Combinatorics, Probability and Computing (2003) 12, 585–598. © 2003 Cambridge University Press
DOI: 10.1017/S096354830300590X Printed in the United Kingdom

On the Erdős–Simonovits–Sós Conjecture about the Anti-Ramsey Number of a Cycle

TAO JIANG[1†] and DOUGLAS B. WEST[2‡]

[1]Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA
(e-mail: `jiangt@muohio.edu`)

[2]Department of Mathematics, University of Illinois, Urbana, IL 61801, USA
(e-mail: `west@math.uiuc.edu`)

Received 15 July 2002; revised 23 July 2003

Given a positive integer n and a family $\mathscr{F}$ of graphs, let $f(n,\mathscr{F})$ denote the maximum number of colours in an edge-colouring of K_n such that no subgraph of K_n belonging to $\mathscr{F}$ has distinct colours on its edges. Erdős, Simonovits and Sós [6] conjectured for fixed k with $k \geqslant 3$ that $f(n,C_k)=\left(\frac{k-2}{2}+\frac{1}{k-1}\right)n+O(1)$. This has been proved for $k \leqslant 7$. For general k, in this paper we improve the previous bound of $(k-2)n-\binom{k-1}{2}$ to $f(n,C_k) \leqslant \left(\frac{k+1}{2}-\frac{2}{k-1}\right)n-(k-2)$. For even k, we further improve it to $\frac{k}{2}n-(k-2)$. We also prove that $f(n,\{C_k,C_{k+1},C_{k+2}\}) \leqslant \left(\frac{k-2}{2}+\frac{1}{k-1}\right)n-1$, which is sharp.

1. Introduction

A subgraph in a colouring of the edges of the complete graph K_n is *polychromatic* if the colours on its edges are distinct; it is a *polychromatic copy of* H if it is also isomorphic to H. Let n be a positive integer, and let $\mathscr{F}$ be a family of graphs. We study the *anti-Ramsey number* $f(n,\mathscr{F})$; this is the maximum number of colours in a colouring of $E(K_n)$ that has no polychromatic copy of any graph in $\mathscr{F}$. (The classical 'Ramsey problem' can be interpreted as finding the minimum number of colours in a colouring of $E(K_n)$ that avoids monochromatic copies of graphs in $\mathscr{F}$.) We write $f(n,H)$ for $f(n,\{H\})$.

Erdős, Simonovits and Sós [6] introduced anti-Ramsey numbers. By relating them to Turán numbers, they showed that $f(n,\mathscr{F})/\binom{n}{2} \to 1-\frac{1}{r-1}$ as $n \to \infty$, where $r=\min\{\chi(H-e): e \in E(H) \text{ and } H \in \mathscr{F}\}$. This determines $f(n,H)$ asymptotically when $r \geqslant 3$.

[†] Research supported by Miami University Faculty Summer Research Grant.

[‡] This material is based upon work supported by the NSA under Award No. MDA904-03-1-0037, which requires the disclaimer that any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the NSA.

When $r = 2$, the limit yields only $f(n,H) = o(n^2)$. This leaves open the asymptotics of anti-Ramsey numbers for bipartite graphs, for graphs that become bipartite upon deletion of an edge, and for families of such graphs. Exact formulas or asymptotics are known for $f(n,H)$ when H is a path [12], a star [8], some types of trees [10], the family of all trees of fixed size [10], or $K_{2,t}$ [2, 7].

Erdős, Simonovits and Sós [6] initiated the study of $f(n,C_k)$. For fixed k with $k \geqslant 3$, they conjectured that $f(n,C_k) = \left(\frac{k-2}{2} + \frac{1}{k-1}\right)n + O(1)$ and proved this for $k = 3$. Alon [1] proved it for $k = 4$, showing that $f(n,C_4) = \lfloor 4n/3 \rfloor - 1$. It is proved for $k \leqslant 7$ in [9]. In Section 6 we explain the relationship between that result and our general bounds.

For general k, Alon [1] proved that $f(n,C_k) \leqslant (k-2)n - \binom{k-1}{2}$. In this paper, we improve this bound to $f(n,C_k) \leqslant \left(\frac{k+1}{2} - \frac{2}{k-1}\right)n - (k-2)$, and for even k we improve it further to $f(n,C_k) \leqslant \frac{k}{2}n - (k-2)$. We also prove that the bound conjectured for $f(n,C_k)$ does hold when we further restrict the colourings on $E(K_n)$ by also forbidding slightly longer cycles. In particular, we prove that $f(n,\{C_k,C_{k+1},C_{k+2}\}) \leqslant \left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - 1$.

2. Preliminaries

Given a graph G, we use $n(G)$ for $|V(G)|$, $e(G)$ for $|E(G)|$, and $G[S]$ for the subgraph induced by vertex set S. A *u,v-path* is a path with endpoints u and v. We use the following notions.

Definition. Given a graph G and a colouring c of $E(G)$, a *representing graph* for c is a spanning subgraph L of G having exactly one edge of each colour under c (L may have isolated vertices). For a family $\mathscr{F}$, an *$\mathscr{F}$-good* colouring is a colouring of the edges of a complete graph with no polychromatic copy of any graph in $\mathscr{F}$. We write *H-good* for $\{H\}$-good.

We begin with the Erdős–Simonovits–Sós construction. When $r = k-1$, the number of colours equals $\left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - 1$, and the construction never uses more than $\left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - 1$ colours. In [9], it is proved that $f(n,C_k) \leqslant \left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - 1$ when $k \leqslant 7$.

Theorem 2.1. ([6]) *If $n = (k-1)q + r$, where $1 \leqslant r \leqslant k-1$, then*

$$f(n,C_k) \geqslant \left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - \left(\frac{k-1-r}{2} + \frac{1}{k-1}\right)r.$$

Proof. Partition the vertices into sets $V_1,\dots,V_q$ of size $k-1$ and one set V_{q+1} of size r. The edges with endpoints in the same set receive $q\binom{k-1}{2} + \binom{r}{2}$ distinct colours. On the remaining edges are q additional colours $c_1,\dots,c_q$, with colour $c_{\min\{i,j\}}$ on the edges with endpoints in V_i and V_j when $i \neq j$. The total $q\binom{k-1}{2} + \binom{r}{2} + q$ equals the given formula when $n = (k-1)q + r$.

Each V_i is too small to contain C_k, and every cycle that visits more than one of the subsets has two edges of colour c_i, where V_i is the smallest-indexed set that it visits. Hence the colouring is C_k-good. □

In this construction, deleting a polychromatic copy of K_{k-1} yields the analogous construction for $n-k+1$ vertices. The difference in the number of colours is $\binom{k-1}{2}+1$, which equals $\left(\frac{k-2}{2}+\frac{1}{k-1}\right)(k-1)$. The idea of deleting (roughly) $k-1$ vertices and using induction motivates our proof. We use this approach in proving the conjecture for $k \leqslant 4$, which also serves as a basis for induction on k in our general result.

Lemma 2.2. ([1]) *Every colouring of $E(K_n)$ having no polychromatic k-cycle or l-cycle also has no polychromatic $(k+l-2)$-cycle. In particular, every C_k-good colouring is also $C_{2+s(k-2)}$-good, for every positive integer s.*

Proof. Let C be a $(k+l-2)$-cycle in such a colouring. Let x and y be vertices at distance $k-1$ along C. The edge xy completes cycles of lengths k and l with the two x,y-paths along C. Since neither of these is polychromatic, C is also not polychromatic.

The second claim is now immediate by induction on s. □

The proof of Lemma 2.2 is essentially the same as in Alon [1]. Theorem 2.3 was proved for C_3 in [6] and for C_4 in [1]; the proof here is simpler.

Theorem 2.3. ([6, 1]) $f(n, C_3) = n-1$ *and* $f(n, C_4) = \lfloor 4n/3 \rfloor - 1$.

Proof. For $k \leqslant 4$, the construction of Theorem 2.1 uses exactly $\lfloor \left(\frac{k-2}{2}+\frac{1}{k-1}\right)n \rfloor - 1$ colours.

For $k=3$, Lemma 2.2 implies that a C_3-good colouring has no polychromatic cycles of any length. Hence a representing graph has at most $n-1$ edges.

For $k=4$, we use induction on n. If $n \leqslant 3$, then $\binom{n}{2} \leqslant \lfloor 4n/3 \rfloor - 1$. For $n \geqslant 4$, let c be a C_4-good colouring of $E(K_n)$. If c uses more than $n-1$ colours, then c has a polychromatic 3-cycle C. Let H be a representing graph containing C. By Lemma 2.2, H has only odd cycles.

Let F be a component of $H - V(C)$. Let $V(C) = \{u, v, w\}$. If $N_H(u)$ and $N_H(v)$ both intersect $V(F)$, then H has three pairwise internally disjoint u,v-paths, which yields an even cycle. If $x, y \in N_H(u) \cap V(F)$, then the nontrivial x,y-path P in F must have odd length. Now v, w, u, x and v, u, y, P form edge-disjoint v,x-paths of odd length in H. Adding vx completes a polychromatic even cycle in c with one of them.

Thus H has at most one edge from C to each component F of $H - V(C)$. By the induction hypothesis, $e(F) \leqslant \lfloor 4n(F)/3 \rfloor - 1$. Summing over components of $H - V(C)$ and counting edges to $V(C)$ yields

$$e(H) \leqslant 3 + \sum_F \lfloor 4n(F)/3 \rfloor \leqslant 3 + \lfloor 4(n-3)/3 \rfloor = \lfloor 4n/3 \rfloor - 1. \qquad \square$$

In applying induction on the number of vertices, we will want to limit the number of edges between a cycle and the rest of the vertices in a representing graph when some cycle lengths are forbidden. Our basic lemma for this setting is of independent interest.

For vertex-disjoint subgraphs J, J' in a graph G, let $E_G[J, J']$ denote the set of edges having one endpoint in $V(J)$ and the other in $V(J')$, and let $e_G(J, J')$ denote its size.

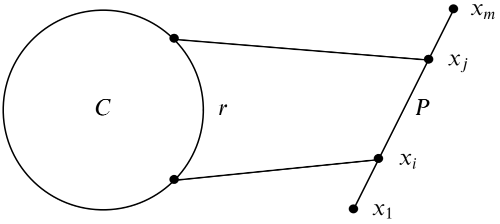

Figure 1. Disjointness of shifted neighbourhoods

We write J as v when the subgraph is a single vertex v. We drop the subscript when only one graph is being discussed.

Lemma 2.4. *Let C be a cycle of length p in a graph G, and let P be a path in $G - V(C)$. If G has no cycle whose length is congruent to 2 modulo p, then $e(P, C) \leqslant p$.*

Proof. Let $x_1, \ldots, x_m$ be the vertices of P, in order, and let $N_i = N(x_i) \cap V(C)$. With respect to a consistent orientation of C, let N_i^s denote the shift of N_i by s positions. If $j - i \equiv r \bmod p$ with $0 \leqslant r \leqslant p-1$ and there is a vertex in $N_i^{m-i} \cap N_j^{m-j}$, then x_i and x_j have neighbours on C that are separated by distance r along C. Replacing this portion of C with the edges from its endpoints to x_i and x_j and the x_i, x_j-path along P yields a cycle in G with length congruent to 2 modulo p (see Figure 1).

Therefore, the sets $N_1^{m-1}, N_2^{m-2}, \ldots, N_m^0$ are pairwise disjoint. Since they all lie in $V(C)$, their sizes sum to at most p. Since $|N_i^s| = |N_i|$, the sum of their sizes is $e(P, C)$. □

The hypothesis on G in Lemma 2.4 can be weakened when a bound is placed on the length of P. That is, the result $e(P, C) \leqslant p$ holds whenever G has no $(p+2)$-cycle if P has length at most p. We will need only the form proved above.

3. A greedy structure

Since $\left(\frac{k-2}{2} + \frac{1}{k-1}\right)$ increases with k, in our induction step we can restrict attention to colourings of $E(K_n)$ with polychromatic cycles shorter than C_k. We will use length $k-2$ when $\mathscr{F} = \{C_k\}$ and length $k-1$ when $\mathscr{F} = \{C_k, C_{k+1}, C_{k+2}\}$. We focus first on $\mathscr{F} = \{C_k\}$.

Definition. Let C be a polychromatic $(k-2)$-cycle in a C_k-good colouring c of $E(K_n)$, and let H be a representing graph containing C. Let a be the number of chords of C in H, and let b be the number of edges of H with exactly one endpoint in $V(C)$. The list (C, H, F) is *C_k-greedy in c* if C and H are chosen to lexicographically maximize the ordered pair (a, b) and F is a component of $H - V(C)$ with maximum order.

From the defining conditions, a C_k-greedy (C, H, F) has the following properties.

(1) Every colour on an edge of K_n induced by $V(C)$ is on some edge of H induced by $V(C)$.
(2) Every colour on an edge of K_n incident to $V(C)$ is on some edge of H incident to $V(C)$.
(3) No colour appearing in F appears on any edge of K_n incident to $V(C)$.

When we use these properties, we say 'by greediness'.

Property (1) of greediness is used in Theorem 5.5 and in the proofs of the optimal bounds for $k \leqslant 7$ in [9]. Property (2) is used heavily in the subsequent lemmas here. Most of the lemmas do not use property (3); we use this when reducing the proof of the bound for general k (Theorem 4.1) to the case where $H - V(C)$ has no edges. The variations in the lemmas for even k are used for the improvement of the general bound for even k (again Theorem 4.1).

Lemma 3.1. *Let (C, H, F) be C_k-greedy in c. Let P be an x, y-path in F, and let u be a vertex of C. If the length of P is a multiple of $k - 2$, then $c(ux) = c(uy)$. If k is even, $ux \in E(H)$, and the length of P is an odd multiple of $(k - 2)/2$, then $c(ux) = c(u'y)$, where u' is the vertex opposite u on C.*

Proof. The first statement does not assume that ux or uy lies in $E(H)$. Nevertheless, greediness implies that $c(ux)$ and $c(uy)$ do not appear on the edges of P. If $c(ux) \neq c(uy)$, then adding ux and uy to P produces a polychromatic cycle whose length is congruent to 2 modulo k. Lemma 2.2 forbids this, so $c(ux) = c(uy)$.

For the second statement, greediness again forbids $c(u'y)$ from the edges of P. Also, $c(u'y)$ appears on at most one of the u, u'-paths in C. If $c(ux) \neq c(u'y)$, then adding one of these paths plus ux and $u'y$ to P yields a polychromatic cycle forbidden by Lemma 2.2. □

Lemma 3.2. *Let (C, H, F) be C_k-greedy in c. Let $q = k - 2$ when k is odd and $q = (k - 2)/2$ when k is even. If F has a cycle C' of length at least q, then there exists H' such that (C, H', F) is C_k-greedy in c and all edges of $E_{H'}[C, F]$ are incident to $V(C')$.*

Proof. Let ux be an edge of H with $u \in V(C)$ and $x \in V(F) - V(C')$. Let P be a shortest path in F from x to $V(C')$. Since C' has length at least q, we can extend P along C' to obtain a path P' whose length is a multiple of q. Let y be the endpoint of P' on C'. By repeated application of Lemma 3.1, there is a vertex $v \in V(C)$ such that $c(ux) = c(vy)$. Replace ux with vy. Replacing all of $E_H[C, F - V(C')]$ yields the desired graph H'. □

The *circumference* of a graph is the length of its longest cycle (or is ∞ if the graph is acyclic). Our final tool is based on Woodall's proof [13] (see [3, pp. 137–8] for an exposition) of the Erdős–Gallai bound [5] on the number of edges in a n-vertex graph with circumference at most l. When W is a set of vertices in a graph G, let $e_G(W)$ denote the number of edges of G incident to W.

Lemma 3.3. *Let* (C,H,F) *be* C_k*-greedy in* c*, and let* $p=k-2$*. If* F *has circumference at most* l*, then* $e_H(W)\leqslant\lambda|W|$ *for some* $W\subseteq V(F)$*, where* $\lambda=\max\{1+\frac{p}{2},\frac{l}{2}+\frac{p}{l/2+1}\}$.

Proof. From the set of longest paths in F, choose P to lexicographically maximize the pair (a,d), where P is a u,v-path, $a=e_H(v,C)$, and $d=d_F(u)$. Index the vertices of P as $x_1,\ldots,x_m$ in order, with $u=x_1$ and $v=x_m$.

Let $W=\{x_i\colon x_{i+1}\in N_F(x_1)\}$; note that $|W|=d$ and $x_1\in W$ and $v\notin W$. We claim that $e_H(W)\leqslant\lambda|W|$. Note that $e_H(W)=e_H(W,C)+e_F(W)$.

For $x_i\in W$, let Q_i be the path formed by the v,x_{i+1}-path in P, the edge $x_{i+1}x_1$, and the x_1,x_i-path in P. If x_i has a neighbour in F outside $V(P)$, then Q_i extends, contradicting the choice of P. Therefore, $N_F(W)\subseteq V(P)$. Since Q_i has the same length as P, the choice of x_1 yields $d_F(x_i)\leqslant d$. Hence $e_F(W)\leqslant d^2$.

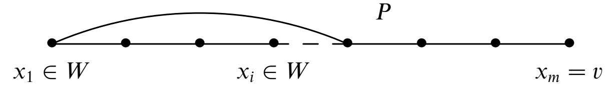

Figure 2. Longest paths from v in F

If x_i has a neighbour x_j with $j>l$ in F, then x_ix_j and $x_{i+1}x_1$ form a cycle of length j with portions of P, contradicting the circumference hypothesis. Therefore, both $W\subseteq L$ and $N(W)\subseteq L$, where $L=\{x_1,\ldots,x_l\}$. The sum $\sum_{x\in W}d_F(x)$ counts each edge of $F[W]$ twice and each edge of $E_F[W,L-W]$ once. Since $d_F(x)\leqslant d$ for $x\in W$, we have $2e_F(W)-e_F(W,L-W)\leqslant d^2$, which yields $e_F(W)\leqslant\frac{1}{2}(d^2+d(l-d))=dl/2$.

Now consider $e_H(W,C)$. By Lemmas 2.2 and 2.4, $e(P,C)\leqslant p$. Since v has the most edges to $V(C)$ among endpoints of longest paths in F, we have $e_H(W,C)\leqslant p\frac{d}{d+1}$.

We have shown that $e_H(W)\leqslant dg(d)$, where $g(d)=\min\{d,l/2\}+p/(d+1)$. We bound $g(d)$. In the range $d\geqslant l/2$, the maximum occurs when $d=l/2$. In the range $d\leqslant l/2$, the maximum occurs when $d=1$ or $d=l/2$. Hence $g(d)\leqslant\max\{1+\frac{p}{2},\frac{l}{2}+\frac{p}{l/2+1}\}$. ☐

4. The bound for general *k*

Our general bound for odd k is valid for all k, but for even k we prove a stronger bound.

Theorem 4.1. *If* $n\geqslant 2$*, then* $f(n,C_k)\leqslant\beta_kn-(k-2)$*, where* $\beta_k=(\frac{k+1}{2}-\frac{2}{k-1})$ *for odd* k *and* $\beta_k=k/2$ *for even* k.

Proof. We use induction on $n+k$. By Theorem 2.3, we may assume that $k\geqslant 5$. When $2\leqslant n\leqslant k-1$, we have $f(n,C_k)=\binom{n}{2}$. Here it suffices to show that $\beta_kn-(k-2)-\binom{n}{2}\geqslant 0$. The left side is $2\beta_k-k+1$ when $n=2$ and is $1+(2\beta_k-k)(k-1)/2$ when $n=k-1$. Both values are nonnegative when $k\geqslant 3$ whether k is odd or even, and hence this quadratic inequality holds for $2\leqslant n\leqslant k-1$.

Hence we may assume that $n\geqslant k\geqslant 5$. Let c be a C_k-good colouring of $E(K_n)$. Since the desired bound exceeds $f(n,C_{k-2})$, we may assume that c has a polychromatic $(k-2)$-cycle.

Hence we may select a C_k-greedy (C, H, F). We define *long cycle* to be a cycle of length at least q as defined in Lemma 3.2 ($q = k-2$ when k is odd, and $q = (k-2)/2$ when k is even).

Case 1: F has a long cycle C'.
By Lemma 3.2, we may assume that all of $E_H[C, F]$ is incident to $V(C')$. Since $V(C')$ lies on a path in F, Lemma 2.4 yields $e_H(C, F) \leqslant k-2$.

Consider the colourings obtained by restricting c to $V(F)$ and to $V(H) - V(F)$. Since $n(F) \geqslant 2$ and $n(H - V(F)) \geqslant 2$, the induction hypothesis applies. Since F and $H - V(F)$ are contained in representing graphs for these colourings, we have the desired bound:

$$e(H) = e(F) + e(H - V(F)) + e_H(C, F) \leqslant \beta_k n - (k-2).$$

Case 2: F has a cycle but no long cycle.
We apply Lemma 3.3 with $l = q-1$ and $p = k-2$. Now $1 + p/2 = k/2$. The quantity $\frac{l}{2} + \frac{p}{l/2+1}$ reduces to β_k when k is odd and to something at most β_k when k is even and at least 8. This case cannot occur when $k = 6$, because then $q = 2$ and every cycle is long. Hence $\max\{1 + \frac{p}{2}, \frac{l}{2} + \frac{p}{l/2+1}\} = \beta_k$ for $k \geqslant 5$, and we obtain $W \subseteq V(F)$ such that $e_H(W) \leqslant \beta_k |W|$. We discard W and apply the induction hypothesis to the restriction of c to $V(H) - W$ to obtain the desired bound on $e(H)$.

Case 3: F is a tree with at least two vertices.
Let v be a vertex of F having the most neighbours in H on C. Let P be a maximal path in F starting at v; let u be its other endpoint. By Lemma 2.4, $e(P, C) \leqslant k-2$. The choice of v yields $e_H(u, C) \leqslant (k-2)/2$. Also u has exactly one neighbour in F, so $d_H(u) \leqslant k/2$. We let $W = \{u\}$ and delete W as in Case 2.

Case 4: $H - V(C)$ has no edges.
The greedy choice of F makes this the only remaining case. Since $H - V(C)$ has no edges, $d_H(u) = e_H(u, C)$ when $u \notin V(C)$. Since $n \geqslant k$ and C is a $(k-2)$-cycle, we can choose distinct vertices u and v outside C. If $d_H(u) > k/2$, then we can choose distinct vertices $x, y \in N_H(u)$, with successors x', y' on C in a consistent orientation of C. If $x', y' \in N_H(v)$, then replacing xx' and yy' on C with $\{xu, uy, x'v, vy'\}$ yields a forbidden k-cycle in H. Hence at most one vertex of C can be both a neighbour of v and a successor of a neighbour of u. This yields $d_H(u) + d_H(v) \leqslant k-1$. Hence $d_H(u) \leqslant k/2$ or $d_H(v) \leqslant k/2$, and we finish as in Case 3. □

5. Forbidding cycles of lengths k through $k+2$

Let $\mathscr{C}_k = \{C_k, C_{k+1}, C_{k+2}\}$. We prove that $f(n, \mathscr{C}_k) \leqslant (\frac{k-2}{2} + \frac{1}{k-1})n - 1$. This is optimal infinitely often, since the construction of Theorem 2.1 shows that equality holds whenever $k-1$ divides n. The proof uses many of the ideas in our general bound on $f(n, C_k)$, but we need analogues of the lemmas in that proof. We say that (C, H, F) is *$\mathscr{C}_k$-greedy in c* if it is C_{k+1}-greedy in c; in particular, C will be a polychromatic $(k-1)$-cycle. We start with a stronger version of Lemma 2.4.

Lemma 5.1. *Let C be a cycle of length p in a graph G having no cycle with length greater than 3 and congruent to any of $\{1,2,3\}$ modulo p, and let P be a path in $G-V(C)$.*

(1) *If u is an endpoint of P, then $e(P,C)+e(u,C)\leqslant p$.*

(2) *If no two consecutive vertices on P have a common neighbour on C, then $e(P,C)\leqslant\lfloor p/2\rfloor$.*

(3) *If S is a subset of $V(P)$ with no two consecutive vertices on P, then $e(S,C)\leqslant\lfloor p/2\rfloor$.*

Proof. Let $x_1,\dots,x_m$ be the vertices of P, in order, and let $N_i=N(x_i)\cap V(C)$. With respect to a consistent orientation of C, let N_i^s denote the shift of N_i by s positions.

(1) As in Lemma 2.4, if there is a vertex in $N_i^{m-i}\cap N_j^{m-j}$, then G has a cycle with length congruent to 2 modulo p. Therefore, the sets $N_1^{m-1},N_2^{m-2},\dots,N_m^0$ are pairwise disjoint. Letting $u=x_m$, N_m^{-1} is also disjoint from all of these, since a common vertex in N_m^{-1} and N_j^{m-j} yields a cycle with length congruent to 1 modulo $p-1$. Since they all lie in $V(C)$, the sizes of these disjoint sets sum to at most p. Since $|N_i^s|=|N_i|$, the sum equals $e(P,C)+e(u,C)$.

(2) Since G has no $(p+1)$-cycle, for each i the sets N_i^r and N_i^{r+1} are disjoint. If $N_i^{m-i}\cup N_i^{m-i+1}$ and $N_j^{m-j}\cup N_j^{m-j+1}$ have a common vertex x, then $x\in N_i^{m-i+1}\cap N_j^{m-j}$ or $x\in N_i^{m-i}\cap N_j^{m-j+1}$. With $i<j$, the first case yields a cycle with length congruent to 1 modulo p. The second case yields a non-triangle cycle with length congruent to 3 modulo p, or has $i=j-1$ and yields consecutive vertices on P with a common neighbour on C. All these cases are forbidden by the hypotheses, so the sets $\{N_i^{m-i}\cup N_i^{m-i+1}\colon 1\leqslant i\leqslant m\}$ are pairwise disjoint subsets of $V(C)$. Each edge from P to C is counted exactly twice in these sets, so $e(P,C)\leqslant\lfloor p/2\rfloor$.

(3) If S contains no two consecutive vertices on P, then in the argument for (2) the case of consecutive vertices with a common neighbour cannot arise. Hence the sets $\{N_i^{m-i}\cup N_i^{m-i+1}\colon x_i\in S\}$ are pairwise disjoint, and $e(C,S)\leqslant\lfloor p/2\rfloor$. □

As in Lemma 2.4, in Lemma 5.1 the hypothesis on G can be weakened to having no cycle with length in $\{p,p+1,p+2\}$ when the length of P is at most $p-1$, but we only need the statement proved above.

We will apply Lemma 5.1 to a graph with an edge-colouring having no polychromatic cycle with length in $\{p+1,p+2,p+3\}$. By Lemma 2.2, such a colouring has no polychromatic non-triangle cycle with length congruent to one of $\{1,2,3\}$ modulo p, so we may apply Lemma 5.1 to a representing graph.

We also need variants of Lemmas 3.1 and 3.2. Henceforth, let $q=\lfloor(k-1)/2\rfloor$.

Lemma 5.2. *Let (C,H,F) be $\mathscr{C}_k$-greedy in c. Let P be an x,y-path in F, and let u be a vertex of C. If the length of P is a multiple of $k-1$, then $c(ux)=c(uy)$. If $ux\in E(H)$, and the length of P is congruent to q modulo $k-1$, then $c(ux)=c(u'y)$, where u' is a vertex at distance q from u on C.*

Proof. The first claim holds by Lemma 3.1 because (C,H,F) is C_{k+1}-greedy in c.

For the second claim, the first allows us to assume that P has length q. Now C_{k+1}-greediness forbids $c(u'y)$ from the edges of P, and $c(u'y)$ can appear on only one of the two u, u'-paths in C. If $c(ux) \neq c(u'y)$, then adding one of these paths plus ux and $u'y$ to P produces a polychromatic cycle of length k or $k+1$, both of which are forbidden. □

Lemma 5.3. *Let (C, H, F) be $\mathscr{C}_k$-greedy in c. If F has a cycle C' of length at least q, then there exists H' such that (C, H', F) is C_k-greedy in c and all of $E_{H'}[C, F]$ is incident to C'.*

Proof. Let ux be an edge of H with $u \in V(C)$ and $x \in V(F) - V(C')$. Let P be a shortest path in F from x to $V(C')$. Since C' has length at least q, we can extend P along C' to obtain a path P' whose length is congruent to q modulo $k-1$. Let y be the endpoint of P' on C'. By Lemma 5.2, $c(ux) = c(u'y)$, where u' is a vertex at distance q from u in C. Replace ux with $u'y$. Replacing all of $E_H[C, F - V(C')]$ in this way yields the desired graph H'. □

Our next lemma extends early results of Ore [11] in the theory of Hamiltonian graphs. For completeness, we give a short self-contained inductive proof. Stronger results are known about panconnected graphs, where an n-vertex graph G is *panconnected* if, whenever $d \leqslant l \leqslant n-1$ for vertices u and v at distance d in G, there is a u, v-path of length l in G ([4] surveys certain types of sufficient conditions for this and many other related properties about paths and cycles in graphs).

Lemma 5.4. *Let G be an n-vertex graph. If $e(\overline{G}) \leqslant n-4$, then G has a u, v-path of length l whenever $u, v \in V(G)$ and $2 \leqslant l \leqslant n-1$. If $e(\overline{G}) \leqslant n-3$, then G has a spanning cycle.*

Proof. We use induction on n; the claim is immediate for $n = 4$.

For $n > 4$, suppose first that $e(\overline{G}) \leqslant n-4$. Since at most $n-4$ edges are missing and K_n has $n-2$ edge-disjoint u, v-paths of length 2, there is a u, v-path of length 2 in G.

If v is not isolated in $\overline{G}$, then $e(\overline{G-v}) \leqslant n-5$. Since $e(\overline{G}) \leqslant n-4$, in $G-u$ there is a neighbour x of v. By the induction hypothesis, for $2 \leqslant l \leqslant n-2$ there is a u, x-path of length l in $G-v$. Append v to obtain the desired path in G.

If v is isolated in $\overline{G}$, then $e(\overline{G-v}) \leqslant (n-1)-3$. By the induction hypothesis, $G-v$ has a spanning cycle. Append v to the end of a path of length $l-1$ along the cycle from u.

Finally, we need a spanning cycle when $e(\overline{G}) = n-3$. This yields $\delta(G) \geqslant 2$. Since the complement of a graph with maximum degree at most 1 has a spanning cycle, we may assume that some vertex x has degree at least 2 in $\overline{G}$. Select $y, z \in N_{\overline{G}}(x)$. Since $e(\overline{G-x}) \leqslant n-5$, we can add the path z, x, y to a spanning y, z-path in $G-x$ to complete a spanning cycle in G. □

To facilitate the inductive proof of our bound on $f(n, \mathscr{C}_k)$, we need a stronger bound in the case when a $\mathscr{C}_k$-greedy (C, H, F) has few edges in $H[V(C)]$. Recall that $q = \lfloor \frac{k-1}{2} \rfloor$.

Theorem 5.5. *If $k \geqslant 3$ and $n \geqslant 2$, then $f(n, \mathscr{C}_k) \leqslant \left(\frac{k-2}{2} + \frac{1}{k-1}\right)n - 1$. Furthermore, if the vertices of each polychromatic $(k-1)$-cycle in a $\mathscr{C}_k$-good colouring c of $E(K_n)$ induce* *edges with at most* $\binom{k-1}{2}-q$ *colours, then the number of colours used in* c *is at most* $\left(\frac{k-2}{2}+\frac{1}{k-1}\right)n-q$.

Proof. We use induction on $n+k$. Let $\alpha_k=\left(\frac{k-2}{2}+\frac{1}{k-1}\right)$.

When $2\leqslant n\leqslant k-1$, we have $f(n,\mathscr{C}_k)=\binom{n}{2}$. For $2\leqslant n\leqslant k-2$, we require the stronger bound, and $\binom{n}{2}\leqslant\alpha_k n-q$ requires $\frac{n-1}{2}+\frac{q}{n}\leqslant\frac{k-2}{2}+\frac{1}{k-1}$. Since $h_q(x)=\frac{x}{2}+\frac{q}{x+1}$ defines a convex function, it suffices to observe that the inequality holds for $n=2$ and $n=k-2$.

If $n=k-1$, then $\binom{n}{2}=\alpha_k n-1$, and the bound holds with equality. Furthermore, the stronger bound holds if the stronger condition holds. If $k\leqslant 4$, then $q=1$, and Theorem 2.3 yields both desired statements.

We may thus assume that $n\geqslant k\geqslant 5$. This yields $\alpha_{k-1}n-1\leqslant\alpha_k n-q$, so it suffices to consider colourings with polychromatic $(k-1)$-cycles. Hence we may let (C,H,F) be $\mathscr{C}_k$-greedy in c. Also let $H'=H[V(C)]$, and let *long cycle* mean a cycle of length at least $q+1$.

Case 1: F has a long cycle C'.
By Lemma 5.3, we may assume that all of $E_H[C,F]$ is incident to $V(C')$. Fix an orientation of C'.

For $x\in V(C')$, let x' denote its successor on C', and let y be the vertex q steps after x on C'. If $x,x'\in N_H(u)$ for some $u\in V(C)$, then let u' denote a vertex at distance q from u on C. By Lemma 5.2, $c(u'y)=c(ux)$. Since H is a representing graph, the colours on $u'y$, ux', and the x',y-path of length $q-1$ on C' are distinct and do not appear on C. Combining these edges with the u,u'-path of length $k-1-q$ on C yields a polychromatic k-cycle.

Hence we may assume that no two consecutive vertices on C' have a common neighbour in H on C. Applying Lemma 5.1(2) to a spanning path in C' yields $e_H(C',C)\leqslant q$.

Let $F'=H-V(F)$. Since $n(F)\geqslant 2$ and $n(F')\geqslant 2$, we can apply the induction hypothesis to the colourings obtained by restricting c to $V(F)$ and to $V(F')$. If $e(H')\leqslant\binom{k-1}{2}-q$, then greediness of (C,H,F) implies that both F and F' have no polychromatic $(k-1)$-cycle inducing more than $\binom{k-1}{2}-q$ edges. Hence we can apply the tighter bound in the induction hypothesis, obtaining

$$e(H)=e(F)+e(F')+e_H(C,F)\leqslant\alpha_k n(F)-q+\alpha_k n(F')-q+e_H(C,C')\leqslant\alpha_k n(H)-q.$$

If $e(H')>\binom{k-1}{2}-q$, then we obtain $e(F)\leqslant\alpha_k n(F)-1$ and $e(F')\leqslant\alpha_k n(F')-1$. If $e(F)\leqslant\alpha_k n(F)-q$, then applying the induction hypothesis as above yields $e(H)\leqslant\alpha_k n(H)-1$, which suffices. If $e(F)>\alpha_k n(F)-q$ and $n(F)\geqslant k$, then the inequality $\alpha_{k-1}n(F)-1\leqslant\alpha_k n(F)-q$ allows us to assume that F contains a $(k-1)$-cycle; we take this as C'. On the other hand, if $n(F)<k$, then certainly $n(C')\leqslant k-1$.

We are left with $e(F)\leqslant\alpha_k n(F)-1$ and $e(F')\leqslant\alpha_k n(F')-1$ and $n(C')\leqslant k-1$. To obtain the desired bound on $e(H)$, it suffices to show that $e_H(C,C')\leqslant 1$. We show first that all of $E_H[C',C]$ is incident to one vertex of C.

We have $e(\overline{H'})\leqslant q-1$. If $q-1\leqslant(k-1)-4$, then Lemma 5.4 applies. This inequality holds when $k\geqslant 6$. Suppose that ux and vy are edges of $E_H[C,C']$, with $u\neq v$ and $x,y\in V(C')$. Let r be the length of an x,y-path in C'. By Lemma 5.4, there is an x,y-path through xu and $V(C)$ and vy that has length l, for each l in $\{4, 5, \dots, k\}$. We have $r + l \equiv 2 \bmod (k-2)$ for some l in this set unless $r = k-3$. In this case, setting $l = 4$ gives us a polychromatic cycle of length $k+1$, which is congruent to 2 modulo $k-1$. Since c is C_k-good and C_{k+1}-good, this violates Lemma 2.2.

When $k = 5$, the argument still applies unless H' consists of a 4-cycle plus one chord uv and the edges from C' arrive at u and v. Now H' has no u,v-path of length 3. However, the remaining chord, whose colour appears in H', lies on two u,v-paths of length 3 sharing no other edge, and one of these paths can be used.

Hence all edges of $E_H[C, C']$ must be incident to a single vertex u in $V(C)$. If $e_H(u, C') > 1$, then choose $x, z \in N_H(u) \cap V(C')$. Let y be a vertex of C' reached by a path of length q from x along C'. By Lemma 5.2, $c(vy) = c(ux)$, where v has distance q from u along C. Now vy and uz have distinct colours that do not appear in H' or C'. Replacing ux with vy in H now contradicts the argument about edges arriving at a single vertex of C.

Case 2: F has no long cycle.
In this case, it suffices by the induction hypothesis to find a set $W \subseteq V(F)$ with $e_H(W) \leqslant \alpha_k |W|$ (as in Theorem 4.1). This proves simultaneously both the overall bound and the stronger bound needed when $e(H')$ is small. For $W \subseteq V(F)$, let $\beta(W) = e_H(W)/|W|$.

Among the paths in F with maximum length, let P (a u,v-path) be chosen to lexicographically maximize the pair (a, d), where $a = e_H(v, C)$ and $d = d_F(u)$. Index the vertices of P as $x_1, \dots, x_m$ in order, with $u = x_1$ and $v = x_m$.

If $d \geqslant q$, then F has a long cycle, so we may assume that $d \leqslant q-1$. If $m > q$, then by Lemma 5.2 we can shift all of $E_H(u, C)$ away from u. Only the d incident edges in F remain incident to u. We obtain $\beta(\{u\}) = d \leqslant q - 1 \leqslant \alpha_k$ for the new representing graph used in place of H. Hence we may assume that $m \leqslant q$. The maximality of P also yields $m > d$.

We consider subcases depending on the value of d.

Subcase 2.1: $d \geqslant 3$ and $m = d+1$.
Since $q \geqslant d+1 \geqslant 4$, this requires $k \geqslant 9$. Also, $m = d+1$ requires $v \in N_F(u)$, so $P + uv$ is a cycle. The cycle must span $V(F)$, since P has maximum length in F. Lemma 5.1(1) now yields $e_H(F, C) \leqslant k-1$. If $W = V(F)$, then

$$\beta(W) \leqslant \frac{\binom{d+1}{2} + k - 1}{|W|} = \frac{\binom{d+1}{2} + k - 1}{d+1} = \frac{d}{2} + \frac{k-1}{d+1}.$$

Consider again the function of d defined by $h_{k-1}(d) = \frac{d}{2} + \frac{k-1}{d+1}$. Over an interval, h_{k-1} is maximized at an endpoint. For $k \geqslant 9$, we compute that $h_{k-1}(3) \leqslant \alpha_k$ and $h_{k-1}(q-1) \leqslant \alpha_k$. Hence $\beta(W) \leqslant h_{k-1}(d) \leqslant \alpha_k$.

Subcase 2.2: $d \geqslant 3$ and $m > d+1$.
Since $q \geqslant m \geqslant d+2 \geqslant 5$, this requires $k \geqslant 11$. Let $W = \{x_i : x_{i+1} \in N_F(x_1)\}$. We have $|W| = d$ and $u \in W$ and $v \notin W$. For $w \in W$, there is a w,v-path of length $m-1$ with vertex set $V(P)$ (see Figure 2). Our choice of u and v thus yields $e_H(w, C) \leqslant e_H(v, C)$ and $d_F(w) \leqslant d$ for all $w \in W$. As in Lemma 3.3, $N_F(W) \subseteq V(P)$.

Computations as in Lemma 3.3 now yield $e_F(W) \leqslant dm/2$. By Lemma 5.1(1), $e_H(P, C) + e_H(v, C) \leqslant k-1$. Now $v \notin W$ yields $e_H(W, C) + e_H(v, C) + e_H(v, C) \leqslant k-1$. By the choice of v, this yields $e_H(W,C) \leqslant \frac{d}{d+2}(k-1)$. Thus

$$\beta(W) = \frac{1}{d}\left[e_F(W) + e_H(W,C)\right] \leqslant \frac{1}{d}\left[\frac{dq}{2} + \frac{d}{d+2}(k-1)\right] \leqslant \frac{k-1}{4} + \frac{k-1}{5}.$$

In the last inequality, we used $d \geqslant 3$. This bound simplifies to $\frac{k-2}{2} + \frac{1}{2} - \frac{k-1}{20}$, which is bounded by $\frac{k-2}{2}$ when $k \geqslant 11$.

Subcase 2.3: $d = 2$.
Let $W = \{x_i : x_{i+1} \in N_F(u)\}$; we have $W = \{u, w\}$ for some $w \in V(P) - \{u, v\}$. As before, there is a longest path in F from w to v, so $e_H(w,C) \leqslant e_H(v,C)$.

If $w \neq x_2$, then $q \geqslant m \geqslant 4$, which requires $k \geqslant 9$. Lemma 5.1(3) yields $e_H(W,C) \leqslant q$. Since $e_F(W) \leqslant 4$, we have $\beta(W) \leqslant (q+4)/2$. This is bounded by α_k when $k \geqslant 10$. If $k = 9$, then $m = 4$ and $uv \in E(F)$. Hence $V(F) = V(P)$, since otherwise we find a longer path in F. Since $d = 2$, $F \neq K_4$; hence $e(F) \leqslant 5$. Also $e_H(F,C) \leqslant k-1$, by Lemma 5.1(1). Thus $\beta(V(F)) \leqslant (5+k-1)/4 \leqslant \alpha_k$.

If $w = x_2$, then $q \geqslant m \geqslant 3$, which requires $k \geqslant 7$. Lemma 5.1(1) yields $e_H(P,C) + e_H(v,C) \leqslant k-1$. Thus $e_H(u,C) + e_H(w,C) + 2e_H(v,C) \leqslant k-1$. The choice of v yields $e_H(W,C) \leqslant q$. Since $e_F(W) \leqslant 3$, we have $\beta(W) \leqslant (q+3)/2$. This is bounded by α_k when $k \geqslant 8$.

For $k = 7$, we have $m \leqslant q = 3$ and thus $F = C_3$. By Lemma 5.1(1), we have $e_H(P,C) + e_H(v,C) \leqslant k-1$. Our choice of v thus yields $e_H(F,C) \leqslant \left\lfloor \frac{3}{4}(k-1) \right\rfloor$, and then $\beta(V(F)) \leqslant \left(\left\lfloor \frac{3}{4}(k-1) \right\rfloor + 3\right)/3 \leqslant \left(\frac{k-2}{2} + \frac{1}{k-1}\right)$.

Subcase 2.4: $d = 1$.
If $m \geqslant 3$ (which by $q \geqslant m$ requires $k \geqslant 7$), then u and v are not consecutive on P. Let $W = \{u, v\}$. By Lemma 5.1(3), $e_H(W,C) \leqslant q$. Since $e_H(u,C) \leqslant e_H(v,C)$, we have $e_H(u,C) \leqslant \lfloor q/2 \rfloor$. We have $\beta(\{u\}) \leqslant \lfloor q/2 \rfloor + 1 \leqslant \alpha_k$ (since $k \geqslant 7$).

If $m = 2$, then $V(F) = \{u, v\}$. By Lemma 5.1(1), $e_H(P,C) + e_H(v,C) \leqslant k-1$. Again $e_H(u,C) \leqslant e_H(v,C)$, so $e_H(P,C) \leqslant \left\lfloor \frac{2}{3}(k-1) \right\rfloor$. Thus, $\beta(V(F)) \leqslant \left(\left\lfloor \frac{2}{3}(k-1) \right\rfloor + 1\right)/2 \leqslant \alpha_k$.

Subcase 2.5: $d = 0$.
In this case, $H - V(C)$ has no edges, by the greedy choice of F. If $e(\overline{H'}) \leqslant q-1$, then as in Case 1 we conclude that $E_H[F,C]$ is incident to a single vertex of C. Thus $e(H) \leqslant \binom{k-1}{2} + (n-k+1) \leqslant n + \binom{k-2}{2} - 1 \leqslant \alpha_k n - 1$.

When $e(\overline{H'}) \geqslant q$, we need to prove the stronger bound. Since $H - V(C)$ has no edges, $d_H(u) = e_H(u,C)$ when $u \notin V(C)$. If $N_H(u)$ has two consecutive vertices on C, then we have a forbidden polychromatic k-cycle. Hence $d_H(u) \leqslant q$, which yields

$$e(H) \leqslant \binom{k-1}{2} - q + (n-k+1)q = nq - q + \frac{k-1}{2}(k-2-2q).$$

When k is even, $q = (k-2)/2$ and $\alpha_k = q + \frac{1}{k-1}$. The bound becomes $e(H) \leqslant \alpha_k n - q - \frac{n}{k-1} < \alpha_k n - q$, as desired.

When k is odd, $q = (k-1)/2$ and $\alpha_k = q - \frac{1}{2} + \frac{1}{k-1}$. The bound becomes $e(H) \leqslant \alpha_k n - q + \frac{n-k+1}{2} - \frac{n}{k-1}$, and we need to improve it. Since k is odd, $d_H(u) = q$ requires neighbours and nonneighbours of u to alternate on C. Suppose that this occurs for two vertices $u, v \notin V(C)$. If $N_H(u) \cup N_H(v) = V(C)$, then let w, x, y, z be successive on C with $w, y \in N_H(u)$; replacing $\{wx, yz\}$ with $\{wu, uy, xv, vz\}$ yields a polychromatic $(k+1)$-cycle. Hence $N_H(u) = N_H(v)$; now consider the edge uv. By greediness, $c(uv)$ appears on an edge yb of H with w, x, y, z, a successive on C. Regardless of whether $y \in N_H(u)$ and/or $b \in \{x, z, u, v\}$, in all cases we can replace $\{wx, xy\}$ or $\{xy, yz\}$ or $\{yz, za\}$ with a detour through $\{u, v\}$ to complete a polychromatic k-cycle.

Hence $d_H(u) \leqslant q - 1$ when $u \notin V(C)$, except for at most one vertex. This reduces the upper bound by $n - k$, which is sufficient. □

6. Concluding remarks

Our results suggest several approaches to proving the full Erdős–Simonovits–Sós conjecture. With Theorem 5.5, it suffices to show that an optimal C_k-good colouring also has no polychromatic $(k+1)$-cycle or $(k+2)$-cycle. This condition holds in the construction of Theorem 2.1.

Another approach is to study a greedy structure based on a polychromatic $(k-1)$-cycle as in Theorem 5.5. Again Lemma 5.4 makes it possible to bound $e_H(F, C)$ tightly when $e(\overline{H'}) \leqslant k - 5$, but there remain many cases when $e(\overline{H'})$ is larger. This approach is used in [9] to prove the conjecture for $k \leqslant 7$. For larger k, stronger results about nearly panconnected graphs may make it possible to handle the cases.

Acknowledgement

We thank an anonymous referee for many worthwhile suggestions and correction of several errors in the proof of the second theorem.

References

[1] Alon, N. (1983) On a conjecture of Erdős, Simonovits, and Sós concerning anti-Ramsey theorems. *J. Graph Theory* **1** 91–94.

[2] Axenovich, M. and Jiang, T. Anti-Ramsey numbers for small complete bipartite graphs. To appear in *Ars Combinatoria.*

[3] Bollobás, B. (1978) *Extremal Graph Theory*, Academic Press.

[4] Broersma, H., Ryjáček, Z. and Schiermeyer, I. (2000) Closure concepts: A survey. *Graphs Combin.* **16** 17–48.

[5] Erdős, P. and Gallai, T. (1959) On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungary* **10** 337–356.

[6] Erdős, P., Simonovits, M. and Sós, V. T. (1975) Anti-Ramsey theorems. In *Infinite and Finite Sets* (Keszthely, Hungary, 1973), Vol. 10 of *Coll. Math. Soc. J. Bolyai*, pp. 657–665.

[7] Jiang, T. (2002) Anti-Ramsey numbers of subdivided graphs. *J. Combin. Theory Ser. B* **85** 361–366.

[8] Jiang, T. (2002) Edge-colorings with no large polychromatic stars. *Graphs Combin.* **18** 303–308.

[9] Jiang, T., Schiermeyer, I. and West, D. B. The Erdős–Simonovits–Sós conjecture for $k \leqslant 7$. In preparation.
[10] Jiang, T. and West, D. B. Edge-colorings of complete graphs that avoid polychromatic trees. To appear in *Discrete Math.*
[11] Ore, O. (1963) Hamiltonian connected graphs. *J. Math. Pures. Appl.* **42** 21–27.
[12] Simonovits, M. and Sós, V. T. (1984) On restricted colorings of K_n. *Combinatorica* **4** 101–110.
[13] Woodall, D. (1972) Sufficient conditions for circuits in graphs. *Proc. Lond. Math. Soc.* **24** 739–755.