A Tight Lower Bound for the Steiner Point Removal Problem on Trees

T.-H. Hubert Chan[1,⋆], Donglin Xia[2,⋆⋆],
Goran Konjevod[2,⋆⋆], and Andrea Richa[3,⋆⋆⋆]

[1] Computer Science Department, Carnegie Mellon University
hubert@cs.cmu.edu
[2] Department of Computer Science and Engineering, Arizona State University
{dxia, goran}@asu.edu
[3] Department of Computer Science and Engineering, Arizona State University
aricha@asu.edu

Abstract. Gupta (SODA'01) considered the Steiner Point Removal (SPR) problem on trees. Given an edge-weighted tree T and a subset S of vertices called *terminals* in the tree, find an edge-weighted tree T_S on the vertex set S such that the distortion of the distances between vertices in S is small. His algorithm guarantees that for any finite tree, the distortion incurred is at most 8. Moreover, a family of trees, where the leaves are the terminals, is presented such that the distortion incurred by any algorithm for SPR is at least $4(1 - o(1))$. In this paper, we close the gap and show that the upper bound 8 is essentially tight. In particular, for complete binary trees in which all edges have unit weight, we show that the distortion incurred by any algorithm for the SPR problem must be at least $8(1 - o(1))$.

1 Introduction

The Steiner Point Removal (SPR) problem was first considered by Gupta [1]. An instance of the problem is given by an edge-weighted tree $T = (V, E)$ and a subset $S \subseteq V$ of vertices called *terminals*. Informally, we would like to find an edge-weighted tree T_S on the terminal set S such that the new tree approximates all the distances between terminal pairs in the original tree. Formally, we say that a weighted tree T_S on the set S has *distortion* at most α if for all $u, v \in S$, the condition $d_T(u, v) \leq d_{T_S}(u, v) \leq \alpha \cdot d_T(u, v)$ holds, where $d_G(u, v)$ is the shortest path distance between two nodes u and v in the graph G. We say an instance has *distortion* at most α if such a tree T_S exists. The objective is to find the smallest constant $\alpha > 0$ such that every instance of the SPR Problem has distortion at most α.

In Gupta's original paper [1], it was shown that $\alpha \leq 8$, i.e., there exists a tree T_S with distortion at most 8. This shows that any submetric of a tree metric is

⋆ Supported in part by the NSF CAREER award CCF-0448095, by an Alfred P. Sloan Fellowship, and by a fellowship from the Croucher Foundation.
⋆⋆ Supported in part by NSF grant CCR-0209138.
⋆⋆⋆ Supported in part by NSF CAREER grant CCR-9985284.

J. Diaz et al. (Eds.): APPROX 2006, LNCS 4110, pp. 70–81, 2006.

"close" to a tree metric. Such a result leads to the first combinatorial proof of the fact that a graph of girth g embeds into a tree with distortion at least $\Omega(g)$, as opposed to the topological proof given by Rabinovich and Raz [2].

Moreover, such a result has potential applications in end system multicast [3,4,5,6]. In a multicast routing protocol, a routing tree $T = (V, E)$ is defined on hosts S, which correspond to the terminals, and routers that connect the hosts and forward messages. The edges represent connections between hosts and routers, and their weights correspond to transmission costs. However, most routers are designed to handle only unicast, and hence a virtual routing tree T_S consisting of only the hosts is suggested for implementing the multicast protocol. Thus, it is important that the virtual tree T_S approximates the original costs well, which is ensured by the upper bound result.

The result has also been used subsequently for embedding k-outerplanar metrics into ℓ_1 by Chekuri et al. [7], embedding general metrics into distributions of tree metrics by Fakcharoenphol et al. [8], and solving the metric labeling problem via tree-rounding by Archer et al. [9].

A natural question to ask is whether the upper bound of 8 is tight. The original paper [1] only gives a lower bound of $4(1 - o(1))$ for some family of trees. In this paper, we close this gap and prove the following theorem showing that the upper bound of 8 is essentially tight.

Theorem 1. *For any $\epsilon > 0$, there exists an instance of the Steiner Point Removal Problem with distortion at least $8 - \epsilon$.*

We anticipate that the techniques presented in this paper may also be applicable to the several open problems in this area, in particular, to the open problems listed in Section 5.

1.1 Proof Strategy

Our lower bound examples will be complete binary trees with unit-weight edges, with the leaves being the terminals. We first show in Section 3 that as far as complete binary trees are concerned, the optimal distortion can always be achieved by a *minor* T_S of the original tree $T = (V, E)$, i.e., the tree T_S can be obtained by contracting edges of tree T of the following form: (1) an edge between two non-terminals; (2) an edge between a terminal x and a non-terminal node y, with the resulting merged node keeping the same name (and terminal status) as x. The weight assigned to each edge (x, y) in T_S will be $d_T(x, y)$, the distance between its two endpoints in the original tree T. Note that each node in V will eventually be contracted into a terminal in S. Thus the minor tree T_S can also be characterized by a mapping $f : V \to S$ that maps each vertex in V to the terminal in S to which it eventually contracts. We call such a mapping f a *minor mapping.*

In Section 4, we show that there exists a complete binary tree such that its minors must incur a large distortion, namely $8 - o(1)$. Let us define some notation before giving the general idea on how one can get such a lower bound:

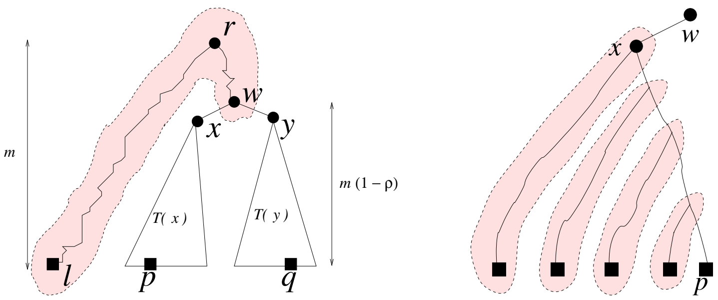
1. Denote by T_n the complete binary tree of height n, having 2^n leaves, with unit-weight edges, and denote by r_n the root of T_n.

2. *Expanding Parameter* $\rho_f(r)$: Suppose the tree T has its root r mapped under f to leaf l, i.e. $l = f(r)$. Suppose that w is a vertex furthest away from the root r in the subtree rooted at the child of r that is not an ancestor of l and $f(w) = l$. Set w to be r if no such vertex exists. The *expanding parameter* $\rho_f(r)$ at r with respect to f is defined to be the ratio $d_T(r, w)/d_T(r, l)$. See Figure 1(a).
3. For each complete binary tree T_n, let ρ_n be the maximum $\rho_f(r_n)$ for all the minor mappings f for T_n with distortion no more than α. Then define $\rho := \limsup_{n\to\infty} \rho_n$.

First we show that $0 < \rho < 1$ (See Claims 4 and 4.). Thus there exists an arbitrarily small constant $\epsilon_1 > 0$ such that $0 < \rho - \epsilon_1 < \rho + \epsilon_1 < 1$. Then by the definition of ρ, there exists an arbitrarily large integer m such that $\rho - \epsilon_1 < \rho_m < \rho + \epsilon_1$. Now consider the complete binary tree T_m and the minor mapping f with distortion no more than α that achieves $\rho_f(r_m) = \rho_m$. As shown in Figure 1(a), let w be the lowest vertex that achieves the *expanding parameter* $\rho_f(r_m)$, vertices x and y be the children of vertex w, and $T(x)$ and $T(y)$ be subtrees rooted at x and y respectively.

The idea is to find leaves p and q in the subtrees $T(x)$ and $T(y)$ respectively such that the distortion exhibited by the pair (p, q) is large. First observe that the distance in T_m between any leaf in $T(x)$ and any leaf in $T(y)$ is $2m(1 - \rho_m) < 2m(1 - (\rho - \epsilon_1))$. Next, we want to argue that there is a leaf p in the subtree $T(x)$ such that the distance between p and $f(r_m)$ in the minor tree $f(T_m)$ is larger than $\frac{2m}{\rho+\epsilon_1}(1 - \epsilon_2)$ for any constant $\epsilon_2 > 0$ if m is large enough. Symmetrically, we can also find such a leaf q in the subtree $T(y)$, thereby the distance between p and q in the minor tree $f(T_m)$ is larger than $\frac{4m}{\rho+\epsilon_1}(1 - \epsilon_2)$ Therefore the distortion according the minor mapping f must be larger than $\frac{2}{(1-(\rho-\epsilon_1))(\rho+\epsilon_1)}(1 - \epsilon_2) \geq 8(1 - \epsilon_2)$. Since the distortion of f is no more than α, we get the lower bound $\alpha > 8 - o(1)$.

We still need to determine how to find such a leaf p in the subtree $T(x)$. We will use a recursive algorithm on the roots of the subtrees considered, starting with the subtree $T(x)$. First we limit p to be one of the leaves in $T(x)$, whose distances to $f(r_m)$ in T_m are all $2m$. Then, we limit p to be one of the leaves of $T(x)$ in the subtree of x that does not contain $f(x)$; the distances of those leaves to $f(x)$ in T_m are all $2m(1 - \rho_m) - 2 \gtrapprox 2m(1 - (\rho + \epsilon_1))$. In general, as shown in Figure 1(b), we limit p to be one of the leaves of the subtree of $T(z)$ (initially $z = x$) that does not contain $f(z)$; we then let z be the root of the corresponding subtree, and recurse. Roughly speaking, the heights of these trees are no less than m, $m(1 - (\rho + \epsilon_1))$, $m(1 - (\rho + \epsilon_1))^2$, $m(1 - (\rho + \epsilon_1))^3$, $\cdots$, respectively, if m is large enough (See Lemma 1 for a formal proof). Thus the distance between p and $f(r_m)$ in the minor tree $f(T_m)$ must be larger than $\frac{2m}{\rho+\epsilon_1}(1 - \epsilon_2)$, where $\epsilon_2 > 0$ can be any constant and m is large enough. Therefore our algorithm finds such a leaf p, and it follows that $\alpha > 8 - o(1)$.



(a) Expanding parameter for the root of T_m

(b) Select p in $T(x)$

Fig. 1. The Minor Construction for Tree T_m (Shadow areas refer to components contracted to a terminal)

2 Notation

In this section, we will introduce and formalize some additional notation that will be used in Sections 3 and 4. Suppose T is a tree with edge set E and a positive distance associated with each edge. We denote the distance of the unique shortest path between two vertices u and v by $d_T(u, v)$. We use $\mathsf{L}(T)$ to denote the set of leaves, i.e. the degree-one vertices in T.

As defined in Section 1.1, we denote by T_n the complete binary tree of height n, having 2^n leaves with unit weight edges. We denote by r_n the root of T_n and the terms *child, parent, ancestor and descendant* are used with their usual meanings. From now on, we restrict the SPR Problem to such trees, with the leaves being the terminals.

Formally, we say f is a transformation from T to $\widehat{T}$, if $\widehat{T} = (\mathsf{L}(T), \widehat{E})$ is a tree on the vertex set $\mathsf{L}(T)$, and each edge $(u, v) \in \widehat{E}$ has weight $d_T(u, v)$. The distortion of such a transformation is

$$\mathsf{D}(f) := \max_{x \neq y \in \mathsf{L}(T)} \frac{d_{\widehat{T}}(x, y)}{d_T(x, y)}.$$

A transformation f from T to $\widehat{T}$ is *minor* if $\widehat{T}$ is a minor of T, i.e. $\widehat{T}$ can be obtained from T by edge contractions. Note that a minor transformation f for a tree T can be equivalently viewed as a mapping $f : \mathsf{V}(T) \to \mathsf{L}(T)$ that maps each vertex to the terminal to which it eventually contracts. We call such f a *minor mapping.*

3 Restricting to Minor Transformations

In this section, we show that in order to obtain a lower bound on the distortion of transformations for complete binary trees, it suffices to consider minor transformations.

The radius of a tree T is given by $R(T) = \min_{u \in \mathsf{V}(T)} \max_{v \in \mathsf{V}(T)} d_T(u, v)$. A *center point* of T is a vertex $u_0 \in \mathsf{V}(T)$ such that $R(T) = \max_{v \in \mathsf{V}(T)} d_T(u_0, v)$.

Theorem 2. *For any $n \geq 0$ and for any transformation f of T_n, there exists a minor transformation f' such that*

(a) the distortion of f' does not increase, $\mathsf{D}(f') \leq \mathsf{D}(f)$;
(b) the radius does not increase, $R(f'(T_n)) \leq R(f(T_n))$;
(c) the terminal $f'(r_n)$ is a center point of $f'(T_n)$.

Proof: We argue by induction on n. The case $n = 0$ is trivial. For the case $n = 1$, there is only one transformation for T_1, which is minor and satisfies the requirements.

Assume the result holds true for any T_k, where $k < n$. Consider some transformation $f : T_n \rightarrow \widehat{T}_n$.

We denote by $[n]$ the set of integers $\{0, 1, \ldots, n\}$.

For any $x \in \mathsf{L}(T_n)$ and $i \in [n]$, denote by $T_i(x)$ the i-level complete binary subtree of T_n which contains x; denote the root of $T_i(x)$ by $r_i(x)$. For any $x \in \mathsf{L}(T_n)$ and $i \in [n]$, denote by $S_i(x)$ the minimal subtree of $\widehat{T}_n$ that includes all the vertices in $\mathsf{L}(T_i(x))$. Let k be the maximum integer such that for any $x \in \mathsf{L}(T_n)$, $\mathsf{V}(S_k(x)) \subseteq \mathsf{L}(T_{n-1}(x))$. Since $k = 0$ satisfies the above conditions, such a k exists. Note that $k < n$; otherwise, $\mathsf{L}(T_n(x)) \subseteq \mathsf{V}(S_k(x)) \subseteq \mathsf{L}(T_{n-1}(x))$, which is a contradiction.

From the maximality of k, there exists $u \in \mathsf{L}(T_n)$ such that $\mathsf{V}(S_{k+1}(u)) \nsubseteq \mathsf{L}(T_{n-1}(u))$. Also, there exists $v \in \mathsf{L}(T_{k+1}(u))$ such that $T_k(v) \neq T_k(u)$ and the u-v path in $\widehat{T}_n$ uses some vertex not in $\mathsf{L}(T_{n-1}(u))$. Let vertex $w \notin \mathsf{L}(T_{n-1}(u))$ be the first such vertex on the path from u to v, and $u' \in \mathsf{L}(T_{n-1}(u))$ be the previous vertex of w on the path. Since $T_{n-1}(u') \neq T_{n-1}(w)$, it follows that (u', w) has weight $2n$.

Claim. Edge (u', w) is an edge of weight $2n$ that separates $S_k(u)$ and $S_k(v)$ in $\widehat{T_n}$.

Proof of Claim 3: By the definition of k, $\mathsf{V}(S_k(u)) \subseteq \mathsf{L}(T_{n-1}(u))$ and $\mathsf{V}(S_k(v)) \subseteq \mathsf{L}(T_{n-1}(v))$. Since $w \notin \mathsf{L}(T_{n-1}(u))$, edge (u', w) separates $S_k(u)$ and v. Since $u' \in \mathsf{L}(T_{n-1}(u))$ and $w \notin \mathsf{L}(T_{n-1}(u))$, then exactly one of them is not in $\mathsf{L}(T_{n-1}(v))$. Since $\mathsf{V}(S_k(v)) \subseteq \mathsf{L}(T_{n-1}(v))$, edge (u', w) separates $S_k(v)$ and u. Therefore edge (u', w) separates $S_k(u)$ and $S_k(v)$. □

Thus in the tree $\widehat{T}_n$, there is a unique path connecting $S_k(u)$ and $S_k(v)$ with all its intermediate vertices not in $\mathsf{V}(S_k(u)) \cup \mathsf{V}(S_k(v))$. Let $u_0 \in \mathsf{V}(S_k(u))$ and $v_0 \in \mathsf{V}(S_k(v))$ be the two endpoints of the path. Then, vertex w is on the u_0-v_0 path and $d_{\widehat{T}_n}(u_0, w) \geq 2n$.

If $k + 1 < n$, then $v \in \mathsf{L}(T_{k+1}(u)) \subseteq \mathsf{L}(T_{n-1}(u))$, thereby $d_{\widehat{T}_n}(v_0, w) \geq 2n$; if $k + 1 = n$, we have the trivial bound $d_{\widehat{T}_n}(v_0, w) \geq 0$.

Consider vertices $u_1 \in \mathsf{V}(S_k(u))$ and $v_1 \in \mathsf{V}(S_k(v))$, which are furthest away from u_0 and v_0 respectively. Hence, we have $d_{\widehat{T}_n}(u_0, u_1) \geq R(S_k(u))$ and $d_{\widehat{T}_n}(v_0, v_1) \geq R(S_k(v))$. Without loss of generality, assume $R(S_k(u)) \leq R(S_k(v))$.

Observing that $d_{\widehat{T}_n}(u_1, v_1) = d_{\widehat{T}_n}(u_1, u_0) + d_{\widehat{T}_n}(u_0, w) + d_{\widehat{T}_n}(w, v_0) + d_{\widehat{T}_n}(v_0, v_1)$, we have

$$\mathsf{D}(f) \geq \frac{d_{\widehat{T}_n}(u_1, v_1)}{d_{T_n}(u_1, v_1)} \geq \begin{cases} \frac{4n+2R(S_k(u))}{2(k+1)} & \text{if } k+1 < n; \\ \frac{2n+2R(S_k(u))}{2(k+1)} & \text{if } k+1 = n \end{cases} \tag{3.1}$$

Also,

$$R(f(T_n)) \geq 2n + R(S_k(u)) \tag{3.2}$$

Next, we construct a transformation g for the subtree $T_k(u)$. We obtain the transformed tree $\widehat{T}_k(u)$ from $S_k(u)$, the minimal subtree in $\widehat{T}_n$ containing $\mathsf{L}(T_k(u))$, by contracting all the vertices $v \notin \mathsf{L}(T_k(u))$ as follows:

1. Contract any edge neither of whose endpoints is in $\mathsf{L}(T_k(u))$.
2. For each remaining vertex $x \notin \mathsf{L}(T_k(u))$, contract one of the edges incident to x.
3. For each edge (x, y) in $\widehat{T}_k(u)$ set its weight as $d_{T_k(u)}(x, y)$, i.e. $d_{T_n}(x, y)$.

The following claim states the properties of the transformation g. Its proof is technical and will be deferred to the end of the section.

Claim. Suppose the transformation g from $T_k(u)$ to the tree $\widehat{T}_k(u) = (\mathsf{L}(T_k(u)), \widehat{E})$ is as described above. Then, the distortion $\mathsf{D}(g) \leq \mathsf{D}(f)$ and the radius
$R(g(T_k(u))) \leq R(S_k(u))$.

By the induction hypothesis , there exists a minor transformation g' for $T_k(u)$ such that $\mathsf{D}(g') \leq \mathsf{D}(g)$, $R(g'(T_k(u))) \leq R(g(T_k(u)))$, and $r_k(u)$ is contracted into a center point of $g'(T_k(u))$. By Claim 3, we also have $\mathsf{D}(g) \leq \mathsf{D}(f)$ and $R(g(T_k(u))) \leq R(S_k(u))$. Hence, we have $\mathsf{D}(g') \leq \mathsf{D}(f)$ and $R(g'(T_k(u))) \leq R(S_k(u))$.

We next use the transformation g' to construct a minor transformation f' for T_n. Since all the k-level complete binary subtrees T_k of T_n are isomorphic to $T_k(u)$, the transformation g' also defines a minor transformation for each of these subtrees T_k. Then a minor transformation f' for T_n can be obtained by edge contractions as follows:

1. Remove internal nodes in each T_k via edge contraction using minor transformation g'.
2. Since the $(n-k-1)$-level complete binary subtree rooted at r_n is the remaining component for contraction, we just contract the whole subtree into its adjacent vertex in $g'(T_k(u))$.

Therefore, r_n and $r_k(u)$ are contracted to the same leaf. Hence, r_n is contracted into a center point of $g'(T_k(u))$. In fact, the tree $f'(T_n)$ consists of components $g'(T_k)$ and additional edges connecting the center point of $g'(T_k(u))$ to the center points of the other components. Moreover if $k+1=n$, $f'(T_n)$ only has two components $g'(T_k)$, thereby its diameter is $2n+2\cdot R(g'(T_k(u)))$. And if $k+1<n$, $f'(T_n)$ has more than two components $g'(T_k)$, thereby its diameter is $4n+2\cdot R(g'(T_k(u)))$. Thus

$$\mathsf{D}(f') = \begin{cases} \max(\mathsf{D}(g'), \frac{4n+2\cdot R(g'(T_k(u)))}{2(k+1)}) \text{ if } k+1<n; \\ \max(\mathsf{D}(g'), \frac{2n+2\cdot R(g'(T_k(u)))}{2(k+1)}) \text{ if } k+1=n; \end{cases} \tag{3.3}$$

Thus, by Equation (3.1) and the relationship between the transformations g' and f, we have $\mathsf{D}(f') \leq \mathsf{D}(f)$, proving part (a) of the theorem. Moreover, by Equation (3.2), we obtain part(b)

$$R(f'(T_n)) = 2n + R(g'(T_k(u))) \leq R(f(T_n)), \tag{3.4}$$

and r_n is contracted into a center point of $g'(T_k(u))$, which can be verified to be a center point of $R(f'(T_n))$, hence proving part(c). □

We next give the proof of Claim 3, as promised earlier.

Proof of Claim 3: We first observe that any maximal connected component C in the tree $S_k(u)$ that does not contain any vertex in $\mathsf{L}(T_k(u))$ will be contracted into a vertex of $\mathsf{L}(T_k(u))$.

We will use the following fact about distances between leaves.

Fact 3. *Any edge between two leaves in* $\mathsf{L}(T_k(u))$ *has weight at most* $2k$*; and any edge between a leaf in* $\mathsf{L}(T_k(u))$ *and one outside it has weight at least* $2(k+1)$*.*

1. To show $\mathsf{D}(g) \leq \mathsf{D}(f)$, we prove that $d_{\widehat{T}_k(u)}(x,y) \leq d_{\widehat{T}_n}(x,y)$ for any $x,y \in \mathsf{L}(T_k(u))$.
   Fix any $x,y \in \mathsf{L}(T_k(u))$. Let P be the x-y path in $\widehat{T}_k(u)$ and Q be the x-y path in $S_k(u)$.
   Since any maximal connected component C excluding vertices in $\mathsf{L}(T_k(u))$ in the tree $S_k(u)$ is contracted into one vertex of $\mathsf{L}(T_k(u))$, any maximal subpath Q' of Q excluding vertices in $\mathsf{L}(T_k(u))$ is contracted into some vertex c of $\mathsf{L}(T_k(u))$. By maximality of Q', there exists $a,b \in \mathsf{L}(T_k(u))$ on path Q such that a-Q'-b is a subpath of Q, which would become a subpath a-c-b in P. By Fact 3, the length of this subpath decreases.
   On the other hand, an edge in Q that joins two vertices in $\mathsf{L}(T_k(u))$ remains in P and its weight does not change.
   Hence, it follows that the length of P is at most that of Q.
   Therefore,

$$d_{\widehat{T}_k(u)}(x,y) \leq d_{\widehat{T}_n}(x,y) \text{ for any } x,y \in \mathsf{L}(T_k(u)) \tag{3.5}$$

   Thus $\mathsf{D}(g) \leq \mathsf{D}(f)$.

2. Next we show that $R(g(T_k(u))) \leq R(S_k(u))$.
Let $u_0 \in \mathsf{V}(S_k(u))$ be the center point of $S_k(u)$. By the minimality of $S_k(u)$, this radius must be realized by some vertex in $\mathsf{L}(T_k(u))$.

$$R(S_k(u)) = \max_{x \in L_k(u)} (d_{\widehat{T}_n}(u_0, x)) \tag{3.6}$$

If $u_0 \in \mathsf{L}(T_k(u)) = \mathsf{V}(\widehat{T}_k(u))$, then by Equations (3.5) and (3.6),

$$R(\widehat{T}_k(u)) \leq \max_{x \in \mathsf{L}(T_k(u))} d_{\widehat{T}_k(u)}(u_0, x) \leq \max_{x \in \mathsf{L}(T_k(u))} (d_{\widehat{T}_n}(u_0, x)) = R(S_k(u)).$$

If $u_0 \notin \mathsf{L}(T_k(u)) = \mathsf{V}(\widehat{T}_k(u))$, then let $u_0' \in \mathsf{V}(\widehat{T}_k(u))$ be the vertex into which u_0 is contracted. For any $x \in \mathsf{L}(T_k(u)) = \mathsf{V}(\widehat{T}_k(u))$, let P be the u_0'-x path in $\widehat{T}_k(u)$ and Q be the u_0-x path in $S_k(u)$.
Observe that the initial maximal subpath Q' of Q excluding vertices in $\mathsf{L}(T_k(u))$ is contracted into u_0'. Let u_1 be the first vertex on Q in the direction from u_0 to x such that $u_1 \in \mathsf{L}(T_k(u))$. Hence, the subpath Q'-u_1 becomes a subpath u_0'-u_1 in P, whose length decreases by Fact 3. By Equation (3.5), the length of the remaining subpath of P is at most that of the remaining subpath of Q. Hence, the length of P is at most that of Q.
Therefore,

$$R(\widehat{T}_k(u)) \leq \max_{x \in \mathsf{V}(\widehat{T}_k(u))} d_{\widehat{T}_k(u)}(u_0', x) \leq \max_{x \in \mathsf{V}(S_k(u))} d_{\widehat{T}_n}(u_0, x) = R(S_k(u))$$

Thus, we also have $R(g(T_k(u))) \leq R(S_k(u))$ in this case. □

4 A Lower Bound for Minor Transformations

In view of Theorem 2 in the previous section, we consider only minor transformations for complete binary trees.

Definition 4 (Optimal distortion for minor transformation). We define $\alpha \geq 1$ to be the smallest constant such that for any instance of the SPR Problem, there exists a *minor* transformation that achieves distortion at most α.

Observe that the algorithm given by Gupta [1] indeed produces a minor with distortion at most 8. Hence, the constant α is at most 8. We prove the following theorem, which implies that the constant $\alpha \geq 8$.

Theorem 5. *For any $\epsilon > 0$, the constant $\alpha \geq 8 - \epsilon$.*

Hence, combining Theorems 2 and 5, we obtain the result of Theorem 1, which states that:

> For any $\epsilon > 0$, there exists an instance of the Steiner Point Removal Problem with distortion at least $8 - \epsilon$.

We first introduce some notation. Without causing ambiguity, we use $d(u, v)$ to denote the distance between nodes u and v in the original tree T, and $path(u, v)$ to denote the subset of vertices lying on the unique path between u and v in T. Let v be a vertex in T_n. We denote the subtree rooted at v by $T(v)$, which is identical to $T_{n-d(r_n,v)}$. For u, $v \in \mathsf{L}(T)$, we use $d_f(u, v)$ to denote the distance between them after the transformation f is applied to the tree.

Definition 6. Given a minor mapping $f : \mathsf{V}(T) \to \mathsf{L}(T)$, a vertex v is a *normal vertex* (with respect to f) if v is an ancestor of $f(v)$.

Consider a normal vertex v and suppose $u = f(v)$. Then, v is an ancestor of u and all the vertices along the path from v to u are mapped to u. Recall that $T(v)$ has two branches rooted at v. We wish to measure how far vertices down the branch *not* containing u are mapped to u under f.

Definition 7. For each normal vertex v, its *expanding parameter* with respect to some minor mapping f is defined to be

$$\rho_f(v) := \max\{\tfrac{d(v,w)}{d(v,f(v))} : w \in T(v),\ f(w) = f(v),\ path(v, f(v)) \cap path(v, w) = \{v\}\}.$$

Since our lower bound is obtained from large trees, we consider how the expanding parameter behaves for large values of n.

Definition 8. For each $n \in \mathbb{N}$, let

$$\rho_n := \max\{\rho_f(r_n) \,|\, \text{minor mapping } f : T_n \to \mathsf{L}(T_n),\ \mathsf{D}(f) \leq \alpha\}.$$

Define

$$\rho := \limsup_{n\to\infty} \rho_n. \tag{4.7}$$

Observe that since $\rho_n \in [0, 1]$, it follows the limit supremum $\rho \in [0, 1]$. We show in the next claim that ρ is strictly less than 1.

Claim. The limit supremum $\rho < 1$.

Proof: Assume on the contrary that $\rho = 1$. Then, by the definition of limit supremum ρ, there exists an integer n such that $\rho_n \geq 7/8$. Thus by the definition of ρ_n, there exists a minor mapping f on T_n with $\mathsf{D}(f) \leq \alpha$ such that $\rho_f(r_n) \geq 7/8$.

Let w be a vertex that attains $\rho_f(r_n)$. Since every leaf of T_n is mapped into itself and $w \neq f(w)$, w is not a leaf. Then let p and q be two leaves from different branches of the subtree $T(w)$. Thus $d(p, q) = 2(1-\rho_f(r_n))n \leq n/4$. On the other hand, $d_f(p, q) = d_f(p, f(w)) + d_f(f(w), q) \geq 4n$. Thus $\mathsf{D}(f) \geq \frac{d_f(p,q)}{d(p,q)} \geq \frac{4n}{n/4} \geq 16$, contradicting $\mathsf{D}(f) \leq \alpha \leq 8$. Thus $\rho < 1$. □

The following lemma shows the relationship between the expanding parameter ρ_n and the distorted distance d_f. Intuitively, if the expanding parameters for normal vertices of large heights are small, then there exists some vertex whose distorted distance to the image of the root is large.

Lemma 1. *Suppose* $0 < \beta < 1$ *and* $N_0 \in \mathbb{N}$ *such that for any integer* $n > N_0$, *the expanding parameter* $\rho_n \leq \beta$. *Then, for any real* $0 < \epsilon < 1$, *there exists integer* $N > N_0$ *such that for any integer* $m \geq N$ *and any minor mapping* f *on tree* T_m *with distortion* $\mathsf{D}(f) \leq \alpha$, *there exists a leaf* p *in* T_m *such that the distorted distance*

$$d_f(p, f(r_m)) \geq \frac{2m}{\beta}(1-\epsilon).$$

Furthermore, if $\rho_f(r_m) > 0$, *then* $\mathsf{D}(f) \geq \frac{2(1-\epsilon)}{\beta(1-\rho_f(r_m))}$.

Proof: Given any real $\epsilon > 0$, fix a large enough integer k such that $(1-\beta)^k \leq \frac{\epsilon}{2}$. Let N be large enough such that $\frac{k}{N} \leq \frac{\epsilon}{2}$ and $(1-\beta)^k(N+\frac{1}{\beta}) - \frac{1}{\beta} > N_0$.

Let $m \geq N$ and let f be a minor mapping on T_m with $\mathsf{D}(f) \leq \alpha$. We define sequences of vertices $\{v_i\}_{i=0}^{k}$ and $\{w_i\}_{i=0}^{k-1}$ in T_m as follows. Let $v_0 = r_m$, and w_0 be the vertex that attains $\rho_f(v_0)$ under the minor mapping f with $\mathsf{D}(f) \leq \alpha$. For $1 \leq i \leq k$, let v_i be a child of vertex w_{i-1} such that $f(w_{i-1}) \notin T(v_i)$, and hence v_i is normal. Let w_i be the vertex that attains $\rho_f(v_i)$, for $1 \leq i < k$. Let h_i be the height of $T(v_i)$ for $0 \leq i \leq k$.

Claim. For $0 \leq i < k$, the height $h_i \geq (1-\beta)^i(m+\frac{1}{\beta}) - \frac{1}{\beta} > N_0$.

Proof of Claim 4: The claim is trivial for $i = 0$. Assume that $h_{i-1} \geq (1-\beta)^{i-1}(m+\frac{1}{\beta}) - \frac{1}{\beta} > N_0$, for some $0 < i < k$. Observe that $h_i + 1 + \rho_f(v_{i-1})h_{i-1} = h_{i-1}$ and $\rho_f(v_{i-1}) \leq \beta$, since $h_{i-1} > N_0$. Then $h_i = (1-\rho_f(v_{i-1}))h_{i-1} - 1 \geq (1-\beta)\{(1-\beta)^{i-1}(m+\frac{1}{\beta}) - \frac{1}{\beta}\} - 1 = (1-\beta)^i(m+\frac{1}{\beta}) - \frac{1}{\beta} > N_0$. □

Thus, we set $p := f(v_k)$ and from Claim 4, we have

$$\begin{aligned} d_f(f(r_m), p) &= 2\sum_{i=0}^{k-1} h_i \geq 2\sum_{i=0}^{k-1}\{(1-\beta)^i(m+\frac{1}{\beta}) - \frac{1}{\beta}\} \\ &= 2(m+\frac{1}{\beta})\frac{1-(1-\beta)^k}{\beta} - \frac{2k}{\beta} \geq \frac{2m}{\beta}\cdot\{1-(1-\beta)^k - \frac{k}{m}\} \\ &\geq \frac{2m}{\beta}(1-\epsilon), \end{aligned} \tag{4.8}$$

where the last inequality follows from $(1-\beta)^k \leq \frac{\epsilon}{2}$ and $\frac{k}{m} \leq \frac{k}{N} \leq \frac{\epsilon}{2}$.

Furthermore, if $\rho_f(r_m) > 0$, then $m \cdot \rho_f(r_m) > 0$. Thus w_0 is a proper descendant of r_m. Note that p is a leaf of $T(w_0)$ and $T(w_0)$ has two branches. Thus by symmetry, there exists another leaf q such that p and q are in the different branches of $T(w_0)$ and $d_f(q, f(r_m)) \geq \frac{2m}{\beta}(1-\epsilon)$. Observing that $f(w_0) = f(r_m)$, the distorted distance $d_f(p,q) = d_f(p, f(r_m)) + d_f(f(r_m), q) \geq \frac{4m}{\beta}(1-\epsilon)$, and the original distance $d(p,q) = 2m(1-\rho_f(r_m))$. Therefore, the distortion $\mathsf{D}(f) \geq \frac{d_f(p,q)}{d(p,q)} \geq \frac{2(1-\epsilon)}{\beta(1-\rho_f(r_m))}$. □

Using Lemma 1, we can show that the limit supremum $\rho > 0$.

Claim. The limit supremum $\rho > 0$.

Proof of Lemma 4: On the contrary, suppose $\rho = 0$. Let $\beta = 1/32$. By the definition of limit supremum ρ, there exists N_0 such that for any $n > N_0$, $\rho_n < \beta$. Then by Lemma 1, for $\epsilon = 1/2$, there exists $m > N_0$ such that for any minor mapping f on T_m with $\mathsf{D}(f) \le \alpha$, there exists a leaf p in T_m such that $d_f(p, f(r_m)) \ge \frac{2m}{\beta}(1-\epsilon) = \frac{m}{\beta}$. Thus $\mathsf{D}(f) \ge \frac{d_f(p,f(r_m))}{d(p,f(r_m))} \ge \frac{m}{2m\beta} = \frac{1}{2\beta} = 16$, which contradicts $\mathsf{D}(f) \le \alpha \le 8$. Thus $\rho > 0$. □

Now, we are ready to prove the main theorem of this section.

Proof of Theorem 5: Let $\epsilon > 0$. Without loss of generality, we can assume $\epsilon < 1$. Suppose on the contrary, we have $\alpha < 8 - \epsilon$.

Since $0 < \rho < 1$, let $\epsilon_1 < \min\{\epsilon/48, \rho\}$ be a positive small constant such that $\rho + \epsilon_1 < 1$. By the definition of limit supremum ρ, there exists $N_0 > 0$ such that for all $n > N_0$, $\rho_n < \rho + \epsilon_1$. Then by Lemma 1, for $\epsilon_2 = \epsilon/24$ there exists N such that for any $m \ge N$ and any minor mapping f on tree T_m with distortion $\mathsf{D}(f) \le \alpha$ and $\rho_f(r_m) > 0$ we have $\mathsf{D}(f) \ge \frac{2(1-\epsilon_2)}{(\rho+\epsilon_1)(1-\rho_f(r_m))}$.

By the definition of limit supremum ρ, there exists arbitrarily large m such that $\rho_m > \rho - \epsilon_1 > 0$. Hence, we can choose m such that $m > N$. By the definition of ρ_m, there exists a minor mapping f on tree T_m with distortion $\mathsf{D}(f) \le \alpha$ and $\rho_f(r_m) = \rho_m > \rho - \epsilon_1 > 0$. Thus, the constant α is at least

$$\begin{aligned}
\mathsf{D}(f) &\ge \tfrac{2(1-\epsilon_2)}{(\rho+\epsilon_1)(1-\rho_f(r_m))} \\
&\ge \tfrac{2(1-\epsilon_2)}{(\rho+\epsilon_1)(1-(\rho-\epsilon_1))} \ge \tfrac{8(1-\epsilon_2)}{(1+2\epsilon_1)^2} && \text{(The denominator is min when } \rho = \tfrac{1}{2}.\text{)} \\
&\ge \tfrac{8(1-\epsilon_2)}{(1+\epsilon_2)^2} \ge 8(1-3\epsilon_2) && \text{(Note: } 2\epsilon_1 \le \epsilon_2\text{; as } \epsilon_2 \ge 0,\ \tfrac{1-\epsilon_2}{(1+\epsilon_2)^2} \ge 1-3\epsilon_2\text{)} \\
&= 8 - \epsilon,
\end{aligned}$$

obtaining the desired contradiction. Hence, for all $\epsilon > 0$, the constant $\alpha \ge 8 - \epsilon$. □

5 Open Problems

We conclude the paper by outlining some directions for future work.

1. Of course one final goal would be to consider the SPR problem on general graphs. Formally, there are two main questions to be addressed: (1) we would like to determine what is the smallest α (possibly depending on the size of input), such that given any edge weighted graph $G = (V, E)$ and a set of terminals $S \subset V$, there is a way to remove non-terminals by edge contractions to produce a minor $H = (S, E')$ where for any pair of terminals (u, v), $d_G(u, v) \le d_H(u, v) \le \alpha \cdot d_G(u, v)$; and (2) we would like to devise a constructive algorithm that outputs such a minor $H = (S, E')$ with distortion at most α. Since this task may prove to be quite hard to accomplish on general graphs, one could first consider other restricted classes of graphs — such as outerplanar graphs, planar graphs, series-parallel graphs, etc. — as an intermediate step. Note that no algorithm with proven nontrivial bounds on distortion for these classes of graphs is known.

2. Another interesting question is to be able to determine the *approximation bound on the optimal distortion* of a given algorithm for the SPR problem. For example, it would be interesting to determine, given any instance of the SPR problem on trees, how far from the optimal distortion for that instance can the distortion obtained by Gupta's algorithm [1] be (in that paper, Gupta only shows an absolute bound on the distortion of his algorithm; this paper confirms that for some instances of the problem, this is the best distortion possible).
3. We can also ask a similar question as that in Problem 1 in a probabilistic framework. What is the smallest α such that given any weighted graph $G = (V, E)$ and a set of terminals $T \subset V$, there exists a distribution $\mathcal{H}$ of minors $\{H = (T, E')\}$ such that $d_G(u, v) \leq E_{\mathcal{H}}[d_H(u, v)] \leq \alpha \cdot d_G(u, v)$? This task may be easier to accomplish than that in Problem 1, since some upper bounds on α under a probabilistic framework exist in the literature. For example, it follows from [7] that k-outerplanar graph can be embedded into a probability distribution over spanning trees with $O(c^k)$ distortion for some absolute constant c, implying that $\alpha = O(c^k)$ for k-outerplanar graphs; and a recent result by Elkin et. al. [10] shows that for general graphs, $\alpha = O(\log^2 n \log\log n)$, which is later improved to $O(\log^2 n)$ by Dhamdhere et. al. [11], shows that for general graphs, $\alpha = O(\log^2 n)$. Can we do any better?

References

1. Gupta, A.: Steiner points in tree metrics don't (really) help. SODA (2001) 220–227
2. Rabinovich, Y., Raz, R.: Lower bounds on the distortion of embedding finite metric spaces in graphs. Discrete Comput. Geom. **19**(1) (1998) 79–94
3. Chu, Y., Rao, S., Zhang, H.: A case for end system multicast. In: Proceedings of ACM Sigmetrics, Santa Clara, CA. (2000)
4. Xie, J., Talpade, R.R., Mcauley, A., Liu, M.: Amroute: ad hoc multicast routing protocol. Mob. Netw. Appl. **7**(6) (2002)
5. Chawathe, Y.: Scattercast: an adaptable broadcast distribution framework. Multimedia Syst. **9**(1) (2003)
6. Francis, P.: Yoid: Extending the internet multicast architecture. (2000)
7. Chekuri, C., Gupta, A., Newman, I., Rabinovich, Y., Sinclair, A.: Embedding k-outerplanar graphs into ℓ_1. In: Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms. (2003) 527–536
8. Fakcharoenphol, J., Rao, S., Talwar, K.: A tight bound on approximating arbitrary metrics by tree metrics. In: Proceedings of the thirty-fifth ACM symposium on Theory of computing, ACM Press (2003) 448–455
9. Archer, A., Fakcharoenphol, J., Harrelson, C., Krauthgamer, R., Talwar, K., Tardos, E.: Approximate classification via earthmover metrics. In: In 15th Annual ACM-SIAM Symposium on Discrete Algorithms. (2004) 1072–1080
10. Elkin, M., Emek, Y., Spielman, D.A., Teng, S.H.: Lower-stretch spanning trees. In: Proceedings of the 37th Annual ACM Symposium on Theory of Computing. (2005) 494–503
11. Dhamdhere, K., Gupta, A., Räcke, H.: (Improved embeddings of graph metrics into random trees)