A counterexample to the coarse Menger conjecture

Tung Nguyen[1]
Princeton University,
Princeton, NJ 08544, USA

Alex Scott[2]
University of Oxford,
Oxford, UK

Paul Seymour[3]
Princeton University,
Princeton, NJ 08544, USA

November 19, 2023; revised January 12, 2024

[1]Supported by AFOSR grants A9550-19-1-0187 and FA9550-22-1-0234, and by NSF grants DMS-1800053 and DMS-2154169.

[2]Supported by EPSRC grant EP/X013642/1

[3]Supported by AFOSR grants A9550-19-1-0187 and FA9550-22-1-0234, and by NSF grants DMS-1800053 and DMS-2154169.

Abstract

It was conjectured, independently by two sets of authors, that for all integers $k, d \geq 1$ there exists $\ell > 0$, such that if S, T are subsets of vertices of a graph G, then either there are k paths between S, T, pairwise at distance at least d, or there is a set $X \subseteq V(G)$ with $|X| \leq k - 1$ such that every path between S, T contains a vertex with distance at most ℓ from some member of X. The result is known for $k \leq 2$, but we will show that it is false for all $k \geq 3$, even if G is constrained to have maximum degree at most three. We also give a proof of the result when $k = 2$ that is simpler than the previous proofs.

1 Introduction

If X is a subset of the vertex set of a graph G, or a subgraph of G, and the same for Y, then $d(X,Y)$ denotes the distance between X,Y, that is, the number of edges in the shortest path of G with one end in X and the other in Y. We are interested in when there are k paths in G between two given sets that are pairwise at distance at least d. The following conjecture was proposed by Albrechtsen, Huynh, Jacobs, Knappe, and Wollan [1], and independently by Georgakopoulos and Papasoglu [3]:

1.1 False conjecture: *For all integers $k,d \geq 1$ there exists $\ell > 0$ with the following property. Let G be a graph and let $S,T \subseteq V(G)$; then either*

- *there are k paths between S,T, pairwise at distance at least d; or*
- *there is a set $X \subseteq V(G)$ with $|X| \leq k-1$ such that every path between S,T contains a vertex with distance at most ℓ from some member of X.*

Both sets of authors proved the result for $k=2$, but the $k \geq 3$ case remained open. Here we give a counterexample with $d=k=3$, and with maximum degree at most three. We also give a simpler proof of the $k=2$ case.

The case $k=3$ was of special interest, because it is easy to see that if the result is true when $d=3$ then it is true for general k (apply the result when $d=3$ to the dth power of G).

The case $d=2$ and general k is still open. Another related conjecture, still open, is 1.1 with the bound on $|X|$ relaxed:

1.2 Conjecture: *For all k,d there exist ℓ,n such that, with G,S,T as before, either the first bullet of 1.1 holds, or there is a set $X \subseteq V(G)$ with $|X| \leq n$ such that every path between S,T contains a vertex with distance at most ℓ from some member of X.*

When $d=2$, this last conjecture was shown to be true for graphs of bounded maximum degree by Hendrey, Norin, Steiner and Turcotte [4], and independently by Gartland, Koorhonen and Lokshtanov [2].

Agelos Georgakopoulos brought to our attention the following variant of 1.1, which also remains open:

1.3 Conjecture: *For all integers $k,d \geq 1$ there exists $\ell > 0$ with the following property. Let G be a graph and let $S,T \subseteq V(G)$; then either*

- *there are k paths between S,T, pairwise at distance at least d; or*
- *there is a vertex x of G such that there do not exist $k-1$ paths between S and T, pairwise at distance at least ℓ and each with distance at least ℓ from x.*

2 The counterexample

In this section, for each value of $\ell > 0$, we give an example of a graph G, and two subsets S,T of $V(G)$, such that there do not exist three paths between S,T, pairwise at distance at least 3, and for every $X \subseteq V(G)$ with $|X| \leq 2$, there is a path P between S,T such that its distance from X is more

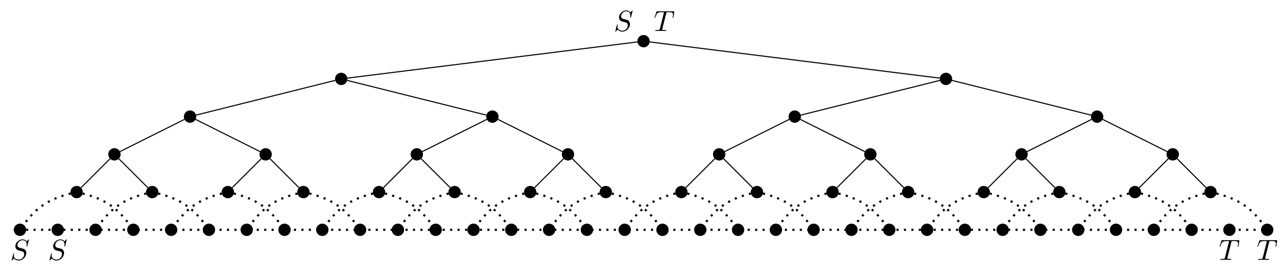

Figure 1: The dotted curves represent long paths.

than ℓ. It is illustrated in figure 1. The two sets S, T both have size three, and there is a vertex in $S \cap T$.

To show that this is a counterexample, we need to check that for every $X \subseteq V(G)$ with $|X| \leq 2$, there is a path P between S, T such that its distance from X is more than ℓ; and that there do not exist three paths between S, T, pairwise at distance at least 3. The first is easy, so let us do it now.

2.1 *Let $\ell \geq 1$ be an integer, and let G be as in figure 1, where the binary tree has depth more than $2\ell + 2$, and each of the dotted curves represents a path of length more than 2ℓ. Let S, T be as shown in the figure. If $X \subseteq V(G)$ with $|X| \leq 2$, then there is a path P between S, T such that $d(X, P) > \ell$.*

Proof. Let r be the root of the binary tree (the vertex at the top of the figure). Since $r \in S \cap T$, we may assume there exists $x_1 \in X$ such that its distance from r is at most ℓ; and since all of the paths represented by the dotted curves (let us call them "dotted paths") have distance at least $2\ell + 1$ from r (because the binary tree has depth more than $2\ell + 2$), they all have distance more than ℓ from x_1. Since the path (M say) at the bottom of the figure is between S, T, we may assume that there exists $x_2 \in X$ with distance at most ℓ from M. The set of vertices of G with distance at most ℓ from x_2 is either a subset of the vertex set of one of the dotted paths in M, or it contains exactly one end v of a dotted path of M, and consists of v together with subsets of the interiors of the (at most three) dotted paths incident with v. In either case, there is an end v of one of the dotted paths in M, such that the set of vertices of G with distance at most ℓ from x_2 is a subset of the set consisting of v together with the interiors of the dotted paths incident with v. But for every choice of v, there is a path between S, T made by a union of dotted paths, none of them incident with v; and consequently this path has distance more than ℓ from X. This proves 2.1. ∎

To prove the second statement, that there do not exist three paths between S, T, pairwise at distance at least d, we will prove something stronger, by induction on the depth of the binary tree; and for that we need to allow the binary tree to have small depth, and we need to set up some notation, so let us define the graph more carefully.

Let $k \geq 2$ be an integer. Take a uniform binary tree B with depth k. Thus, B has $2^k - 1$ vertices, and a root, and every path from the root to one of the leaves has exactly $k - 1$ edges. Now add two more vertices, and let Z consist of the set of leaves of B together with the two new vertices; and add a path M with vertex set Z, as shown in figure 2. We call the resulting graph G_k. (To get the graph of the counterexample, we need to replace the edges of $G_{2\ell+3}$ incident with vertices in Z by long paths, but let us not do that yet.) For each vertex v of B that is not a leaf, there is a copy of some G_h with root v, formed by v and its descendents in B, and two extra vertices; we will apply the inductive hypothesis to these smaller graphs. Let M have ends s_1, t_2, and let s_2, t_1 be the neighbours in M of s_1, t_2 respectively.

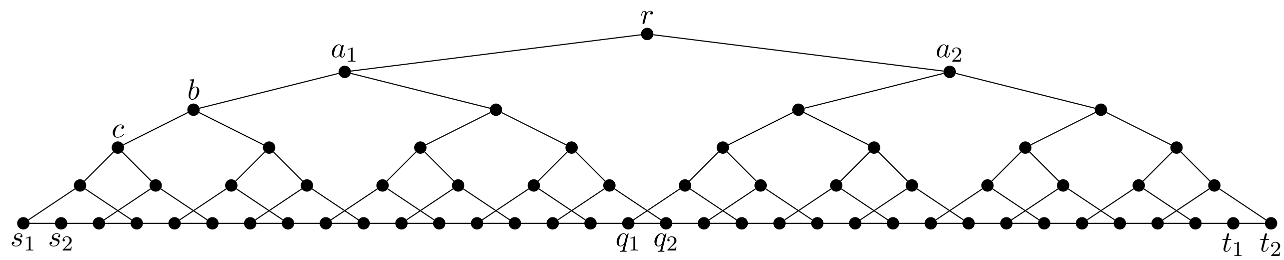

Figure 2: The graph G_k. (In this figure, $k = 6$.) The vertex b is some vertex on the path between a_1 and s_1; it need not be adjacent to a_1, and it might equal a_1. Its child on that path is c.

We will show:

2.2 *For each integer $k \geq 2$, if P, Q are vertex-disjoint paths of G_k between $\{s_1, s_2\}$ and $\{t_1, t_2\}$, then either*

- *there is a path of length at most two between $V(P)$ and $V(Q)$ such that none of its edges have ends in Z, or*
- *one of P, Q is the path of B between s_1, t_2.*

Proof. We proceed by induction on k. If $k = 1$ the result is clear, so we assume that $k \geq 2$ and the result holds for all smaller values of k. Suppose that P, Q are vertex-disjoint paths of G_k between $\{s_1, s_2\}$ and $\{t_1, t_2\}$, and the first bullet of the theorem is false. Let r be the root of B, let a_1 be the neighbour of r in the path of B between r, s_1, and define a_2 similarly. Suppose first that neither of P, Q contains r. Then we may assume that P contains q_1 and Q contains q_2, where q_1, q_2 are as shown in the figure; and so from the inductive hypothesis applied to the two copies of G_{k-1} with roots a_1 and a_2, P contains the path of B between s_1, q_2, and Q contains the path of B between q_1, t_2. So $a - 1 \in V(P)$ and $a_2 \in V(Q)$; but r is adjacent to both a_1, a_2, a contradiction.

So r belongs to one of P, Q, say to P. We need to show that P contains the two subpaths of B between r, s_1 and between r, t_2. Suppose it does not contain the first, say; and let us call this path A_1. Choose a maximal common subpath of P, A_1, with ends a_1, b say. (Possibly $b = a_1$.) We see that $b \neq s_1$, since $A \not\subseteq P$; and b is not adjacent to s_1, since one of P, Q must use the edge of B incident with s_1. Thus the subpath of A between b, s_1 has length h say, where $h \geq 2$. Let c be the child of b in $V(A)$. It follows that neither of P, Q contains c; and this contradicts the inductive hypothesis, applied to the copy of G_{h-1} with root c. This proves 2.2. ∎

From 2.2, it follows easily that the graph of figure 1, with S, T as shown, has the property that there do not exist three paths between S, T, pairwise at distance at least 3. Because if P, Q, R are three such paths, one of them contains the root r, and so consists just of the vertex r; and so the other two cannot use r, and so contradict 2.2.

3 A simpler proof

In this section we give a proof of 1.1 when $k = 2$, that is simpler than the proofs previously announced. We need some lemmas about intervals. An *interval* is a pair (a, b) of integers with $a \le b$, and its *length* is $b - a$. If $\mathcal{H}$ is a set of intervals that can be written $\mathcal{H} = \{(a_i, b_i) : 1 \le i \le t\}$ for some $t \ge 1$, such that $0 \le a_1 < a_2 < \cdots < a_t \le n$, and $0 \le b_1 < \cdots < b_t \le n$, we call this the *standard form* for $\mathcal{H}$. Thus $\mathcal{H}$ has a standard form if and only there do not exist distinct $(a, b), (c, d) \in \mathcal{H}$ with $a \le c \le d \le b$. (See figure 3.)

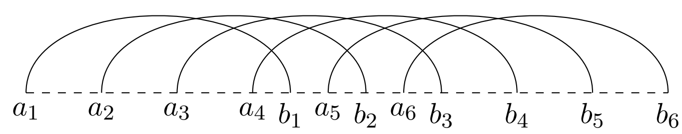

Figure 3: Intervals in standard form (the curves join the ends of the intervals).

If $(a, b), (c, d)$ are intervals, we say (a, b) *captures* (c, d) if $a \le c \le d \le b$; and a set $\mathcal{H}$ of intervals *captures* (c, d) if there exists $(a, b) \in \mathcal{H}$ that captures (c, d). For integers ℓ, n with $0 < \ell \le n$, we say a set $\mathcal{H}$ of intervals is *ℓ-powerful* in $(0, n)$ if

- $0 \le a \le b \le n$ for all $(a, b) \in \mathcal{H}$; and
- $\mathcal{H}$ captures every interval of length ℓ captured by $(0, n)$.

Our objective is, starting with an ℓ-powerful set $\mathcal{H}$ of intervals, to find a subset $\mathcal{H}'$ of $\mathcal{H}$ which is still ℓ'-powerful for some ℓ' not much smaller than ℓ, such that all the ends of all the intervals in $\mathcal{H}'$ are far apart. (Not exactly: in standard form, we cannot arrange that $a_{i+2} - b_i$ is large for each i, and it might even be negative, and we cannot arrange that $a_2 - a_1$ and $b_t - b_{t-1}$ are large, but we will make all the other differences large). We begin with:

3.1 *If $0 < \ell \le n$ and $\mathcal{H}$ is a set of intervals that is ℓ-powerful in $(0, n)$, and minimal with this property, then in standard form $a_j \ge b_i - \ell + 2$ for all $i, j \in \{1, \ldots, t\}$ with $j \ge i + 2$.*

Proof. (See figure 4.)

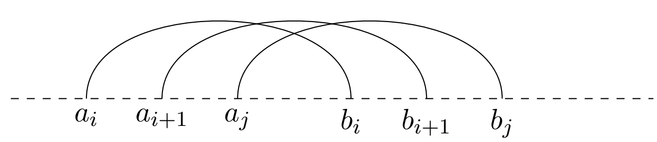

Figure 4: For the proof of 3.1.

Let $1 \le i, j \le t$ with $j \ge i + 2$. From the minimality, there exists $h \in \{0, n - \ell\}$ such that (a_{i+1}, b_{i+1}) captures $(h, h + \ell)$, and no other member of $\mathcal{H}$ captures $(h, h + \ell)$. Since (a_i, b_i) does not capture $(h, h + \ell)$, and $a_i < a_{i+1} \le h$, it follows that $b_i < h + \ell$; and since (a_j, b_j) does not capture $(h, h + \ell)$, and $b_j \ge b_{i+1} \ge h + \ell$, it follows that $a_j > h$. Hence $a_j \ge b_i - \ell + 2$. This proves 3.1. ∎

If $\mathcal{H}$ can be written in standard form, and $\mathcal{H}' \subseteq \mathcal{H}$, then $\mathcal{H}'$ can also be written in standard form. Some inequalities about the corresponding numbers a_i, b_i are *hereditary*; that is, if they hold for $\mathcal{H}$ then they also hold for all its subsets $\mathcal{H}'$. For instance, the inequalities of 3.1 are hereditary. Let $\mathcal{H}' \subseteq \mathcal{H}$, and $\mathcal{H} = \{(a_i, b_i) : 1 \le i \le t\}$ and $\mathcal{H}' = \{(a'_i, b'_i) : 1 \le i \le t'\}$, in standard form. If we know that $a_j \ge b_i - \ell + 2$ for all $i, j \in \{1, \dots, t\}$ with $j \ge i + 2$, then it follows that $a'_j \ge b'_i - \ell + 2$ for all $i, j \in \{1, \dots, t'\}$ with $j \ge i + 2$. Most of the other inequalities we will prove are hereditary, and we leave verifying this to the reader.

3.2 *If $0 < 2\ell \le n$ and $\mathcal{H}$ is a set of intervals that is 2ℓ-powerful in $(0, n)$, then there exists $\mathcal{H}' \subseteq \mathcal{H}$, minimally ℓ-powerful in $(0, n)$, that can be written $\mathcal{H}' = \{(a_i, b_i) : 1 \le i \le t\}$ in standard form, such that*

- *$a_j \ge b_i - \ell + 2$ for $1 \le i, j \le t$ with $j \ge i + 2$; and*
- *$b_i - b_{i-1} \ge \ell$ for $1 < i < t$, and $b_{i-1} - a_i \ge \ell$ for $1 < i \le t$.*

Proof. Choose $\mathcal{H}_1 \subseteq \mathcal{H}$, minimal such that it is 2ℓ-powerful in $(0, n)$. Let us choose $\mathcal{H}_2 \subseteq \mathcal{H}_1$, such that

- $\mathcal{H}_2$ captures the intervals $(0, 2\ell)$ and $(n - 2\ell, n)$;
- if $(a, b) \in \mathcal{H}_2$ and $b + \ell < n$ then some member of $\mathcal{H}_2$ captures the interval $(b - \ell, b + \ell)$;
- $\mathcal{H}_2$ is minimal with these properties.

It follows that we can write $\mathcal{H}_2$ in standard form, $\mathcal{H}_2 = \{(a_i, b_i) : 1 \le i \le t\}$ for some $t \ge 1$; and $a_1 = 0$ and $b_t = n$; and the first bullet of the theorem holds, by 3.1. We claim that:

(1) *For $1 < i < t$, $b_{i-1} \le n - \ell$ and (a_i, b_i) captures $(b_{i-1} - \ell, b_{i-1} + \ell)$, and there is no $j \in \{1, \dots, t\}$ with $j \ne i$ such that (a_j, b_j) captures $(b_{i-1} - \ell, b_{i-1} + \ell)$.*

We prove this by induction on i, so we may assume that either $i = 2$, or the claim holds for $i - 1$. Since $a_i > a_1 \ge 0$ and $b_i < b_t = n$, (a_i, b_i) captures neither of $(0, 2\ell), (n - 2\ell, n)$; and so from the minimality of $\mathcal{H}_2$, there exists $h \in \{1, \dots, t\}$ such that $b_h \le n - \ell$ and (a_i, b_i) captures $(b_h - \ell, b_h + \ell)$, and there is no $j \in \{1, \dots, t\}$ with $j \ne i$ such that (a_j, b_j) captures $(b_h - \ell, b_h + \ell)$. It follows that $h < i$, since $b_h < b_i$; and $h > i - 2$, since if $h \le i - 2$ then from the inductive hypothesis, there is no $j \ne h + 1$ such that (a_j, b_j) captures $(b_h - \ell, b_h + \ell)$. Hence $h = i - 1$. This completes the induction, and so proves (1).

Consequently $b_i - b_{i-1} \ge \ell$ and $b_{i-1} - a_i \ge \ell$ for $1 < i < t$. For the second bullet we still need to check that $b_{t-1} - a_t \ge \ell$. This is true if $b_{t-1} \ge n - \ell$, since (a_t, b_t) captures $(n - 2\ell, n)$. If $b_{t-1} < n - \ell$, then some (a_j, b_j) captures $(b_{t-1} - \ell, b_{t-1} + \ell)$; and $j = t$ since $b_j > b_{t-1}$. This proves that $\mathcal{H}_2$ satisfies the second bullet.

We claim that $\mathcal{H}_2$ is ℓ-powerful in $(0, n)$. To see this, let $0 \le h \le n - \ell$; we must show that some member of $\mathcal{H}_2$ captures $(h, h + \ell)$. We may assume that $h < n - 2\ell$, since some member of $\mathcal{H}_2$ captures $(n - 2\ell, n)$. Choose $i \in \{1, \dots, t\}$ maximal such that $a_i \le h$. We may assume that $b_i < h + \ell$, since otherwise (a_i, b_i) captures $(h, h + \ell)$. Hence $b_i < n - \ell$, and so (a_{i+1}, b_{i+1}) captures

$(b_i - \ell, b_i + \ell)$, by (1). From the choice of i, $a_{i+1} > h$; but $a_{i+1} \le b_i - \ell$, so $b_i \ge a_{i+1} + \ell > h + \ell$, a contradiction. This proves that $\mathcal{H}_2$ is ℓ-powerful in $(0, n)$. Choose $\mathcal{H}' \subseteq \mathcal{H}_2$, minimal such that it is ℓ-powerful in $(0, n)$. Then by 3.1, the first bullet of the theorem holds; and the second still holds, since it is hereditary. This proves 3.2. ∎

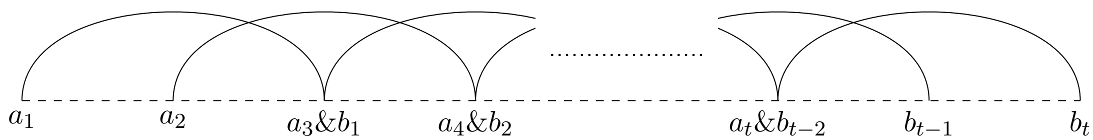

Figure 5: The output of 3.3. a_3, b_1 have been drawn in the same place because we do not know which is larger.

3.3 *If $0 < 4\ell \le n$ and $\mathcal{H}$ is a set of intervals that is 4ℓ-powerful in $(0, n)$, then there exists $\mathcal{H}' \subseteq \mathcal{H}$, ℓ-powerful in $(0, n)$, which can be written $\mathcal{H}' = \{(a_i, b_i) : 1 \le i \le t\}$ in standard form, such that the order of the numbers $a_1, \dots, a_t$ and $b_1, \dots, b_t$ is:*

$$0 = a_1 < a_2 < \min(a_3, b_1) < \max(a_3, b_1) < \min(a_4, b_2) < \max(a_4, b_2) < \cdots$$
$$\cdots < \min(a_t, b_{t-2}) < \max(a_t, b_{t-2}) < b_{t-1} < b_t = n.$$

(See figure 5.) Every two of $a_1, \dots, a_t, b_1, \dots, b_t$ differ by at least ℓ, except possibly the pairs (a_1, a_2), (b_{t-1}, b_t), and (a_i, b_{i-2}) for $3 \le i \le t$.

Proof. From 3.2 with ℓ replaced by 2ℓ, there exists $\mathcal{H}_1 \subseteq \mathcal{H}$, 2ℓ-powerful in $(0, n)$, that satisfies the inequalities of 3.2. By 3.2 again, applied to $\mathcal{H}_1$, with the order of $\{0, \dots, n\}$ reversed, there exists $\mathcal{H}_2 \subseteq \mathcal{H}_1$, minimally ℓ-powerful in $(0, n)$, such that in standard form,

- $b_i - b_{i-1} \ge 2\ell$ for $1 < i < t$, and $b_{i-1} - a_i \ge 2\ell$ for $1 < i \le t$ (since these inequalities are true for $\mathcal{H}_1$ and are hereditary);
- $a_{i+1} - a_i \ge \ell$ for $1 < i < t$; and
- $a_j \ge b_i - \ell + 2$ for all $i, j \in \{1, \dots, t\}$ with $j \ge i + 2$.

It follows that $b_j - b_i \ge \ell$ for all $i, j \in \{1, \dots, t\}$ with $j > i$ and $(i, j) \ne (t - 1, t)$; that $a_j - a_i \ge \ell$ for all $i, j \in \{1, \dots, t\}$ with $j > i$ and $(i, j) \ne (1, 2)$. Now let $i, j \in \{1, \dots, t\}$; we want to show that $|a_j - b_i| \ge \ell$, unless $i = j - 2$. If $i > j$, then $b_i \ge b_j + \ell \ge a_j + \ell$ as required. If $j = i$, the minimality of $\mathcal{H}_2$ implies that $b_i - a_i \ge \ell$, as required. If $i = j - 1$, we have $b_i - a_j \ge 2\ell$ from the first bullet above, as required. Now suppose that $i \le j - 3$. Hence $b_{j-2} \le a_j + \ell - 2$, and $b_{j-2} \ge b_{j-3} + 2\ell$; so $a_j + \ell - 2 \ge b_{j-3} + 2\ell$, that is, $a_j \ge b_{j-3} + \ell + 2 \ge b_i + \ell$ as required. This proves 3.3. ∎

Now let us deduce 1.1 when $k = 2$. As we said earlier, it suffices to handle the case when $d = 3$, so we will prove:

3.4 *Let G be a non-null graph and let $S, T \subseteq V(G)$; then either*

- *there are two paths between S, T with distance at least 3; or*
- *there exists $x \in V(G)$ such that every path between S, T contains a vertex with distance at most 161 from x.*

Proof. We assume the second outcome is false. Let $c \geq 7$ and $\ell \geq 2c + 5$ be integers: we will prove that the theorem holds with 161 replaced by $8\ell+c+2$. We may assume that G is connected, and there is a path between S, T; let R be such a path with minimum length, and let its vertices be $r_1, \ldots, r_{n-1}$ in order, where $r_1 \in S$ and $r_{n-1} \in T$. There is a path P between S, T with $d(r_1, P) > 8\ell + c + 2$, and we can assume that $d(P, R) \leq 2$, so $n - 2 \geq 8\ell + c + 1$. In particular, $S \cap T = \emptyset$ and there are no edges between S, T.

Let W be the set of all vertices with distance at most c from R. We call the set of all vertices v with $d(v, R) = c$ the *surface*. For each $v \in V(G)$, if $v \in S$ and $d(v, R) \geq c + 4$, let $a(v) = 0$, and otherwise let $a(v)$ be the smallest $i \in \{1, \ldots, n - 1\}$ such that there is a path between v and r_i of length $d(v, R)$. Similarly, if $v \in T$ and $d(v, R) \geq c + 4$ let $b(v) = n$, and otherwise let $b(v)$ be the largest $i \in \{1, \ldots, n - 1\}$ such that there is a path between v and r_i of length $d(v, R)$. For each $w \in W$, it follows from the choice of R that $0 \leq b(w) - a(w) \leq 2d(w, R) \leq 2c$.

Let $\mathcal{C}$ be the set of components of $G \setminus W$, and for each $C \in \mathcal{C}$, let $N(C)$ be the set of vertices in W with a neighbour in $V(C)$. Thus $N(C) \neq \emptyset$, and $N(C)$ is a subset of the surface. Let $a(C)$ be the smallest value of $a(v)$ for $v \in V(C)$, and define $b(C)$ similarly. Thus $0 \leq a(C) \leq b(C) \leq n$ for each $C \in \mathcal{C}$.

(1) *The set of intervals $\{(a(C), b(C)) : C \in \mathcal{C}\}$ is 16ℓ-powerful in $(0, n)$.*

Let $8\ell \leq m \leq n - 8\ell$; we must show that there exists $C \in \mathcal{C}$ such that $(a(C), b(C))$ captures $(m - 8\ell, m + 8\ell)$. Let B be the set of all vertices of G with distance at most $8\ell + c + 2$ from r_m. Let W_1 be the set of all $w \in W \setminus B$ with $a(w) \leq m$, and let W_2 be the set of all $w \in W \setminus B$ with $b(w) \geq m$. Thus $W_1 \cup W_2 = W \setminus B$. If $w \in W_1$, then since $d(w, r_m) > 8\ell + c + 2$ and $d(w, r_{a(w)}) \leq c$, it follows that $d(r_{a(w)}, r_m) > 8\ell+2$, and so $a(w) \leq m - 8\ell - 3$; and similarly if $w \in W_2$ then $b(w) \geq m + 8\ell + 3$. Thus, if $w_1 \in W_1$ and $w_2 \in W_2$, then

$$b(w_2) - a(w_1) \geq (m + 8\ell + 3) - (m - 8\ell - 3) = 16\ell + 6.$$

There is a path between w_1 and S of length at most $a(w_1) + c - 1$, and a path between w_2 and T of length at most $|E(R)| - b(w_2) + c - 1$, and, since there is no path between S, T of length less than $|E(R)|$, it follows that

$$\begin{aligned} d(w_1, w_2) &\geq |E(R)| - (a(w_1) + c - 1) - (|E(R)| - b(w_2) + c - 1) \\ &= b(w_2) - a(w_1) - 2(c - 1) \geq (16\ell + 6) - 2(c - 1) \geq 2. \end{aligned}$$

Consequently $W_1 \setminus B$ and $W_2 \setminus B$ are disjoint, and there are no edges between them. Moreover, if $w_1 \in S$ and $w_2 \in W_2$ then $b(w_2) \geq m + 8\ell + 3 \geq 16\ell + 3$, and similarly

$$d(w_1, w_2) \geq |E(R)| - (|E(R)| - b(w_2) + c - 1) = b(w_2) - (c - 1) \geq (16\ell + 3) - 2(c - 1) \geq 2.$$

So $S \cap W_2 = \emptyset$ and there are no edges between them, and the same for W_1, T. Thus $W_1 \cup S$ is disjoint from $W_2 \cup T$ and there are no edges between them.

Since the second outcome of the theorem is false, there is an $S-T$ path P in G disjoint from B. Consequently, there is a subpath P' of P with first vertex in $W_1 \cup S$, last vertex in $W_2 \cup T$, and with no other vertex in $W \cup B$. Since $W_1 \cup S$ is disjoint from $W_2 \cup T$ and there are no edges between them, P' has an internal vertex, which is therefore not in W. But the only vertices of P' in W are ends of P', and so there exists $C \in \mathcal{C}$ such that $V(P') \subseteq V(C) \cup W$. We will show that $a(C) \leq m - 8\ell$ and $b(C) \geq m + 8\ell$. Let $w_1 \in W_1 \cup S$ and $w_2 \in W_2 \cup T$ be the ends of P'. Either $w_1 \in W_1$, or $w_1 \in S \setminus W$ and $d(w_1, R) \geq c + 4$, or $w_1 \in S \setminus W$ and $d(w_1, R) \leq c + 3$, and we claim that $a(w_1) \leq m - 8\ell$ in each case. In the first case, w_1 has a neighbour $v \in V(C)$; and since $d(v, R) = c+1 = d(w_1, R)+1$, it follows that $a(C) \leq a(v) \leq a(w_1) \leq m-8\ell-3 \leq m-8\ell$ as required. In the other two cases, $w_1 \in V(C)$, and so $a(C) \leq a(w_1)$. In the second case, $a(w_1) = 0 \leq m - 8\ell$ as required. In the third case, when $w_1 \in S$ and $d(w_1, R) \leq c + 3$, it follows that $d(w_1, r_{a(w_1)}) \leq c + 3$, and since $d(w_1, r_m) > 8\ell + c + 2$, this implies that $a(w_1) \leq m - 8\ell$ as required. This proves that $a(C) \leq m - 8\ell$ in all three cases. Similarly $b(C) \geq m + 8\ell$. This proves (1).

From 3.3, with ℓ replaced by 4ℓ, there exists $\mathcal{D} \subseteq \mathcal{C}$ such that the set of intervals $\{(a(C), b(C)) : C \in \mathcal{D}\}$ is 4ℓ-powerful in $(0, n)$, and the members of $\mathcal{D}$ can be numbered $D_1, \dots, D_t$ such that:

- $0 = a(D_1) < a(D_2) < \cdots < a(D_t) \leq n$, and $0 \leq b(D_1) < b(D_2) < \cdots < b(D_t) = n$; and
- $a(D_i) - b(D_{i-3}) \geq 4\ell$ for $4 \leq i \leq t$.

(We are not using the full strength of 3.3 here; that will come later.)

(2) *For $1 \leq i, j \leq t$, if $j \geq i + 3$ then $d(D_i, D_j) \geq 4\ell - 2c + 2$.*

Suppose not. Hence there exist $w \in N(D_i)$ and $w' \in N(D_j)$ such that $d(w, w') < 4\ell - 2c$; and there is a path between $w, V(R)$ of length c, and its end in R is some r_k where $k \leq b(D_i)$. Similarly there is a path between $w', V(R)$ of length c, and its end in R is some $r_{k'}$ where $k \geq a(D_j)$. Thus there exist $k \leq b(D_i)$ and $k' \geq a(D_j)$ such that there is a path between $r_k, r_{k'}$ of length less than 4ℓ. Since R is a shortest path, it follows that $4\ell - 1 \geq k' - k \geq a(D_j) - b(D_i) \geq 4\ell$, a contradiction. This proves (2).

Let Δ be the union of the vertex sets of the members of $\mathcal{D}$. If $X \subseteq V(G)$, let $\Delta(X)$ be the set of $D \in \mathcal{D}$ with $X \cap N(D) \neq \emptyset$. Let us say a *joint* is a subset $X \subseteq V(G)$, inducing a connected subgraph, such that

- $|\Delta(X)| \geq 2$;
- if $|\Delta(X)| = 2$ then $|X| \leq 3$;
- if $|\Delta(X)| \leq 3$ then $|X| \leq 8$;
- $|X| \leq 29$; and
- every vertex of $X \cap W$ belongs to a path of $G[X]$ with both ends in the surface.

(3) *For every joint X, $|\Delta(X)| \leq 3$, and hence $|X| \leq 8$; and $d(X, R) \geq c - (|X| - 1)/2$.*

Suppose that $|\Delta(X)| \geq 4$; then we can choose i, j with $1 \leq i, j \leq t$ and $j \geq i + 3$ such that $D_i, D_j \in \Delta(X)$. Since $G[X]$ is connected, there is a path between D_i, D_j of length at most $|X| + 1 \leq 30$, contrary to (2), since $4\ell - 2c + 2 > 30$. Thus $|\Delta(X)| \leq 3$. Moreover, if $v \in X \cap W$, then since v belongs to a path of $G[X]$ with both ends in the surface, there is a path from v to the surface of length at most $(|X| - 1)/2$; and since every path from the surface to R has length at least c, it follows that $d(v, R) \geq c - (|X| - 1)/2$. This proves (3).

Let Z be the union of all the joints. Each component of $G[Z \cup \Delta]$ is called a *supercomponent*. Each supercomponent includes at least one member of $\mathcal{D}$, but may be composed of many joints and members of $\mathcal{D}$.

(4) *Let F_1, F_2 be distinct supercomponents, and let Q be a path with ends $f_1 \in V(F_1)$ and $f_2 \in V(F_2)$. If both f_1, f_2 belong to joints then $|Q| \geq 16$; if exactly one of f_1, f_2 belongs to a joint X then $|Q| \geq 8$, and $|Q| \geq 24$ if $|\Delta(X)| \geq 3$; and if neither of f_1, f_2 belong to joints then $|Q| \geq 6$. In any case, $d(F_1, F_2) \geq 5$.*

Let Q^* be the set of vertices in the interior of Q. Suppose first that f_i belongs to a joint X_i for $i = 1, 2$. Then $X_1 \cup X_2 \cup Q^*$ is not a joint (since F_1, F_2 are distinct supercomponents). But $|\Delta(X_1 \cup X_2 \cup Q^*)| \geq 4$ (since $|\Delta(X_i)| \geq 2$ for $i = 1, 2$, and $\Delta(X_1) \cap \Delta(X_2) = \emptyset$), and every vertex of $(X_1 \cup X_2 \cup Q^*) \cap W$ belongs to a path of $G[X_1 \cup X_2 \cup Q^*]$ with both ends in the surface; and so $|X_1 \cup X_2 \cup Q^*| \geq 30$. Since $|X_1|, |X_2| \leq 8$ by (3), it follows that $|Q^*| \geq 14$, and so $|Q| \geq 16$ as claimed. Thus we may assume that $f_2 \in D'$ for some $D' \in \mathcal{D}$ included in F_2, and so $V(Q) \cap N(D') \neq \emptyset$, and hence $D' \in \Delta(Q^*)$. Next, suppose that $f_1 \in X_1$ for some joint X_1. Then $|\Delta(X_1 \cup Q^*)| \geq 3$, and since $X_1 \cup Q^*$ is not a joint (because F_1, F_2 are distinct supercomponents), it follows that either $|\Delta(X_1 \cup Q^*)| = 3$ and $|X_1 \cup Q^*| \geq 9$, or $|X_1 \cup Q^*| \geq 30$. In the first case, $|\Delta(X_1)| = 2$, so $|X_1| \leq 3$, and therefore $|Q^*| \geq 6$ and $|Q| \geq 8$; and in the second case, $|X_1| \leq 8$ by (3), and so $|Q^*| \geq 22$ and $|Q| \geq 24$, as claimed. Finally, we may assume that $f_1 \in D$ for some $D \in \mathcal{D}$ included in F_1, and so $D, D' \in \Delta(Q^*)$, and since Q^* is not a joint, $|Q^*| \geq 4$ and so $|Q| \geq 6$. This proves (4).

(5) *If F_1, F_2 are distinct supercomponents, and A is a path of length at most $c + 1$ between F_2 and R, then $d(F_1, A) \geq 3$.*

Suppose that $f_1 \in V(F_1)$ and $v \in V(A)$ have distance at most two. Let f_2 be the end of A in $V(F_2)$. Suppose first that $f_1 \in \Delta$. Consequently every path from f_1 to $V(R)$ has length more than c, and so the subpath of A between v and R has length at least $c - 1$; and therefore the subpath from v to f_2 has length at most two. Hence there is a path Q of length at most four between f_1, f_2; so $|Q| \leq 5$, contrary to (4). This proves that $f_1 \notin \Delta$, and so $f_1 \in X_1$ for some joint X_1 of F_1. By (3), $d(f_1, R) \geq c - (|X_1| - 1)/2$, and so the subpath of A between v, R has length at least $c - (2 + (|X_1| - 1)/2)$. Hence the subpath of A between v, f_2 has length at most $3 + (|X_1| - 1)/2$. Consequently there is a path Q between f_1, f_2 of length at most $5 + (|X_1| - 1)/2$, and so $|Q| \leq 6 + (|X_1| - 1)/2$. Since $|X_1| \leq 8$, and so $|Q| \leq 9$, it follows from (4) that f_2 is not in a joint and $|\Delta(X_1)| = 2$. But then $|X_1| \leq 3$, so $|Q| \leq 7$, contrary to (4). This proves (5).

Let $\mathcal{F}$ be the set of all supercomponents. For each $F \in \mathcal{F}$, let $a(F)$ be the minimum of $a(v)$ over all $v \in V(F)$, and define $b(F)$ similarly. Since for every $D \in \mathcal{D}$, there exists $F \in \mathcal{F}$ with $D \subseteq F$, and hence with $a(F) \le a(D) \le b(D) \le b(F)$, it follows that the set of intervals $\{(a(F), b(F)) : F \in \mathcal{F}\}$ is 4ℓ-powerful. By 3.3, there exists $\mathcal{H} \subseteq \mathcal{F}$ such that $\{(a(F), b(F)) : F \in \mathcal{H}\}$ is ℓ-powerful, and it can be numbered as $\{H_1, \ldots, H_s\}$, where, writing $a(H_i) = a_i$ and $b(H_i) = b_i$ for each i:

(6) *The order of the numbers $a_1, \ldots, a_s$ and $b_1, \ldots, b_s$ is:*

$$0 = a_1 < a_2 < \min(a_3, b_1) < \max(a_3, b_1) < \min(a_4, b_2) < \max(a_4, b_2) < \cdots$$
$$\cdots < \min(a_s, b_{s-2}) < \max(a_s, b_{s-2}) < b_{s-1} < b_s = n.$$

Every two of $a_1, \ldots, a_s, b_1, \ldots, b_s$ differ by at least ℓ, except possibly the pairs (a_1, a_2), (b_{s-1}, b_s), and (a_i, b_{i-2}) for $3 \le i \le s$.

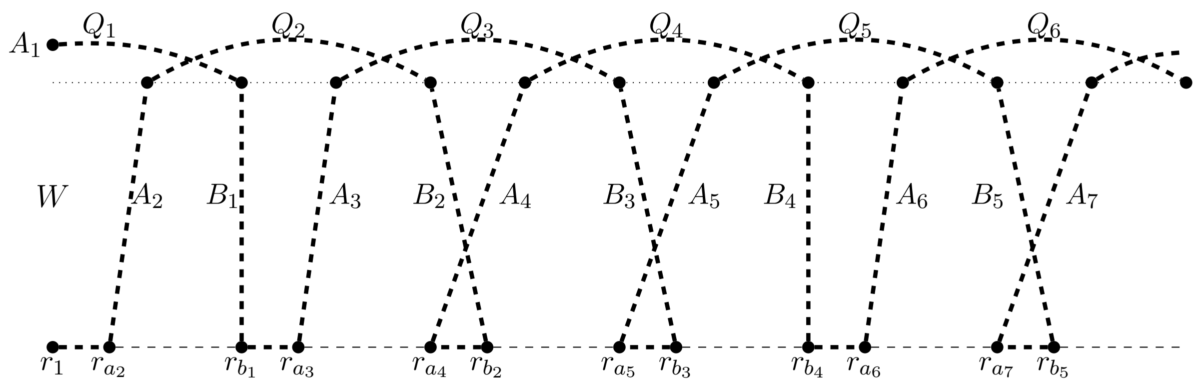

Figure 6: Various paths. The dotted line is the boundary of W. All the dashed lines represent paths, and the one at the bottom is R. The thick dashed lines are the paths we will use to construct a pair of $S-T$ paths with distance three. The paths Q_i have been drawn as though their interiors are disjoint from W, but they might not be; each is a path of a supercomponent, but the supercomponent might penetrate into W a short distance, via joints. In particular the ends of each Q_i have been drawn exactly on the surface, but they might be slightly deeper in W, or just outside of W.

Now we are ready to construct a pair of $S-T$ paths that are distance three apart. (See figure 6.) Since $a(H_1) = 0$, there is a vertex v of H_1 with $a(v) = 0$, and hence with $v \in S$ and $d(v, R) \ge c + 4$; let A_1 be the one-vertex path consisting of this vertex. For $2 \le i \le s$, let A_i be a shortest path with one end in $V(H_i)$ and the other equal to r_{a_i}. Similarly, define B_s to be a one-vertex path with vertex some $v \in V(H_s) \cap T$ with $b(v) = 0$; and for $1 \le i < s$ let B_i be a shortest path with one end in $V(H_i)$ and the other equal to r_{b_i}. It follows that A_i, B_i have length at most $c+1$. Let Q_i be a path of H_i joining the ends of A_i and B_i in H_i. For $1 \le i, j \le n-1$, let $R(i, j) = R(j, i)$ be the subpath of R with ends r_i, r_j. Define $b_0 = 1$ and $a_{s+1} = n - 1$. It follows that the union of the paths

- $A_i \cup Q_i \cup B_i$ for $i \in \{1, \ldots, s\}$, odd; and
- $R(b_i, a_{i+2})$ for $i \in \{0, \ldots, s-1\}$, odd

contains an $S-T$ path. (It might not itself be a path, because for instance the paths B_1, A_3 might intersect). Also, the union of the paths

- $A_j \cup Q_j \cup B_j$ for $j \in \{1, \ldots, s\}$, even; and
- $R(b_j, a_{j+2})$ for $j \in \{0, \ldots, s-1\}$, even

contains an $S-T$ path. We must check that these paths have distance at least three. This is true if $s = 1$, so we assume that $s \geq 2$. We have to check that:

(i) $d(Q_i, Q_j) \geq 3$ for all $i, j \in \{1, \ldots, s\}$ with i odd and j even;

(ii) $d(A_i, A_j), d(A_i, B_j), d(B_i, A_j), d(B_i, B_j) \geq 3$ for all $i, j \in \{1, \ldots, s\}$ with i odd and j even;

(iii) $d(R(b_i, a_{i+2}), R(b_j, a_{j+2})) \geq 3$ for all $i, j \in \{0, \ldots, s-1\}$ with i odd and j even;

(iv) $d(Q_i, A_j), d(Q_i, B_j), d(A_i, Q_j), d(B_i, Q_j) \geq 3$ for all $i, j \in \{1, \ldots, s\}$ with i odd and j even;

(v) $d(Q_i, R(b_j, a_{j+2})) \geq 3$ for all i, j with $i \in \{1, \ldots, s\}$, odd and $j \in \{0, \ldots, s-1\}$, even; and $d(R(b_i, a_{i+2}), Q_j) \geq 3$ for all i, j with $i \in \{0, \ldots, s-1\}$, odd and $j \in \{1, \ldots, s\}$, even; and

(vi) $d(A_i, R(b_j, a_{j+2})), d(B_i, R(b_j, a_{j+2})) \geq 3$ for all i, j with $i \in \{1, \ldots, s\}$, odd and $j \in \{0, \ldots, s-1\}$, even; and $d(R(b_j, a_{j+2}), A_j), d(R(b_j, a_{j+2}), B_j) \geq 3$ for all i, j with $i \in \{0, \ldots, s-1\}$, odd and $j \in \{1, \ldots, s\}$, even.

Statement (i) follows from (4). For statement (ii), suppose first that $i, j \neq 1, s$. Since every vertex in A_i has distance at most $c+1$ from r_{a_i}, and the same for B_i, A_j, B_j, it suffices to check that $d(r_{a_i}, r_{a_j}) \geq 2c+5$ and so on. But these four distances are $|a_i - a_j|, |a_i - b_j|, |b_i - a_j|, |b_i - b_j|$ respectively, and so are all at least $\ell \geq 2c+5$, by (6). If $i = 1$ and $j \neq s$, the same argument shows that $d(B_1, A_j), d(B_1, B_j) \geq 3$, but we need to check that $d(A_1, A_j), d(A_1, B_j) \geq 3$. In this case $V(A_1)$ is a vertex $v \in S$, with $d(v, R) \geq c+4$. Since A_j, B_j both have length at most $c+1$, it follows that $d(v, A_j), d(v, B_j) \geq 3$. The argument is similar if $i = s$ or $j = s$. This proves (ii).

For statement (iii), note that the subpaths $R(b_i, a_{i+2}), R(b_j, a_{j+2})$ of R are disjoint (by (6)), and the distance between them is at least $\ell \geq 3$ (again by (6)). Statement (iv) follows from (5). For statement (v), (3) implies that every path from a supercomponent to R has length at least $c-(8-1)/2 \geq 3$. Finally, for statement (vi), we will prove that $d(A_i, R(b_j, a_{j+2})) \geq 3$ for all i, j with $i \in \{1, \ldots, s\}$, odd and $j \in \{0, \ldots, s-1\}$, even (the other statement is proved similarly). If $i = 1$ then the claim is true since $d(A_1, R) \geq c+4 \geq 3$, so we assume that $i \geq 3$. The distance between r_{a_i} and $R(b_j, a_{j+2})$ is at least ℓ, by (6); and the distance between r_{a_i} and each vertex of A_i is at most $c+1$, and so the distance between A_i and $R(b_j, a_{j+2})$ is at least $\ell - c - 1 \geq 3$. This proves 3.4. ■

References

[1] S. Albrechtsen, T. Huynh, R. W. Jacobs, P. Knappe, and P. Wollan, "The induced two paths problem", `arXiv:2305.04721v3`.

[2] P. Gartland, T. Koorhonen and D. Lokshtanov, "On induced versions of Menger's theorem on sparse graphs", `arXiv:2309.08169`.

[3] A. Georgakopoulos and P. Papasoglu, "Graph minors and metric spaces", `arXiv:2305.07456`.

[4] K. Hendrey, S. Norin, R. Steiner, J. Turcotte, "On an induced version of Menger's theorem", `arXiv:2309.07905`.