k-1) + 4i(i-1)\right)4^m + o(4^m). \end{aligned}$$

$\square$

By Proposition 2.1 and equation (2) we obtain the following.

Corollary 2.2 *For any $i \in \{0, 1, \ldots, k-1\}$ we have*

$$\lim_{m\to\infty} \frac{S_{\mathcal{A}}(k, n_i(m))}{n_i(m)^2} = \frac{\frac{1}{3}k(k-1) + i(i-1)}{(k+i)^2}.$$

Denote the above limit by $t_{\mathcal{A}}(i)$.

Remark 2.3 Note that the existence of the limit $t_{\mathcal{A}}(i)$ means that the behavior of $\frac{S_{\mathcal{A}}(k,x)}{x^2}$, as a function of x with domain equal to the sequence $n_1(1) < \cdots < n_k(1) < n_1(2) < \cdots < n_k(2) < \cdots < n_i(m) < \cdots$, becomes close to periodic (with period k) for m large. In particular, $t_{\mathcal{A}}(0) = t_{\mathcal{A}}(k)$.

2.2 Sequence $\mathcal{B} = \{b_n\}_{n=1}^{\infty}$

Now, we define the second auxiliary sequence. For an even number k, let $\mathcal{B} = \{b_n\}_{n=1}^{\infty}$ be a sequence of integers with $b_n \in \{1, \ldots, k\}$ such that

$$b_n = \begin{cases} a_n + \frac{k}{2} \mod k, & \text{if } a_n + \frac{k}{2} \neq 0 \mod k, \\ k, & \text{otherwise.} \end{cases}$$

Proposition 2.4 *For any $i \in \{0, 1, \ldots, k-1\}$ we have*

$$S_{\mathcal{B}}(k, n_i(m)) = \begin{cases} \left(k^2 - \frac{4}{3}k + 2i(k+2i-2)\right) 4^m + o(4^m), & \text{if } 0 \le i \le \frac{k}{2}, \\ \left(3k^2 - \frac{10}{3}k + (2i-k-2)(2i-k)\right) 4^m + o(4^m), & \text{otherwise.} \end{cases}$$

Proof. First, we will find a formula for $S_{\mathcal{B}}(k, n_0(m))$. Recall that the m-th block is now of the form $2^m \otimes \{\frac{k}{2}+1, \ldots, k, 1, \ldots, \frac{k}{2}\}$. The number of pairs $b_\alpha < b_\beta$, where b_α belongs to the first $m-1$ blocks and $b_\beta = j+1$, for $j = 1, \ldots, k-1$, and belongs to the last block, is equal to

$$j(2 + 2^2 + \cdots + 2^{m-1})2^m.$$

Setting $S_{\mathcal{B}}(k, n_0(0)) = 0$ yields for $m \ge 1$,

$$S_{\mathcal{B}}(k, n_0(m)) = S_{\mathcal{B}}(k, n_0(m-1)) + \sum_{j=1}^{k-1} j(2 + 2^2 + \cdots + 2^{m-1})2^m + 2\binom{\frac{k}{2}}{2}(2^m)^2,$$

where the last quantity counts the pairs $b_\alpha < b_\beta$ of the last block. Hence,

$$S_{\mathcal{B}}(k, n_0(m)) = S_{\mathcal{B}}(k, n_0(m-1)) + k(k-1)2^m(2^{m-1}-1) + \frac{k}{2}\left(\frac{k}{2}-1\right)4^m,$$

and by induction

$$\begin{aligned} S_{\mathcal{B}}(k, n_0(m)) &= k(k-1)\sum_{j=1}^{m} 2^j(2^{j-1}-1) + \frac{k}{2}\left(\frac{k}{2}-1\right)\sum_{j=1}^{m} 4^j \\ &= \frac{2}{3}k(k-1)4^m + \frac{2}{3}k\left(\frac{k}{2}-1\right)4^m + o(4^m) \\ &= \left(k^2 - \frac{4}{3}k\right)4^m + o(4^m). \end{aligned}$$

Now suppose that $i \in \{0, \ldots, \frac{k}{2}\}$. Then,

$$\begin{aligned} S_{\mathcal{B}}(k, n_i(m)) &= S_{\mathcal{B}}(k, n_0(m)) + \sum_{j=\frac{k}{2}}^{i+\frac{k}{2}-1} j(2+2^2+\cdots+2^m)2^{m+1} + \binom{i}{2}(2^{m+1})^2 \\ &= S_{\mathcal{B}}(k, n_0(m)) + i(k+i-1)2^{m+1}(2^m-1) + 2i(i-1)4^m \\ &= S_{\mathcal{B}}(k, n_0(m)) + 2i(k+2i-2)4^m + o(4^m) \\ &= \left(k^2 - \frac{4}{3}k + 2i(k+2i-2)\right)4^m + o(4^m). \end{aligned}$$

Similarly, for $i \in \{\frac{k}{2}+1, \ldots, k-1\}$, we get

$$\begin{aligned} S_{\mathcal{B}}(k, n_i(m)) &= S_{\mathcal{B}}(k, n_{\frac{k}{2}}(m)) + \sum_{j=1}^{i-\frac{k}{2}-1} j(2+2^2+\cdots+2^{m+1})2^{m+1} \\ &= S_{\mathcal{B}}(k, n_{\frac{k}{2}}(m)) + (2i-k-2)(2i-k)4^m + o(4^m) \\ &= \left(3k^2 - \frac{10}{3}k + (2i-k-2)(2i-k)\right)4^m + o(4^m). \end{aligned}$$

□

By Proposition 2.4 and equation (2) we get the following.

Corollary 2.5 *For any $i \in \{0, 1, \ldots, k-1\}$ we have*

$$t_{\mathcal{B}}(i) = \lim_{m\to\infty} \frac{S_{\mathcal{B}}(k, n_i(m))}{n_i(m)^2} = \begin{cases} \frac{\frac{1}{4}k^2 - \frac{1}{3}k + \frac{i}{2}(k+2i-2)}{(k+i)^2}, & \text{if } 0 \le i \le \frac{k}{2}, \\ \frac{\frac{3}{4}k^2 - \frac{5}{6}k + \frac{1}{4}(2i-k-2)(2i-k)}{(k+i)^2}, & \text{otherwise.} \end{cases}$$

The meaning of the existence of $t_{\mathcal{B}}(i)$ is similar as in Remark 2.3.

3 Proof of Theorem 1.2

We start with an outline of the proof. First, we redefine the sequence $\mathcal{B}$ by adding k to all its terms, *i.e.*, $b_n := b_n + k$. For an even integer k we construct a new sequence $\mathcal{C} = \{c_n\}_{n=1}^{\infty}$ setting $c_{2l-1} = a_l$ and $c_{2l} = b_l$, for $l \in \mathbb{N}$. Note, $\mathcal{C}$ is a sequence defined on symbols $\{1, 2, \ldots, 2k\}$. It is easy to see that, if $c_i < c_j$, then $c_i, c_j \in \mathcal{A}$, or $c_i, c_j \in \mathcal{B}$, or $c_i \in \mathcal{A}$ and $c_j \in \mathcal{B}$. Hence,

$$S_{\mathcal{C}}(2k, 2l) = S_{\mathcal{A}}(k, l) + S_{\mathcal{B}}(k, l) + \left|\{(a_i, b_j) \mid a_i \in \{a_n\}_{n=1}^{l}, b_j \in \{b_n\}_{n=1}^{l}, 1 \le i \le j \le l\}\right|.$$

Consequently,

$$\frac{S_{\mathcal{C}}(2k, 2l)}{(2l)^2} \ge \frac{1}{4}\left(\frac{S_{\mathcal{A}}(k, l)}{l^2} + \frac{S_{\mathcal{B}}(k, l)}{l^2}\right) + \frac{1}{8}. \tag{3}$$

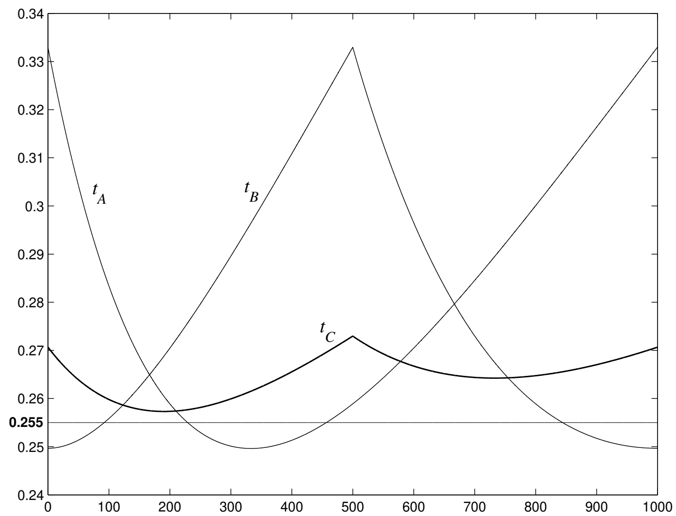

Figure 1: Density functions of sequences $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$.

Figure 1 describes the behavior of the density function for the sequences $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ on one block for $k = 1000$ and a large value of m. More precisely, it gives the graphs of $t_{\mathcal{A}}(1000)$, $t_{\mathcal{B}}(1000)$ and $t_{\mathcal{C}}(2000)$ as a function of $i = 1, \ldots, 1000$. Based on this, one can anticipate that $s_{\mathcal{C}}(2000) > 0.255$. Indeed, we will show that for k large enough $(k \geq 162)$ the sequence $\{c_n\}$ implies the statement of Theorem 1.2.

Proof of Theorem 1.2. In view of Lemma 1.5 we need to show that $s_{\mathcal{C}}(2k) > 0.255$. To this end, we will verify that the limit inferior of the right side of (3) is larger than 0.255, as l goes to infinity. For any l, there are integers m and i such that $n_i(m) \leq l < n_{i+1}(m)$. Since $S_{\mathcal{A}}(k, l)$ is increasing in l, we thus have

$$\frac{S_{\mathcal{A}}(k,l)}{l^2} \geq \frac{S_{\mathcal{A}}(k, n_i(m))}{(n_{i+1}(m))^2}, \tag{4}$$

and similarly,

$$\frac{S_{\mathcal{B}}(k,l)}{l^2} \geq \frac{S_{\mathcal{B}}(k, n_i(m))}{(n_{i+1}(m))^2}. \tag{5}$$

In view of Propositions 2.1 and 2.4 and equation (2), for any $i \in \{0, \ldots, k-1\}$, we have,

$$\lim_{m\to\infty} \frac{S_{\mathcal{A}}(k, n_i(m))}{(n_{i+1}(m))^2} = \frac{\frac{k(k-1)}{3} + i(i-1)}{(k+i+1)^2},$$

and

$$\lim_{m\to\infty} \frac{S_{\mathcal{B}}(k, n_i(m))}{(n_{i+1}(m))^2} = \begin{cases} \frac{k^2 - \frac{4}{3}k + 2i(k+2i-2)}{4(k+i+1)^2}, & \text{if } i \in \{0, \ldots, \frac{k}{2} - 1\}, \\ \frac{3k^2 - \frac{10}{3}k + (2i-k-2)(2i-k)}{4(k+i+1)^2}, & \text{if } i \in \{\frac{k}{2}, \ldots, k-1\}. \end{cases}$$

Consequently,

$$\liminf_{l\to\infty}\left(\frac{S_{\mathcal{A}}(k,l)}{l^2}+\frac{S_{\mathcal{B}}(k,l)}{l^2}\right)\geq \min_i\left\{\lim_{m\to\infty}\frac{S_{\mathcal{A}}(k,n_i(m))}{(n_{i+1}(m))^2}+\lim_{m\to\infty}\frac{S_{\mathcal{B}}(k,n_i(m))}{(n_{i+1}(m))^2}\right\}.$$

Hence, it suffices to verify that for any $i\in\{0,\ldots,\frac{k}{2}-1\}$

$$\frac{\frac{k(k-1)}{3}+i(i-1)}{(k+i+1)^2}+\frac{k^2-\frac{4}{3}k+2i(k+2i-2)}{4(k+i+1)^2}>0.52, \tag{6}$$

and for any $i\in\{\frac{k}{2},\ldots,k-1\}$

$$\frac{\frac{k(k-1)}{3}+i(i-1)}{(k+i+1)^2}+\frac{3k^2-\frac{10}{3}k+(2i-k-2)(2i-k)}{4(k+i+1)^2}>0.52, \tag{7}$$

for k large enough. Multiplying both sides of (6) and (7) by $(k+i+1)^2$ one can verify that inequality (6) holds for $k\geq 162$ and inequality (7) for $k\geq 35$. □

4 Proof of Theorem 1.3

To prove Theorem 1.3 we need to refine some of the estimates made above. Observe that our main "tool" was the fact that for any l there are integers $m\in\mathbb{N}$ and $i\in\{0,\ldots,k-1\}$ such that $n_i(m)\leq l<n_{i+1}(m)$, and consequently that (4) and (5) hold. In order to strengthen inequalities (4) and (5) we choose an integer r, which is a power of 2, and further subdivide the interval $\{n_i(m),\ldots,n_{i+1}(m)-1\}$ into r disjoint intervals of the same length, *i.e.*,

$$\{n_i(m),\ldots,n_{i+1}(m)-1\}=\bigcup_{j=0}^{r-1}\left\{n_i(m)+\frac{j}{r}2^{m+1},\ldots,n_i(m)+\frac{j+1}{r}2^{m+1}-1\right\}.$$

Let $n_i^j(m)=n_i(m)+\frac{j}{r}2^{m+1}$ for $j\in\{0,1,\ldots,r\}$. Note, $n_i^r(m)=n_{i+1}^0(m)$ and

$$n_i^j(m)=\left(2k+2i+2\frac{j}{r}\right)2^m+o(2^m). \tag{8}$$

Then, the following two statements hold.

Proposition 4.1 *Let r be a power of 2. Then, for any $i\in\{0,1,\ldots,k-1\}$ and $j\in\{0,1,\ldots,r-1\}$ we have*

$$S_{\mathcal{A}}(k,n_i^j(m))=\left(\frac{4}{3}k(k-1)+4i(i-1)+8i\frac{j}{r}\right)4^m+o(4^m),$$

for m sufficiently large.

Proof. Note that

$$\begin{aligned} S_{\mathcal{A}}(k, n_i^j(m)) &= S_{\mathcal{A}}(k, n_i(m)) + i(2 + 2^2 + \cdots + 2^{m+1})\frac{j}{r}2^{m+1} \\ &= S_{\mathcal{A}}(k, n_i(m)) + 8i\frac{j}{r}4^m + o(4^m), \end{aligned}$$

which in view of Proposition 2.1 yields the required statement. □

Proposition 4.2 *Let r be a power of 2. Then, for any $i \in \{0, 1, \ldots, k-1\}$ and $j \in \{0, 1, \ldots, r-1\}$ we have*

$$S_{\mathcal{B}}(k, n_i^j(m)) = \begin{cases} \left(k^2 - \frac{4}{3}k + 2i(k + 2i - 2) + \frac{j}{r}(2k + 8i)\right) 4^m + o(4^m), & \text{if } 0 \le i \le \frac{k}{2} - 1, \\ \left(3k^2 - \frac{10}{3}k + (2i - k - 2)(2i - k) + \frac{j}{r}(8i - 4k)\right) 4^m + o(4^m), & \text{otherwise.} \end{cases}$$

for m sufficiently large.

Proof. For $i \in \{0, \ldots, \frac{k}{2} - 1\}$ we get

$$\begin{aligned} S_{\mathcal{B}}(k, n_i^j(m)) &= S_{\mathcal{B}}(k, n_i(m)) + \left(\frac{k}{2} + i\right)(2 + 2^2 + \cdots + 2^m)\frac{j}{r}2^{m+1} + i2^{m+1}\frac{j}{r}2^{m+1} \\ &= S_{\mathcal{B}}(k, n_i(m)) + \frac{j}{r}(2k + 8i)4^m + o(4^m). \end{aligned}$$

Similarly, for $i \in \{\frac{k}{2}, \ldots, k-1\}$, we have

$$\begin{aligned} S_{\mathcal{B}}(k, n_i^j(m)) &= S_{\mathcal{B}}(k, n_i(m)) + \left(i - \frac{k}{2}\right)(2 + 2^2 + \cdots + 2^{m+1})\frac{j}{r}2^{m+1} \\ &= S_{\mathcal{B}}(k, n_i(m)) + \frac{j}{r}(8i - 4k)4^m + o(4^m). \end{aligned}$$

The required statement follows now from Proposition 2.4. □

Based on the above propositions we will prove Theorem 1.3.

Proof of Theorem 1.3. In view of Lemma 1.5 we need to show that $s_{\mathcal{C}}(16) > \frac{1}{4}\left(1 - \frac{1}{16}\right)$. In order to prove it, we will show that the limit inferior of the right side of (3) is strictly greater than $\frac{1}{4}\left(1 - \frac{1}{16}\right)$, as l goes to infinity. For any l there are integers $m \in \mathbb{N}$, $i \in \{0, \ldots, k-1\}$ and $j \in \{0, \ldots, r-1\}$ such that $n_i^j(m) \le l < n_i^{j+1}(m)$, and consequently

$$\frac{S_{\mathcal{A}}(k, l)}{l^2} \ge \frac{S_{\mathcal{A}}(k, n_i^j(m))}{(n_i^{j+1}(m))^2}, \tag{9}$$

and similarly,

$$\frac{S_{\mathcal{B}}(k, l)}{l^2} \ge \frac{S_{\mathcal{B}}(k, n_i^j(m))}{(n_i^{j+1}(m))^2}. \tag{10}$$

Let $k=8$ and $r=64$. Then, one can check[1] that for any $i \in \{0,\ldots,\frac{k}{2}-1\}$ and $j \in \{0,\ldots,r-1\}$

$$\frac{\frac{k(k-1)}{3}+i(i-1)+2i\frac{j}{r}}{\left(k+i+\frac{j+1}{r}\right)^2}+\frac{k^2-\frac{4}{3}k+2i(k+2i-2)+\frac{j}{r}(2k+8i)}{4\left(k+i+\frac{j+1}{r}\right)^2}>\frac{28}{64}, \tag{11}$$

and for any $i \in \{\frac{k}{2},\ldots,k-1\}$ and $j \in \{0,\ldots,r-1\}$

$$\frac{\frac{k(k-1)}{3}+i(i-1)+2i\frac{j}{r}}{\left(k+i+\frac{j+1}{r}\right)^2}+\frac{3k^2-\frac{10}{3}k+(2i-k-2)(2i-k)+\frac{j}{r}(8i-4k)}{4\left(k+i+\frac{j+1}{r}\right)^2}>\frac{28}{64}. \tag{12}$$

Hence, by Propositions 4.1 and 4.2, equation (8) and inequalities (9) and (10) we obtain

$$\liminf_{l\to\infty}\left(\frac{S_{\mathcal{A}}(k,l)}{l^2}+\frac{S_{\mathcal{B}}(k,l)}{l^2}\right)\geq\min_{i,j}\left\{\lim_{m\to\infty}\frac{S_{\mathcal{A}}(k,n_i^j(m))}{(n_i^{j+1}(m))^2}+\lim_{m\to\infty}\frac{S_{\mathcal{B}}(k,n_i^j(m))}{(n_i^{j+1}(m))^2}\right\}>\frac{28}{64},$$

which in view of (3) yields the statement of Theorem 1.3, *i.e.*,

$$s_{\mathcal{C}}(16)>\frac{1}{4}\cdot\frac{28}{64}+\frac{1}{8}=\frac{1}{4}\left(1-\frac{1}{16}\right).$$

□

Remark 4.3 Analogously, one can show that Conjecture 1.1 fails for any $k \geq 24$. In order to do it, take a sequence $\mathcal{C}$ of $2k$ symbols from the proof of Theorem 1.3, for $k \geq 12$ and even. Then an approach similar to the one used in Theorem 1.3 yields

$$s_{\mathcal{C}}(2k)>\frac{1}{4}\left(1-\frac{1}{2k+3}\right).$$

5 Proof of Theorem 1.4

For a given k, let $\mathcal{C}=\{c_n\}_{n=1}^{\infty}$ be a sequence of integers with $c_n \in \{1,\ldots,k\}$. Let $t>1$ be an integer and $\varepsilon=\frac{1}{t}$. Let

$$\mathbb{N}=\bigcup_{i=1}^{k}N^{(i)} \tag{13}$$

be a partition such that $c_n=i$ for every $n \in N^{(i)}$. We further subdivide each $N^{(i)}=\{n_1^i<n_2^i<\cdots\}$, $i=1,\ldots,k$, into

$$N^{(i)}=\bigcup_{j=1}^{t}N^{(i,j)},$$

[1] The authors used MATLAB [4] to verify (11) and (12).

where $N^{(i,j)} = \{n^i_j < n^i_{t+j} < n^i_{2t+j} < \cdots\}$. Now, we construct a new sequence $\mathcal{D} = \{d_n\}_{n=1}^{\infty}$ with values in $\{1, \ldots, \tilde{k}\}$, for $\tilde{k} = kt$, such that

$$d_n = (i-1)t + j,$$

for $n \in N^{(i,j)}$. Let

$$\mathcal{S}_{\mathcal{D}}(l) = \{(n_1, n_2) \mid 1 \leq n_1 < n_2 \leq l, i_1 < i_2 \text{ for } n_1 \in N^{(i_1)}, n_2 \in N^{(i_2)}\}.$$

Note,

$$S_{\mathcal{C}}(k, l) = |\mathcal{S}_{\mathcal{D}}(l)|. \tag{14}$$

For any $i \in \{1, \ldots, \tilde{k}\}$, let r_i be a function defined on $\mathbb{N}$ such that

$$r_i(l) = \frac{|\{n \in \mathbb{N} \mid d_n = i \text{ and } n \leq l\}|}{l}.$$

Consequently, for any $l \in \mathbb{N}$, we have $\big(r_1(l), \ldots, r_{\tilde{k}}(l)\big) \in [0, \varepsilon]^{\tilde{k}}$ and $\sum_{i=1}^{\tilde{k}} r_i(l) = 1$. Since the set $[0, \varepsilon]^{\tilde{k}}$ is compact, the sequence $\big\{\big(r_1(l), \ldots, r_{\tilde{k}}(l)\big)\big\}_{l=1}^{\infty}$ has a limit point in $[0, \varepsilon]^{\tilde{k}}$, say $\big(r_1, \ldots, r_{\tilde{k}}\big)$. Therefore, there exists a sequence $\{l_j\}_{j=1}^{\infty}$ such that

$$\lim_{j \to \infty} \big(r_1(l_j), \ldots, r_{\tilde{k}}(l_j)\big) = \big(r_1, \ldots, r_{\tilde{k}}\big).$$

Obviously, $\big(r_1, \ldots, r_{\tilde{k}}\big) \in [0, \varepsilon]^{\tilde{k}}$ and $\sum_{i=1}^{\tilde{k}} r_i = 1$. Thus, the set $\mathcal{R} = \{1, \ldots, \tilde{k}\}$ can be divided into 3 disjoint sets for some $i_1, i_2 \in \{1, \ldots, \tilde{k}\}$ as follows:

$$\begin{aligned}
\mathcal{R}_1 &= \{1, \ldots, i_1\}, \\
\mathcal{R}_2 &= \{i_1 + 1, \ldots, i_2\}, \\
\mathcal{R}_3 &= \{i_2 + 1, \ldots, \tilde{k}\},
\end{aligned}$$

and

$$\Big| \sum_{i \in \mathcal{R}_q} r_i - \frac{1}{3} \Big| < \varepsilon,$$

for $q = 1, 2, 3$. Consequently, there exists j_0 such that for all $j \geq j_0$ the similar inequality holds, *i.e.*,

$$\Big| \sum_{i \in \mathcal{R}_q} r_i(l_j) - \frac{1}{3} \Big| < \varepsilon, \tag{15}$$

for $q = 1, 2, 3$. Let $\mathcal{S}^1_{\mathcal{D}}(l)$ denote the set of pairs $(n_1, n_2) \in \mathcal{S}_{\mathcal{D}}(l)$ for which $d_{n_1}, d_{n_2} \in \mathcal{R}_q$, for some $q \in \{1, 2, 3\}$. Similarly, let $\mathcal{S}^2_{\mathcal{D}}(l)$ denote the set of pairs $(n_1, n_2) \in \mathcal{S}_{\mathcal{D}}(l)$ for which $d_{n_1} \in \mathcal{R}_{q_1}$ and $d_{n_2} \in \mathcal{R}_{q_2}$ for some $q_1 \neq q_2 \in \{1, 2, 3\}$. Hence,

$$|\mathcal{S}_{\mathcal{D}}(l)| = |\mathcal{S}^1_{\mathcal{D}}(l)| + |\mathcal{S}^2_{\mathcal{D}}(l)|. \tag{16}$$

Furthermore, since

$$|\mathcal{S}_{\mathcal{D}}^1(l_j)| \le \sum_{q=1}^{3} \binom{l_j \sum_{i\in\mathcal{R}_q} r_i(l_j)}{2},$$

and (15) we infer that

$$\limsup_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}^1(l_j)|}{{l_j}^2} = \frac{1}{3^2} \limsup_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}^1(l_j)|}{\left(\frac{l_j}{3}\right)^2} \le \frac{1}{3^2}\cdot\frac{1}{2}\cdot 3 + O(\varepsilon) = \frac{1}{6} + O(\varepsilon). \tag{17}$$

On the other hand, due to the result of J. Czipszer, P. Erdős and A. Hajnal (see Theorem 2 in [1]), which states that Conjecture 1.1 is true for $k=3$, we obtain that

$$\liminf_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}^2(l_j)|}{{l_j}^2} \le \frac{1}{4}\left(1-\frac{1}{3}\right) = \frac{1}{6}. \tag{18}$$

Consequently, by (14), (16), (17) and (18) we get

$$\begin{aligned}\liminf_{l\to\infty} \frac{S_{\mathcal{C}}(k,l)}{l^2} = \liminf_{l\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}(l)|}{l^2} &\le \liminf_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}(l_j)|}{{l_j}^2} \\ &\le \limsup_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}^1(l_j)|}{{l_j}^2} + \liminf_{j\to\infty} \frac{|\mathcal{S}_{\mathcal{D}}^2(l_j)|}{{l_j}^2} \le \frac{1}{3} + O(\varepsilon).\end{aligned}$$

Since this is true with ε arbitrarily small, we infer that

$$s_{\mathcal{C}}(k) \le \frac{1}{3}.$$

This completes the proof of Theorem 1.4. □

Remark 5.1 Slightly modifying the above proof, one can show that for given k and m we have

$$p(k) \le p(m) + \frac{1}{2m}. \tag{19}$$

Indeed, note that the set $\{1,\ldots,\tilde{k}\}$ can be partitioned into m pairwise disjoint sets

$$\{1,\ldots,\tilde{k}\} = \bigcup_{q=1}^{m} \mathcal{R}_q,$$

which satisfy

$$\Big|\sum_{i\in\mathcal{R}_q} r_i(l_j) - \frac{1}{m}\Big| < \varepsilon,$$

for every $q\in\{1,\ldots,m\}$ and every $j\ge j_0$, for some fixed j_0. Consequently, the left sides of (17) and (18) can be bounded by $\frac{1}{2m}+O(\varepsilon)$ and $p(m)$, respectively. Hence,

$$s_{\mathcal{C}}(k) \le p(m) + \frac{1}{2m},$$

i.e., (19) holds.

Note that if Conjecture 1.1 would be true, say for $p(4)$, one could combine it with (19) to improve Theorem 1.4. More precisely, we would have

$$p(k) \leq \frac{1}{4}\left(1 - \frac{1}{16}\right) + \frac{1}{8} = \frac{5}{16},$$

for every $k \geq 2$.

6 Concluding remarks

One can extend the definition of $p(k)$ by allowing k to be also infinity, with the corresponding parameter

$$p(\infty) = \sup_{G^{(\infty)}} \left(\liminf_{l \to \infty} \frac{|E(G^{(l)})|}{l^2} \right),$$

where the supremum is taken over all graphs $G^{(\infty)}$ without infinite increasing paths. J. Czipszer, P. Erdős and A. Hajnal showed in [1] that

$$\frac{1}{4} + \frac{1}{36} \leq p(\infty) \leq \frac{1}{4} + \frac{3}{16}.$$

In this paper we showed that for k large enough the path Turán number satisfies

$$\frac{1}{4} + \frac{1}{200} \leq p(k) \leq \frac{1}{4} + \frac{1}{12}.$$

Determining the precise values of $p(k)$ and $p(\infty)$ does not seem to be an easy problem. However, it is easy to see that for any fixed k

$$p(k) < p(\infty). \tag{20}$$

Indeed, let $\mathcal{C} = \{c_n\}_{n=1}^{\infty}$ be a sequence of k symbols. Furthermore, let $\mathcal{D} = \{d_n\}_{n=1}^{\infty}$ be a sequence from the proof of Theorem 1.4 for some integer t. Then, $\mathcal{D}$ is a sequence of kt symbols. Recalling the definition of $N^{(i)}$ (*cf.* (13)), for $i = 1, \ldots, k$, set

$$l_i(l) = \left|\{\alpha \in N^{(i)} \mid \alpha \leq l\}\right|,$$

and

$$E_i(l) = \left|\{(\alpha, \beta) \in N^{(i)} \times N^{(i)} \mid \ d_\alpha < d_\beta \text{ and } 1 \leq \alpha < \beta \leq l\}\right|.$$

Since $N^{(i)}$ induces a subsequence of the form $1|2|\ldots|\mathrm{t}|1|2|\ldots|\mathrm{t}|1|2|\ldots|\mathrm{t}|\ldots$ we infer that $E_i(l) = \frac{1}{4}\left(1 - \frac{1}{t}\right)l_i(l)^2 + o\left(l_i(l)^2\right)$ and $\sum_{i=1}^{k} l_i(l) = l$. Consequently,

$$\sum_{i=1}^{k} E_i(l) = \frac{1}{4}\left(1 - \frac{1}{t}\right)\sum_{i=1}^{k}\left(l_i(l)^2 + o\left(l_i(l)^2\right)\right) \geq \frac{1}{4}\left(1 - \frac{1}{t}\right) k \left(\frac{l}{k}\right)^2 + o(l^2),$$

and

$$\frac{\sum_{i=1}^{k} E_i(l)}{l^2} \geq \frac{1}{4}\left(1-\frac{1}{t}\right)\frac{1}{k}+o(1).$$

Hence,

$$s_{\mathcal{D}}(kt) \geq s_{\mathcal{C}}(k)+\frac{1}{4}\left(1-\frac{1}{t}\right)\frac{1}{k},$$

and finally by Lemma 1.5 we obtain (20), *i.e.*,

$$p(k)=s(k)<s(kt)=p(kt)\leq p(\infty).$$

Letting $t \to \infty$, we have just showed that

$$p(k)+\frac{1}{4k} \leq p(\infty).$$

Acknowledgements

We would like to thank Jan Zich, who verified our computations of inequalities (6), (7), (11) and (12). We also owe special thanks to the referee for his or her very valuable comments.

References

[1] J. Czipszer, P. Erdős, and A. Hajnal, *Some extremal problems on infinite graphs*, Publications of the Math. Inst. of the Hungarian Academy of Sci. Ser. A **7** (1962), 441–456.

[2] P. Erdős, *Problems and results in combinatorial analysis*, Combinatorics (Proc. Symp. Pure Math., vol. XIX) Amer. Math. Soc., Providence, R.I., 1971, pp. 77–89.

[3] P. Erdős, *The Art of Counting, Selected Writings*, The MIT Press, Cambridge, 1973.

[4] The MathWorks, Inc., Matlab, `http://www.mathworks.com/products/matlab/`.