Sublinear Longest Path Transversals

James A. Long Jr.*[1], Kevin G. Milans†[1], and Andrea Munaro‡[2]

[1]Department of Mathematics, West Virginia University, USA
[2]School of Mathematics and Physics, Queen's University Belfast, UK

April 15, 2021

Abstract

We show that connected graphs admit sublinear longest path transversals. This improves an earlier result of Rautenbach and Sereni and is related to the fifty-year-old question of whether connected graphs admit longest path transversals of constant size. The same technique allows us to show that 2-connected graphs admit sublinear longest cycle transversals.

1 Introduction

A classical exercise in graph theory is to show that if P and Q are longest paths in a connected graph, then the vertex sets of P and Q have non-empty intersection (see [8], exercise 1.2.40). In 1966, Gallai [2] asked whether this result could be strengthened to assert that the family of all longest paths in a connected graph G has non-empty intersection. It turns out the answer is no, as shown by Walther [6] with a 25-vertex counterexample. A 12-vertex counterexample, due to Walther and Voss [7] and independently Zamfirescu [10], is obtained from the Petersen graph by replacing one vertex v with an independent set $\{v_1, v_2, v_3\}$ such that each v_i becomes an endpoint of an edge incident to v (see Figure 1).

Since Gallai's question has a negative answer, a single vertex is generally insufficient to meet every longest path in a connected graph G. A *longest path transversal* in G is a set of vertices that intersects every longest path. Such a set is a transversal in the hypergraph on $V(G)$ whose edges are the vertex sets of longest paths in G. Let $\mathrm{lpt}(G)$ be the minimum size of a longest path transversal in G. The graph G_0 in Figure 1 is a connected 12-vertex graph with $\mathrm{lpt}(G_0) = 2$. Grünbaum [3] constructed a connected 324-vertex graph G with $\mathrm{lpt}(G) = 3$. Soon afterward, Zamfirescu [10] found such a graph with 270 vertices. Walther [6] and Zamfirescu [9] asked if $\mathrm{lpt}(G)$ is bounded for connected graphs G, and this remains

*jalong@mix.wvu.edu
†milans@math.wvu.edu
‡a.munaro@qub.ac.uk

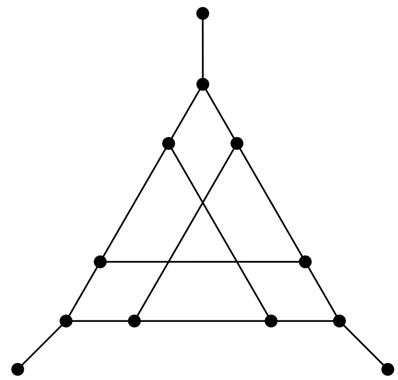

Figure 1: The graph G_0: a 12-vertex graph with $\mathrm{lpt}(G_0) = 2$.

open. In fact, it is not known whether there is a connected graph G with $\mathrm{lpt}(G) \geq 4$. Let G be a connected graph. Since a connected graph does not contain vertex-disjoint longest paths, every partition of $V(G)$ into two sets has a part that contains no longest path in G, forcing the other part to be a longest path transversal. Applying this to a partition of $V(G)$ into two parts of nearly equal size gives $\mathrm{lpt}(G) \leq \lceil n/2 \rceil$ when G is an n-vertex connected graph. It is not too difficult to improve this argument to obtain $\mathrm{lpt}(G) \leq \lceil n/4 \rceil$. Rautenbach and Sereni [4] showed that $\mathrm{lpt}(G) \leq \lceil \frac{n}{4} - \frac{n^{2/3}}{90} \rceil$ for every connected n-vertex graph G. We show that $\mathrm{lpt}(G) \leq 8n^{3/4}$ when G is an n-vertex connected graph, implying that connected graphs have sublinear longest path transversals.

Let $\mathrm{lct}(G)$ be the minimum size of a set of vertices S such that S intersects every longest cycle in G. Analogously to the case of longest paths in 1-connected graphs, every pair of longest cycles in a 2-connected graph intersect. The Petersen graph G is 2-connected and $\mathrm{lct}(G) = 2$. With no connectivity assumptions, Thomassen [5] showed that $\mathrm{lct}(G) \leq \lceil n/3 \rceil$ for each n-vertex graph G. The bound is sharp when G is a disjoint union of triangles and nearly sharp in the 1-connected case when G is obtained from a star with $(n-1)/3$ leaves by replacing each leaf with a triangle. On the other hand, Rautenbach and Sereni [4] proved that if G is 2-connected, then $\mathrm{lct}(G) \leq \lceil \frac{n}{3} - \frac{n^{2/3}}{36} \rceil$. We show that $\mathrm{lct}(G) \leq 20n^{3/4}$ when G is 2-connected (Corollary 2).

The problems of finding small longest path transversals and small longest cycle transversals are special cases of a general problem that we aim to address. Given a multigraph F and an edge $e \in E(F)$ with endpoints u and v, the *subdivision operation* produces a new multigraph F' in which e is replaced by a path uwv through a new vertex w in F'. A *subdivision* of F is a graph obtained from F via a sequence of zero or more subdivision operations. For a multigraph R and a graph G, an R-subdivision in G is a subgraph of G isomorphic to a subdivision of R. We ask for a small set of vertices in G that intersects every R-subdivision in G of maximum size. The cases of longest path transversals and longest cycle transversals arise as $R = P_2$ and $R = C_2$ (the multigraph 2-vertex cycle), respectively. We prove that for each connected multigraph R, if the family $\mathcal{F}$ of maximum R-subdivisions in G is pairwise intersecting, then $\mathcal{F}$ admits a transversal of size at most $Cn^{3/4}$, where C is a constant depending on R.

2 Maximum subdivision transversals

Let R be a multigraph. Recall that an R-subdivision in G is a subgraph of G isomorphic to a subdivision of R, and a *maximum R-subdivision* is an R-subdivision F in G that maximizes $|V(F)|$. An *R-transversal* of G is a set of vertices intersecting each maximum R-subdivision. Let $\tau_R(G)$ be the minimum size of an R-transversal in G.

Given sets of vertices X and Y of G, an *(X,Y)-separator* is a set of vertices S such that no path in $G-S$ has one endpoint in X and the other endpoint in Y. We allow an (X,Y)-separator to contain vertices in X and Y. An *(X,Y)-connector* is a collection of vertex-disjoint paths $\{P_1,\dots,P_k\}$ such that each P_i has one endpoint in X, the other endpoint in Y, and the interior vertices of P_i are outside $X \cup Y$. A variant of Menger's Theorem asserts that the minimum size of an (X,Y)-separator equals the maximum size of an (X,Y)-connector (see, e.g., Theorem 3.3.1 in [1]).

Our next result shows that when the maximum R-subdivisions in a graph G pairwise intersect, G has sublinear R-transversals. We make no attempt to optimize the multiplicative constant 8 or the dependence on m.

Theorem 1. *Let R be a connected m-edge multigraph with $m \ge 1$ and let G be an n-vertex graph. If the maximum R-subdivisions in G pairwise intersect, then $\tau_R(G) \le 8m^{5/4}n^{3/4}$.*

Proof. Let $m = |E(R)|$ and let $\varepsilon = 2(m/n)^{1/4}$. We may assume that $m \le n$, since otherwise we may take $V(G)$ as our R-transversal. Let $\mathcal{F}$ be the family of maximum R-subdivisions in G. An *ε-partial transversal* is a triple (H,X,Y) such that H is a subgraph of G, $X = V(G) - V(H)$, $Y \subseteq X$ with $|Y| \le \varepsilon|X|$, and each $F \in \mathcal{F}$ is a subgraph of H or contains a vertex in Y. Given an ε-partial transversal (H,X,Y), we either obtain an ε-partial transversal (H',X',Y') with $|V(H')| < |V(H)|$ or we produce an R-transversal with at most $8m^{5/4}n^{3/4}$ vertices. Starting with $(H,X,Y) = (G,\varnothing,\varnothing)$ and iterating gives the result.

Let (H,X,Y) be an ε-partial transversal, and let $\mathcal{F}_0$ be the set of $F \in \mathcal{F}$ such that F is a subgraph of H. We may assume that H contains vertex-disjoint paths P_1 and P_2 each of size $\lceil \varepsilon n \rceil$. Otherwise, every path in H has size less than $2\lceil \varepsilon n \rceil$, and so each $F \in \mathcal{F}_0$ has at most $2m\lceil \varepsilon n \rceil$ vertices. Since $\mathcal{F}_0$ is pairwise intersecting, we have that $V(F) \cup Y$ is an R-transversal for each $F \in \mathcal{F}_0$. It follows that $\tau_R(G) \le |Y| + 2m\lceil \varepsilon n \rceil \le \varepsilon n + 2m\lceil \varepsilon n \rceil \le (2m+1)\varepsilon n + 2m \le (2m+2)\varepsilon n \le 4m\varepsilon n = 8m^{5/4}n^{3/4}$.

Suppose that H has a $(V(P_1),V(P_2))$-separator S of size at most $\varepsilon^2 n$. Since graphs in $\mathcal{F}_0$ are connected, each $F \in \mathcal{F}_0$ has a vertex in S or is contained in some component of $H-S$. Also, since $\mathcal{F}_0$ is pairwise intersecting, at most one component H' of $H-S$ contains graphs in $\mathcal{F}_0$. Since S is a separator, H' is disjoint from at least one of $\{P_1,P_2\}$. With $X' = V(G) - V(H')$ and $Y' = Y \cup S$, we have $|X'| - |X| \ge \varepsilon n$ and $|Y'| = |Y| + |S| \le \varepsilon|X| + \varepsilon^2 n \le \varepsilon|X| + \varepsilon(|X'| - |X|) \le \varepsilon|X'|$. It follows that (H',X',Y') is an ε-partial transversal. Also $|V(H')| < |V(H)|$ since $|X'| > |X|$.

Otherwise, by Menger's Theorem, H has a $(V(P_1),V(P_2))$-connector $\mathcal{P}$ with $|\mathcal{P}| \ge \varepsilon^2 n$. Let $\mathcal{P}'$ be the set of paths in $\mathcal{P}$ of size at most $2/\varepsilon^2$. Note that $|\mathcal{P}'| \ge |\mathcal{P}|/2$, or else $\mathcal{P}$ has at least $(\varepsilon^2 n)/2$ paths of size more than $2/\varepsilon^2$, contradicting that the paths in $\mathcal{P}$ are disjoint. So we have $|\mathcal{P}'| \ge |\mathcal{P}|/2 \ge (\varepsilon^2/2)n = 2m^{1/2}n^{1/2} \ge 2$. Combining P_1 with two paths in $\mathcal{P}'$ whose endpoints in $V(P_1)$ are as far apart as possible and a segment of P_2 gives a cycle C_0 such that $(\varepsilon^2/2)n \le |V(C_0)| \le 2\lceil \varepsilon n \rceil + 4/\varepsilon^2 - 4 \le 2\varepsilon n + 4/\varepsilon^2$, where the lower bound

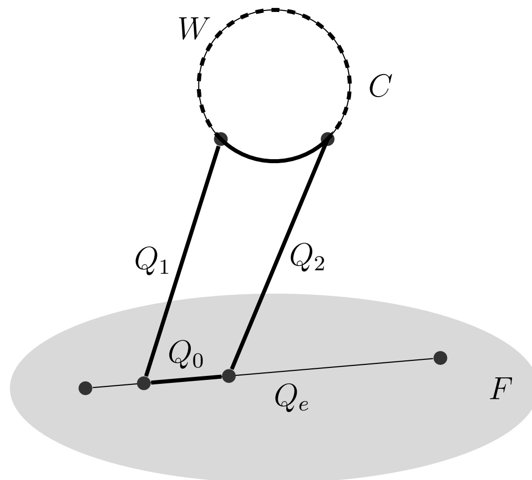

Figure 2: $(V(C), V(F))$-connector case. The subpath W of the cycle C is dashed, and the cycle D is displayed in bold.

counts vertices in $V(P_1) \cap V(C_0)$ and the upper bound counts at most $2\lceil \varepsilon n \rceil$ vertices in $(V(P_1) \cup V(P_2)) \cap V(C_0)$, at most $4/\varepsilon^2$ vertices on the paths in $\mathcal{P}'$ linking P_1 and P_2, and observing that the 4 endpoints of the linking paths are counted twice.

Let C be a longest cycle in H subject to $|V(C)| \leq 2\varepsilon n + 4/\varepsilon^2$, let $\ell = |V(C)|$, and note that $\ell \geq |V(C_0)| \geq (\varepsilon^2/2)n$. If $V(C)$ intersects each subgraph in $\mathcal{F}_0$, then $Y \cup V(C)$ witnesses $\tau_R(G) \leq |V(C)| + |Y| \leq (2\varepsilon n + 4/\varepsilon^2) + \varepsilon n = 3\varepsilon n + (n/m)^{1/2} < 8m^{5/4}n^{3/4}$. Otherwise, choose $F \in \mathcal{F}_0$ that is disjoint from C. We may assume $|V(F)| \geq \ell$, or else $Y \cup V(F)$ witnesses that $\tau_R(G) \leq |V(F)| + |Y| < (2\varepsilon n + 4/\varepsilon^2) + \varepsilon n < 8m^{5/4}n^{3/4}$.

If H has a $(V(C), V(F))$-separator T of size at most $\varepsilon\ell$, then we obtain an ε-partial transversal as follows. At most one component H' of $H - T$ contains graphs in $\mathcal{F}_0$. Let $X' = V(G) - V(H')$ and let $Y' = Y \cup T$. Since H' is disjoint from one of $\{C, F\}$, it follows that $|X'| - |X| \geq \ell$. We compute $|Y'| = |Y| + |T| \leq \varepsilon|X| + \varepsilon\ell \leq \varepsilon|X| + \varepsilon(|X'| - |X|) \leq \varepsilon|X'|$. Hence (H', X', Y') is an ε-partial transversal with $|V(H')| < |V(H)|$.

Otherwise, H has a $(V(C), V(F))$-connector $\mathcal{Q}$ with $|\mathcal{Q}| \geq \varepsilon\ell$. We use $\mathcal{Q}$ to obtain a contradiction. For $e \in E(R)$, let Q_e be the path in F corresponding to e, and let $\mathcal{Q}_e$ be the set of paths in $\mathcal{Q}$ which have an endpoint in Q_e. Since $|E(R)| = m$, it follows that $|\mathcal{Q}_e| \geq |\mathcal{Q}|/m \geq \varepsilon\ell/m$ for some edge $e \in E(R)$. Let $\mathcal{Q}'$ be the set of paths in $\mathcal{Q}_e$ of size at most $\frac{2mn}{\varepsilon\ell}$, and note that $|\mathcal{Q}'| \geq |\mathcal{Q}_e|/2 \geq \frac{\varepsilon\ell}{2m}$, or else $\mathcal{Q}_e$ has at least $\frac{\varepsilon\ell}{2m}$ paths of size more than $\frac{2mn}{\varepsilon\ell}$, a contradiction. The endpoints of paths in $\mathcal{Q}'$ divide Q_e into $|\mathcal{Q}'| - 1$ edge-disjoint subpaths. Choose $Q_1, Q_2 \in \mathcal{Q}'$ to minimize the length of such a subpath Q_0 of Q_e, and note that Q_0 has length at most $\frac{n-1}{|\mathcal{Q}'|-1}$; see Figure 2. Since $m \leq n$, we have $2m \leq 2m^{3/4}n^{1/4} = \frac{\varepsilon^3}{4}n \leq \frac{\varepsilon\ell}{2}$, and hence $\frac{n-1}{|\mathcal{Q}'|-1} < \frac{n}{\frac{\varepsilon\ell}{2m}-1} = \frac{2mn}{\varepsilon\ell - 2m} \leq \frac{4mn}{\varepsilon\ell}$.

The endpoints of Q_1 and Q_2 on C partition C into two subpaths; let W be the longer subpath. If $|E(W)| \geq |E(Q_0)|$, then we would obtain a larger R-subdivision by using Q_1, W, and Q_2 to bypass Q_0. Since F is a maximum R-subdivision, we have $|E(W)| < |E(Q_0)|$. Therefore using Q_1, Q_0, and Q_2 to bypass W gives a cycle D with $|E(D)| > |E(C)|$. By the extremal choice of C, it follows that $|V(D)| > 2\varepsilon n + 4/\varepsilon^2$. On the other hand, $|V(D)| =$

$|E(D)| \leq \frac{\ell}{2} + |E(Q_1)| + |E(Q_0)| + |E(Q_2)| \leq \frac{\ell}{2} + \frac{2mn}{\varepsilon\ell} + \frac{4mn}{\varepsilon\ell} + \frac{2mn}{\varepsilon\ell} = \frac{\ell}{2} + \frac{8mn}{\varepsilon\ell}$.

Therefore $2\varepsilon n + \frac{4}{\varepsilon^2} < |V(D)| \leq \frac{\ell}{2} + \frac{8mn}{\varepsilon\ell} \leq \varepsilon n + \frac{2}{\varepsilon^2} + \frac{8mn}{\varepsilon\ell} \leq \varepsilon n + \frac{2}{\varepsilon^2} + \frac{16m}{\varepsilon^3}$, where the last inequality uses $\ell \geq (\varepsilon^2/2)n$. Simplifying gives $\varepsilon n < \frac{16m}{\varepsilon^3} - \frac{2}{\varepsilon^2} < \frac{16m}{\varepsilon^3}$, and this inequality is violated when $\varepsilon \geq (16m/n)^{1/4}$. □

Applying Theorem 1, we obtain the following corollary.

Corollary 2. *Let G be an n-vertex graph. If G is connected, then* $\mathrm{lpt}(G) \leq 8n^{3/4}$*. If G is 2-connected, then* $\mathrm{lct}(G) \leq 20n^{3/4}$*.*

Proof. When $R = P_2$, an R-transversal is a longest path transversal. It is well known that if G is connected, then the longest paths pairwise intersect. By Theorem 1, we have $\mathrm{lpt}(G) = \tau_R(G) \leq 8n^{3/4}$.

Similarly, when $R = C_2$, an R-transversal is a longest cycle transversal. If G is 2-connected, then the longest cycles pairwise intersect. By Theorem 1, we have $\mathrm{lct}(G) = \tau_R(G) \leq 8 \cdot 2^{5/4} \cdot n^{3/4} \leq 20n^{3/4}$. □

We do not know whether the assumption in Theorem 1 that R is connected is necessary to obtain sublinear R-transversals. To obtain analogues of Corollary 2 for general R, we show that the maximum R-subdivisions pairwise intersect when the connectivity of G is sufficiently large. Recall that a graph G is *k-connected* if $|V(G)| > k$ and $G - S$ is connected for each $S \subseteq V(G)$ with $|S| < k$. Moreover, the *connectivity* of G, denoted $\kappa(G)$, is the maximum k such that G is k-connected.

Lemma 3. *Let R be a connected m-edge multigraph with $m \geq 1$. If $\kappa(G) > m^2$, then the maximum R-subdivisions in G are pairwise intersecting.*

Proof. Suppose for a contradiction that G has disjoint maximum R-subdivisions F_1 and F_2, and let $k = |V(F_1)| = |V(F_2)|$. By Menger's Theorem, there is an $(V(F_1), V(F_2))$-connector $\mathcal{P}$ with $|\mathcal{P}| = \min\{k, m^2 + 1\}$. If $|\mathcal{P}| = k$, then every vertex in F_1 is an endpoint of a path in $\mathcal{P}$, and we obtain an R-subdivision of size more than k by replacing an edge $uv \in E(F_1)$ with a path in $\mathcal{P}$ having u as an endpoint, a path in $\mathcal{P}$ having v as an endpoint, and an appropriate path in the connected subgraph F_2.

So we may assume $|\mathcal{P}| = m^2 + 1$. For each $e \in E(R)$, let $F_i(e)$ be the path in F_i corresponding to e. Since R has no isolated vertices, we may associate each $P \in \mathcal{P}$ with an ordered pair of edges $(e_1, e_2) \in (E(R))^2$ such that P has its endpoint in F_1 in $F_1(e_1)$ and its endpoint in F_2 in $F_2(e_2)$. Since $|\mathcal{P}| > m^2$, some pair (e_1, e_2) is associated with distinct paths $P, Q \in \mathcal{P}$. Let W_i be the subpath of $F_i(e_i)$ whose endpoints are in $V(P) \cup V(Q)$. If $|E(W_1)| \geq |E(W_2)|$, then we modify F_2 to obtain a larger R-subdivision by using P, W_1, and Q to bypass W_2. Similarly, if $|E(W_2)| \geq |E(W_1)|$, then we modify F_1 to obtain a larger R-subdivision by using P, W_2, and Q to bypass W_1. □

Corollary 4. *Let R be a connected m-edge multigraph. If G is an n-vertex graph with $\kappa(G) > m^2$, then $\tau_R(G) \leq 8m^{5/4}n^{3/4}$.*

As it is not known whether there exists a connected graph G with $\mathrm{lpt}(G) > 3$, reducing the gap between our sublinear upper bound on $\mathrm{lpt}(G)$ and the constant lower bound remains a major open problem in the area of longest path transversals.

Acknowledgement

The authors greatly appreciate the careful comments of an anonymous referee.

References

[1] R. Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, 2005.

[2] T. Gallai. Problem 4. In P. Erdős and G. Katona, editors, *Theory of Graphs, Proceedings of the Colloquium Held at Tihany, Hungary, September 1966*, page 362. Academic Press, New York, 1968.

[3] B. Grünbaum. Vertices missed by longest paths or circuits. *Journal of Combinatorial Theory, Series A*, 17(1):31–38, 1974.

[4] D. Rautenbach and J.-S. Sereni. Transversals of longest paths and cycles. *SIAM Journal on Discrete Mathematics*, 28(1):335–341, 2014.

[5] C. Thomassen. Hypohamiltonian graphs and digraphs. In Y. Alavi and D.R. Lick, editors, *Theory and Applications of Graphs*, pages 557–571. Springer Berlin Heidelberg, 1978.

[6] H. Walther. Über die nichtexistenz eines knotenpunktes, durch den alle längsten wege eines graphen gehen. *Journal of Combinatorial Theory*, 6(1):1–6, 1969.

[7] H. Walther and H.-J. Voss. *Über Kreise in Graphen*. Deutscher Verlag der Wissenschaften, 1974.

[8] D. B. West. *Introduction to Graph Theory*. Prentice Hall, 2nd edition, 2001.

[9] T. Zamfirescu. A two-connected planar graph without concurrent longest paths. *Journal of Combinatorial Theory, Series B*, 13(2):116–121, 1972.

[10] T. Zamfirescu. On longest paths and circuits in graphs. *Mathematica Scandinavica*, 38: 211–239, 1976.