Euclidean Bounded-Degree Spanning Tree Ratios*

Timothy M. Chan†
School of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
tmchan@uwaterloo.ca

January 29, 2004

Abstract

Let τ_K be the worst-case (supremum) ratio of the weight of the minimum degree-K spanning tree to the weight of the minimum spanning tree, over all finite point sets in the Euclidean plane. It is known that $\tau_2 = 2$ and $\tau_5 = 1$. In STOC'94, Khuller, Raghavachari, and Young established the following inequalities: $1.103 < \tau_3 \leq 1.5$ and $1.035 < \tau_4 \leq 1.25$. We present the first improved upper bounds: $\tau_3 < 1.402$ and $\tau_4 < 1.143$. As a result, we obtain better approximation algorithms for Euclidean minimum bounded-degree spanning trees.

Let $\tau_K^{(d)}$ be the analogous ratio in d-dimensional space. Khuller et al. showed that $\tau_3^{(d)} < 1.667$ for any d. We observe that $\tau_3^{(d)} < 1.633$.

1 Introduction

The starting point of this work is the following well-known observation [6, 20]: for finite point sets in any metric space, we can construct a spanning path (or cycle) of at most twice the weight of the minimum spanning tree (MST), by doubling the MST edges, taking an Euler tour, and short-cutting repeated vertices. This strategy yields a simple factor-2 approximation algorithm for the traveling salesman path (or tour) problem. It was shown by Fekete et al. [9] that even in the geometric case of the Euclidean metric in the plane, the analysis cannot be improved upon (in other words, short-cutting doesn't help much in the worst case), as there exists point sets whose traveling salesman path weight is more than $2 - \varepsilon$ times the MST weight for any fixed $\varepsilon > 0$. However, by using other lower bounds besides the MST weight, it is possible to obtain better approximation guarantees for the traveling salesman problem in polynomial time, as was demonstrated by Christofides [5] (factor 3/2) for general metrics, and Arora [1] and Mitchell [14] (factor $1 + \varepsilon$) for the Euclidean metric in fixed dimensions.

The focus of the present paper is on the following generalization of the traveling salesman path problem (which corresponds to the $K = 2$ case): given K, find a spanning tree, of minimum weight such that the maximum degree is at most K. The degree constraint is natural to consider, since

*A preliminary version of this paper has appeared in *Proc. 19th ACM Sympos. Comput. Geom.*, 2003.
†This work was supported in part by an NSERC Research Grant.

high-degree nodes in networks are in many ways undesirable. The $K = 3$ case is especially appealing, since once rooted, a degree-3 tree becomes a binary tree.

For this bounded-degree spanning tree problem, Christofides' algorithm no longer gives a 3/2 approximation factor, and at present, a gneralization of Arora or Mitchell's polynomial-time approximation scheme is not known either [2] (although we have recently learned of a $(1+\varepsilon)$-factor algorithm by Arora and Chang [3] that has a superpolynomial running time of the form $n^{O(\log^c n)}$). We thus return to the idea of constructing a solution by traversing the MST and analyzing the weight of the solution as a factor of the MST weight. The doubling strategy still applies; in fact, it is possible, using the triangle inequality alone, to get an approximation factor of $2 - (K-2)/(K_{\max}-2)$ [9], where $K_{\max}$ is the maximum degree of an MST, thus showing that "short-cutting" does help for $K \geq 3$. The analysis is tight for arbitrary metric spaces. For the Euclidean metric in the plane, every point set already possesses an MST of maximum degree 5 [15], so this yields factors 5/3 and 4/3 for $K = 3$ and $K = 4$ respectively.

In as early as 1984, Papadimitriou and Vazirani [16] asked whether the geometry of the Euclidean case (besides the triangle inequality) can be exploited to prove better approximation factors for bounded-degree spanning trees. Khuller, Raghavachari, and Young [11] took an in-depth look into this question and managed to achieve factors 3/2 and 5/4 for $K = 3$ and $K = 4$ respectively in the plane. Since then, no improvements have been made, despite frequent references to their work [2, 4, 8, 9, 12, 17, 18].

We report the first progress in eight years: in the Euclidean plane, there always exists degree-3 and degree-4 spanning trees with weights within factors 1.402 and 1.143 respectively of the MST weight. Such trees can be constructed in polynomial (in fact, near-linear) time.

Immediately, we obtain a factor-1.402 and factor-1.143 approximation algorithm for the minimum Euclidean degree-3 and degree-4 spanning tree problem in the plane. Note that Papadimitriou and Vazirani [16] have shown the NP-hardness of the minimum Euclidean degree-3 spanning tree problem, but the status of the corresponding degree-4 problem remains open. However, regardless of algorithmic implications, our result is important in that it provides new information on a universal constant (the largest ratio of the minimum degree-3/-4 spanning tree weight to the MST weight) similar to the the Steiner ratio (the smallest ratio of the minimum Steiner tree weight to the MST weight) [7] and other constants studied in discrete geometry (such as [10]).

The new algorithms are not complicated and involve some interesting, cleverer recursive tree constructions. Their analyses, though, require more cases and demand techniques more versatile than those of Khuller et al.'s; still, with proper planning, we get proofs that are (hopefully) not too difficult to verify. We briefly review Khuller et al.'s previous algorithm in the next section and explain why 3/2 and (a value close to) 5/4 are particularly difficult barriers to break. In Sections 3 and 4 we present the new recursive algorithms and analyses.

In d-dimensional Euclidean space, the study of these worst-case ratios is even more vital, because the maximum degree of an MST can be much larger (a constant that depends exponentially on d [19]). In their paper, Khuller et al. [11] analyzed a simple algorithm and proved a remarkable 5/3 upper bound for degree-3 spanning trees in *any* number of dimensions. In Section 5, we mention how the bound can be reduced slightly to $\frac{2}{3}\sqrt{6} < 1.633$ using essentially the same algorithm.

2 Khuller, Raghavachari, and Young's approach

To facilitate comparisons, we begin with a recursive interpretation of Khuller et al.'s approach [11].

2.1 Degree-3 spanning trees

We are given an MST T of an n-point set in the plane, which we may assume [15] has maximum degree at most 5. Let $w(T)$ denote the sum of the edge weights in T. Root T at a fixed non-degree-5 vertex, say a leaf, so that each vertex has at most 4 children.

Khuller et al.'s approach can be viewed as a recursive algorithm that transforms the rooted tree T into a new degree-3 spanning tree, with the inductive hypothesis that *the root v of T should have degree 1 in the new tree.* The algorithm is simple:

> Pick a permutation $v_1, \ldots, v_k$ of the children of v in T. Recursively transform the subtrees rooted at $v_1, \ldots, v_k$. Then replace the edges $vv_1, \ldots, vv_k$ by $vv_1, v_1v_2, \ldots, v_{k-1}v_k$, as illustrated by Figure 1.

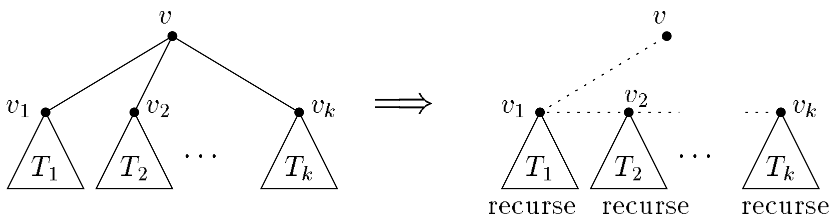


Figure 1: Khuller et al.'s degree-3 algorithm. (Pictures drawn graph-theoretically, not geometrically.)

Showing that the new spanning tree has weight at most $1.5w(T)$ amounts to showing at every step the existence of a "good" permutation $v_1, \ldots, v_k$ satisfying the inequality

$$|vv_1| + |v_1v_2| + \cdots + |v_{k-1}v_k| \;\leq\; 1.5\sum_{i=1}^{k} |vv_i|,$$

or, more informally, that the ratio of the weight of some path starting at v to the weight of a star is bounded by 1.5. Theoretically, the proof can be carried out in $O(1)$ time because $k \leq 4$ only (and the theory of the reals is decidable). Khuller et al. used a combination of nontrivial ideas in order to obtain a "presentable" proof.

2.2 Degree-4 spanning trees

Khuller et al.'s algorithm for producing a degree-4 spanning tree of weight at most $1.25w(T)$ is similar, except for a weaker inductive hypothesis: *the root v of T now should have degree at most 2 in the new tree.* Instead of adding a path that must start at v, we can now add any path visiting $v, v_1, \ldots, v_k$ to the new tree, as in Figure 2.

In the analysis, it is shown that the ratio of the weight of some such path to the weight of the star ($\sum_{i=1}^{k} |vv_i|$) is bounded by 1.25, again via a careful case study.

2.3 Limits to the approach

The 1.5 bound for degree-3 spanning trees is tight if we insist that the designated root must have degree 1 in the new tree. The example in Figure 3(a) indicates why. Even if all vertices with 2 children were to magically disappear, we still have the configurations of Figure 3(d) (which requires

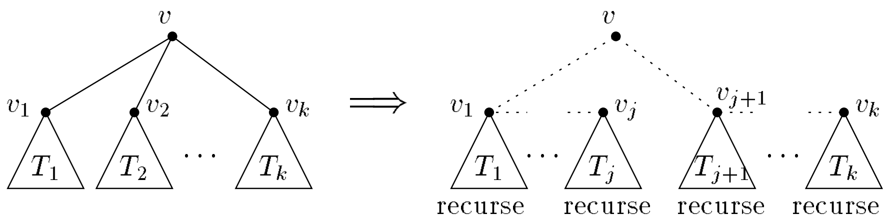


Figure 2: Khuller et al.'s degree-4 algorithm.

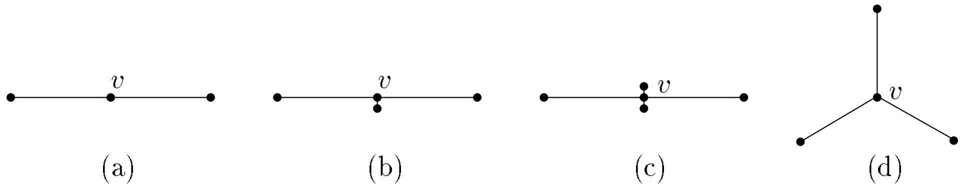


Figure 3: Bad examples.

a ratio of $\frac{1}{3}(2\sqrt{3}+1) > 1.488$) and Figure 3(b,c) (where paths starting at v require a ratio arbitrarily close to 1.5) to contend with.

Khuller et al. did not claim that their 1.25 bound is tight for degree-4 spanning trees, but even if their analysis could be refined, the improvement would be marginal, because the example in Figure 3(d) requires a ratio of $\frac{1}{3}(\sqrt{3}+2) > 1.244$, under the condition that the designated root has degree at most 2 in the new tree.

So, in order to get better results, we need to relax the inductive hypotheses. For instance, as Khuller et al. have noticed, in the recursion to the subtree T_k (but not the other subtrees) in Figure 1, we could allow the root to have degree 2 instead of 1. Similarly, in the recursion to the subtrees T_j and T_k (but not the others) in Figure 2, we could allow their roots to have degree 3 instead of 2. Unfortunately, such refinements do not necessarily translate to improved approximation factors in the worst case. Alternatively, we could try to exploit "upward" information (the parent) instead of just "downward" information (the children) at each vertex v; Khuller et al. even proposed re-rooting T at different vertices. Again, it is unclear how to get general worst-case improvements this way.

3 Better degree-4 spanning trees in the plane

Let $\tau = 1.143$ in this section.

We describe our result for degree-4 spanning trees first, as it is simpler to explain. Our approach is indeed to adopt a weaker inductive hypothesis that consistently permits the root v to have degree at most 3 instead of 2 in the new tree. To do so, we have to recurse not just on subtrees of the original MST, but on trees formed by joining two subtrees together via an "attachment" edge. The key technical idea is to strengthen the weakened hypothesis by forcing this attachment edge to value less, with factor 1 (instead of τ) in the analysis.

3.1 The new approach

To make the above idea precise, we introduce some terms. Given a tree consisting of two types of edges, *charged* and *uncharged*, define the *value* of the tree as the total weight of the charged edges plus τ times the total weight of the uncharged edges.

The input to our recursive algorithm is a tree T with root v, where *all edges are uncharged except possibly for one edge vv'*; the uncharged edges come from subtrees rooted at v and v' in the original degree-5 MST (which we assume is rooted at a fixed leaf). The output is a new tree of maximum degree 4, where *the root v has degree at most 3, all edges are charged, and the value of the tree does not increase.* This last output requirement automatically implies that the new tree weight is at most τ times the original weight.

The algorithm works basically as in Figure 4, where uncharged edges are shown in solid lines and charged edges are shown in dotted lines. (The obvious base case is omitted.) Pick a permutation $v_1, \ldots, v_{k+1}$ of the children of v in T, including v'. First convert $vv_1, \ldots, vv_{k+1}$ to charged edges (except for the already charged vv'). Then do the following:

- **Case $k \leq 2$.** Just recursively transform the subtrees rooted at $v_1, \ldots, v_{k+1}$.
- **Case $k = 3$.** Replace vv_2 by the charged edge v_1v_2, and recursively transform the subtrees rooted at v_1, v_3, v_4. Call $|v_1v_2| - |vv_2|$ the *excess* of the permutation $v_1, \ldots, v_4$.
- **Case $k = 4$.** Replace vv_2 and vv_4 by the charged edges v_1v_2 and v_3v_4, and recursively transform the subtrees rooted at v_1, v_3, v_5. Call $|v_1v_2| - |vv_2| + |v_3v_4| - |vv_4|$ the *excess* of the permutation $v_1, \ldots, v_5$.

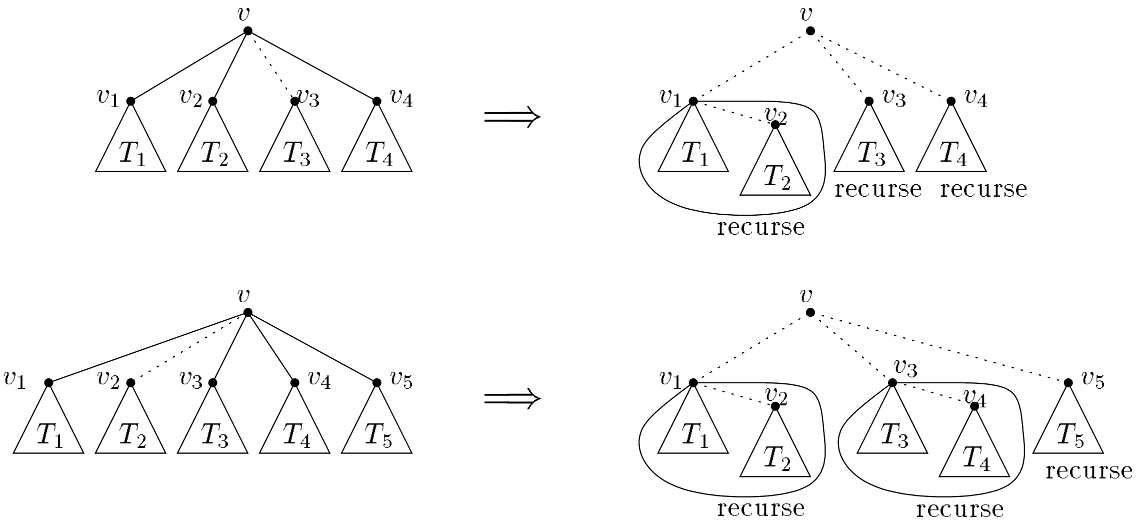


Figure 4: The new degree-4 algorithm in a nutshell.

In both the $k = 3$ and $k = 4$ cases, the change in the tree's value is equal to the excess minus $(\tau - 1)\sum_{v_i \neq v'} |vv_i|$. To ensure that the tree's value can never increase, we will choose the permutation with the smallest excess and prove that the smallest excess is at most $(\tau - 1)\sum_{v_i \neq v'} |vv_i|$.

Clearly, the algorithm runs in linear time, given the initial degree-5 MST (which can be constructed in $O(n \log n)$ time [15]).

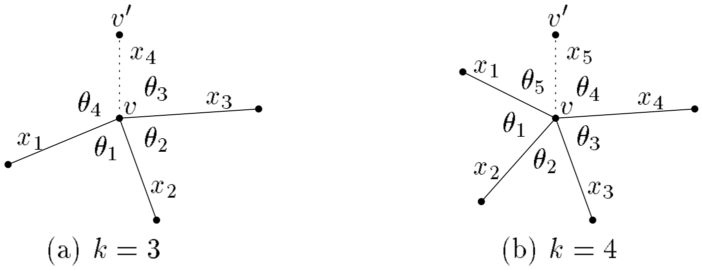


Figure 5: Notation for the degree-4 analysis.

3.2 Preliminaries for the proof

Our analysis, though somewhat lengthy (due to our desire to obtain the lowest constant), relies on very elementary tools—just an angle-sensitive version of the triangle inequality (Lemma 3.1), and a useful min trick (Lemma 3.2):

Lemma 3.1 *If a triangle has sides x, y, z with $x \le y$, and the angle opposite z is θ, then*

$$z \le f(\theta)x + y, \quad \text{where } f(\theta) := \max\{2\sin(\theta/2) - 1,\, 0\}.$$

Proof: When $x = y$, we have $z = 2\sin(\theta/2)x$. As y increases by δ while x is fixed, z can increase by at most δ. □

Lemma 3.2 *If $a_1, \ldots, a_m \ge 0$, then*

$$\min\{a_1x_1, \ldots, a_mx_m\} \le \tfrac{1}{m}\,\text{H.M.}\{a_1, \ldots, a_m\}\,(x_1 + \cdots + x_m),$$

where H.M. *denotes the Harmonic mean.*

Proof: Just take a convex combination: $\min\{a_1x_1, \ldots, a_mx_m\} \le \alpha_1a_1x_1 + \cdots + \alpha_ma_mx_m$, with $\alpha_i = (1/a_i)/[1/a_1 + \cdots + 1/a_m]$. □

3.3 The analysis

- **Case $k = 3$.** Say v_a, v_b, v_c, v' are the children of v in T sorted by angle. Let $x_1 = |vv_a|$, $x_2 = |vv_b|$, $x_3 = |vv_c|$, $x_4 = |vv'|$, $\theta_1 = \angle v_avv_b$, $\theta_2 = \angle v_bvv_c$, $\theta_3 = \angle v_cvv'$, and $\theta_4 = \angle v'vv_a$, as in Figure 5(a).

  We want to show that some permutation has excess less than $(\tau - 1)(x_1 + x_2 + x_3)$. By Lemma 3.1, we have permutations with excesses bounded by:

  $$f(\theta_1)\min\{x_1, x_2\}, \quad f(\theta_2)\min\{x_2, x_3\}, \quad f(\theta_3)\min\{x_3, x_4\}, \quad f(\theta_4)\min\{x_4, x_1\}.$$

  Thus, the minimum excess is at most the minimum of the following values:

  $$\min\{f(\theta_4), f(\theta_1)\}x_1, \quad \min\{f(\theta_1), f(\theta_2)\}x_2, \quad \min\{f(\theta_2), f(\theta_3)\}x_3.$$

  By Lemma 3.2, the minimum excess is at most

  $$\tfrac{1}{3}\,\text{H.M.}\{\min\{f(\theta_4), f(\theta_1)\}, \min\{f(\theta_1), f(\theta_2)\}, \min\{f(\theta_2), f(\theta_3)\}\}\,(x_1 + x_2 + x_3).$$

Since $\min\{\theta_4, \theta_1\} + \min\{\theta_1, \theta_2\} + \min\{\theta_2, \theta_3\} \le 270°$, it can be verified (as the H.M. is no more than the arithmetic mean and f is concave on the interval of interest) that the above coefficient is bounded by $\frac{1}{3} f(90°) = (\sqrt{2} - 1)/3 < 0.139$.

- **Case** $k = 4$. Say v_a, v_b, v_c, v_d, v' are the children of v in T sorted by angle. Let $x_1 = |vv_a|$, $x_2 = |vv_b|$, $x_3 = |vv_c|$, $x_4 = |vv_d|$, $x_5 = |vv'|$, $\theta_1 = \angle v_a v v_b$, $\theta_2 = \angle v_b v v_c$, $\theta_3 = \angle v_c v v_d$, $\theta_4 = \angle v_d v v'$, and $\theta_5 = \angle v' v v_a$, as in Figure 5(b). Because angles between two adjacent MST edges must exceed $60°$ (e.g., see [15]), we know that $\theta_1, \theta_2, \theta_3, \theta_4 + \theta_5 \ge 60°$. Furthermore, because one of the MST edge at v (the parent) is not present in T, we also know the following fact, which will be helpful later (though not necessary to get a new result):

$$\max\{\theta_1, \theta_2, \theta_3, \theta_4 + \theta_5\} \ge 120°. \tag{1}$$

We want to show that some permutation has excess smaller than $(\tau - 1)(x_1 + x_2 + x_3 + x_4)$. By Lemma 3.1, we have permutations with excesses bounded by

$$\begin{aligned}
f(\theta_1) \min\{x_1, x_2\} &+ f(\theta_3) \min\{x_3, x_4\}, \\
f(\theta_2) \min\{x_2, x_3\} &+ f(\theta_4) \min\{x_4, x_5\}, \\
f(\theta_3) \min\{x_3, x_4\} &+ f(\theta_5) \min\{x_5, x_1\}, \\
f(\theta_4) \min\{x_4, x_5\} &+ f(\theta_1) \min\{x_1, x_2\}, \\
f(\theta_5) \min\{x_5, x_1\} &+ f(\theta_2) \min\{x_2, x_3\}, \\
f(\theta_3 + \theta_4) \min\{x_3, x_5\} &+ f(\theta_4 + \theta_5) \min\{x_4, x_1\}, \\
f(\theta_4 + \theta_5) \min\{x_4, x_1\} &+ f(\theta_5 + \theta_1) \min\{x_5, x_2\}.
\end{aligned} \tag{2}$$

Consider three subcases (the missing one is symmetric):

  - Subcase $\theta_4, \theta_5 \le 60°$. Then $f(\theta_4) = f(\theta_5) = 0$ and (2) gives the following values:

    $$f(\theta_2) \min\{x_2, x_3\}, \quad f(\theta_3) \min\{x_3, x_4\}, \quad f(\theta_1) \min\{x_1, x_2\}, \quad f(\theta_2) \min\{x_2, x_3\}.$$

    By Lemma 3.2, the minimum excess is at most

    $$\tfrac{1}{4}\,\text{H.M.}\{f(\theta_1),\ \min\{f(\theta_1), f(\theta_2)\},\ \min\{f(\theta_2), f(\theta_3)\},\ f(\theta_3)\}\,(x_1 + x_2 + x_3 + x_4).$$

    Since $\theta_1 + \theta_2 + \theta_3 \le 300°$ implies $\theta_1 + \min\{\theta_1, \theta_2\} + \min\{\theta_2, \theta_3\} + \theta_3 \le 400°$, it can be verified that the above coefficient is bounded by $\frac{1}{4} f(100°) < 0.134$.

  - Subcase $\theta_4 \ge 60°$, $\theta_5 \le 60°$. Here, $f(\theta_5) = 0$ and (2) implies the following upper bounds on the minimum excess:

    $$\begin{aligned}
    f(\theta_3) \min\{x_3, x_4\}, \quad f(\theta_2) \min\{x_2, x_3\}, \quad &\max\{f(\theta_4), f(\theta_1)\}(x_4 + x_1), \\
    &\max\{f(\theta_4 + \theta_5), f(\theta_5 + \theta_1)\}(x_1 + x_2).
    \end{aligned}$$

    By (1), we have the following possibilities:

    * *Subsubcase* $\theta_1 \ge 120°$ *or* $\theta_4 + \theta_5 \ge 120°$. By Lemma 3.2, the minimum excess is at most

      $$\tfrac{1}{3}\,\text{H.M.}\{f(\theta_2), f(\theta_2), f(\theta_3)\}\,(x_2 + x_3 + x_4).$$

      Since $\theta_2 + \theta_3 \le 360° - 120° - 60° = 180°$, it can be confirmed numerically that the above coefficient is maximized near $\theta_2 \approx 95°$, $\theta_3 \approx 85°$ and is below 0.142.

* *Subsubcase* $\theta_2 \geq 120°$. By Lemma 3.2, the minimum excess is at most

$$\tfrac{1}{3}\,\text{H.M.}\{\max\{f(\theta_4+\theta_5), f(\theta_5+\theta_1)\},\ f(\theta_3),\ f(\theta_3)\}\,[(x_1+x_2)+x_3+x_4].$$

  Since $\max\{\theta_4+\theta_5, \theta_5+\theta_1\}+\theta_3 \leq 360° - 120° - 60° = 180°$, the above coefficient is again bounded by 0.142.

* *Subsubcase* $\theta_3 \geq 120°$. By Lemma 3.2, the minimum excess is at most

$$\tfrac{1}{3}\,\text{H.M.}\{\max\{f(\theta_4), f(\theta_1)\},\ f(\theta_2),\ f(\theta_2)\}\,[(x_4+x_1)+x_2+x_3].$$

  Since $\max\{\theta_4, \theta_1\}+\theta_2 \leq 360° - 120° - 60° = 180°$, the above coefficient is also bounded by 0.142.

– Subcase $\theta_4, \theta_5 \geq 60°$. Consider which of the angles $\theta_1, \ldots, \theta_5$ is the largest (the missing subsubcases are symmetric):

* *Subsubcase: θ_1 is the largest.* Now, (2) gives the following upper bounds:

$$\max\{f(\theta_3), f(\theta_5)\}(x_3+x_1), \quad \max\{f(\theta_2), f(\theta_4)\}(x_2+x_4).$$

  By Lemma 3.2, the minimum excess is at most

$$\tfrac{1}{2}\,\text{H.M.}\{\max\{f(\theta_3), f(\theta_5)\}, \max\{f(\theta_2), f(\theta_4)\}\}\,[(x_3+x_1)+(x_2+x_4)].$$

  Since $\max\{\theta_3, \theta_5\}+\max\{\theta_2, \theta_4\} \leq \min\{2\theta_1,\ 360° - 60° - 60° - \theta_1\} \leq 160°$, it can be verified that the above coefficient is bounded by $\frac{1}{2}f(80°) < 0.143$.

* *Subsubcase: θ_5 is the largest.* Similarly, (2) and Lemma 3.2 imply that the minimum excess is at most

$$\tfrac{1}{2}\,\text{H.M.}\{\max\{f(\theta_1), f(\theta_3)\}, \max\{f(\theta_2), f(\theta_4)\}\}\,[(x_1+x_3)+(x_2+x_4)].$$

  By a similar argument, the coefficient is again bounded by 0.143.

* *Subsubcase: θ_2 is the largest.* By (2) and Lemma 3.2, the minimum excess is at most

$$\tfrac{1}{2}\,\text{H.M.}\{\max\{f(\theta_3), f(\theta_5)\}, \max\{f(\theta_4), f(\theta_1)\}\}\,[(x_3+x_1)+(x_4+x_2)].$$

  The coefficient is again bounded by 0.143.

A ratio of 1.143 has thus been established.

Remarks: More precisely, we can replace 1.143 by $1+\frac{1}{2}f(80°) = \sin 40° + \frac{1}{2} = 1.142787\ldots$

There might be room for improvement in the last subcase, by a more detailed case analysis, or by not bounding distances linearly with Lemma 3.1 (which is tight only when $x = y$). However, the room would be small, considering that our analysis for the $k = 3$ case (with ratio > 1.138) is tight under our inductive hypothesis.

4 Better degree-3 spanning trees in the plane

Let $\tau = 1.402$ in this section.

Logically, our approach for degree-3 spanning trees should adopt a similar relaxed condition where the root v is permitted to have degree 2 in the new tree, instead of 1 as in Khuller et al.'s algorithm. More obstacles have to be overcome now, and it is less apparent what inductive hypothesis we should use to keep the analysis manageable. Our idea is to force not just one attachment edge, but a path of attachment edges, to value less, with factor 1 (instead of τ) in the analysis. Additionally, a technical complication arises because of the need to recurse on general trees, not necessarily formed by joining a constant number of MST subtrees.

4.1 The new approach

The input to our recursive algorithm is a tree T with root v, where *all edges are uncharged except possibly for edges along a path from the root* $vv'v'' \cdots$. The output is a new tree of maximum degree 3, where *the root* v *has degree at most 2, all edges are charged, and the value of the tree does not increase.* (The tree's value is defined in the same way as in Section 3.1, except for the different τ.)

The algorithm works according to one of several schemes depicted in Figure 6 (where charged edges are again shown in dotted lines). Pick a permutation $v_1, \ldots, v_k$ of the children of v in T, excluding v'.

- If $|v_{k-1}v_k| \leq |vv_k|$, then apply Scheme A: replace vv_k by the uncharged edge $v_{k-1}v_k$ to lower the root's degree, and repeat.
- If $\min\{|vv_k| + |v_kv'| - |vv'|,\ \tau|v_kv'|\} \leq \tau|vv_k|$, then apply Scheme B: to lower the root's degree, remove vv_k, and if $|vv_k| + |v_kv'| - |vv'| < \tau|v_kv'|$, replace vv' by the charged edges vv_k and v_kv', otherwise insert the uncharged edge $v'v_k$; then repeat. (Notice how this preserves the input requirement that charged edges form a path from the root.)
- If the above schemes are not applicable for any permutation, we consider two recursive schemes. In Scheme C, remove $vv_1, \ldots, vv_k$, insert a path of charged edges $vv_1v_2 \cdots v_k$, and recursively transform the subtrees rooted at v_1 and v'. Define

$$\text{cost} := |vv_1| + |v_1v_2| + \cdots + |v_{k-1}v_k|.$$

- In Scheme D, remove $vv_1, vv_2, \ldots, vv_k$, insert a path of charged edges $vv_1v_2 \cdots v_{k-1}$, and if $|vv_k| + |v_kv'| - |vv'| < \tau|v_kv'|$, replace vv' by the charged edges vv_k and v_kv', otherwise insert the uncharged edge $v'v_k$. Then recursively transform the subtrees rooted at the two children of v. (Again notice how the input requirement is preserved at the two subtrees.) Define

$$\text{cost} := |vv_1| + |v_1v_2| + \cdots + |v_{k-2}v_{k-1}| + \min\{|vv_k| + |v_kv'| - |vv'|,\ \tau|v_kv'|\}.$$

In Schemes C and D, the change in the tree's value is equal to the cost minus $\tau\sum_{i=1}^{k}|vv_i|$. To ensure that the tree's value can never increase, we will choose a permutation and a scheme with the smallest cost and prove that the smallest cost is at most $\tau\sum_{i=1}^{k}|vv_i|$ whenever Schemes A and B are not applicable.

Note that if Schemes A and B are not applicable, then $k \leq 4$ (see Section 4.3). Testing/handling Schemes A and B requires time proportional to v's degree, which is at most n, and so the algorithm

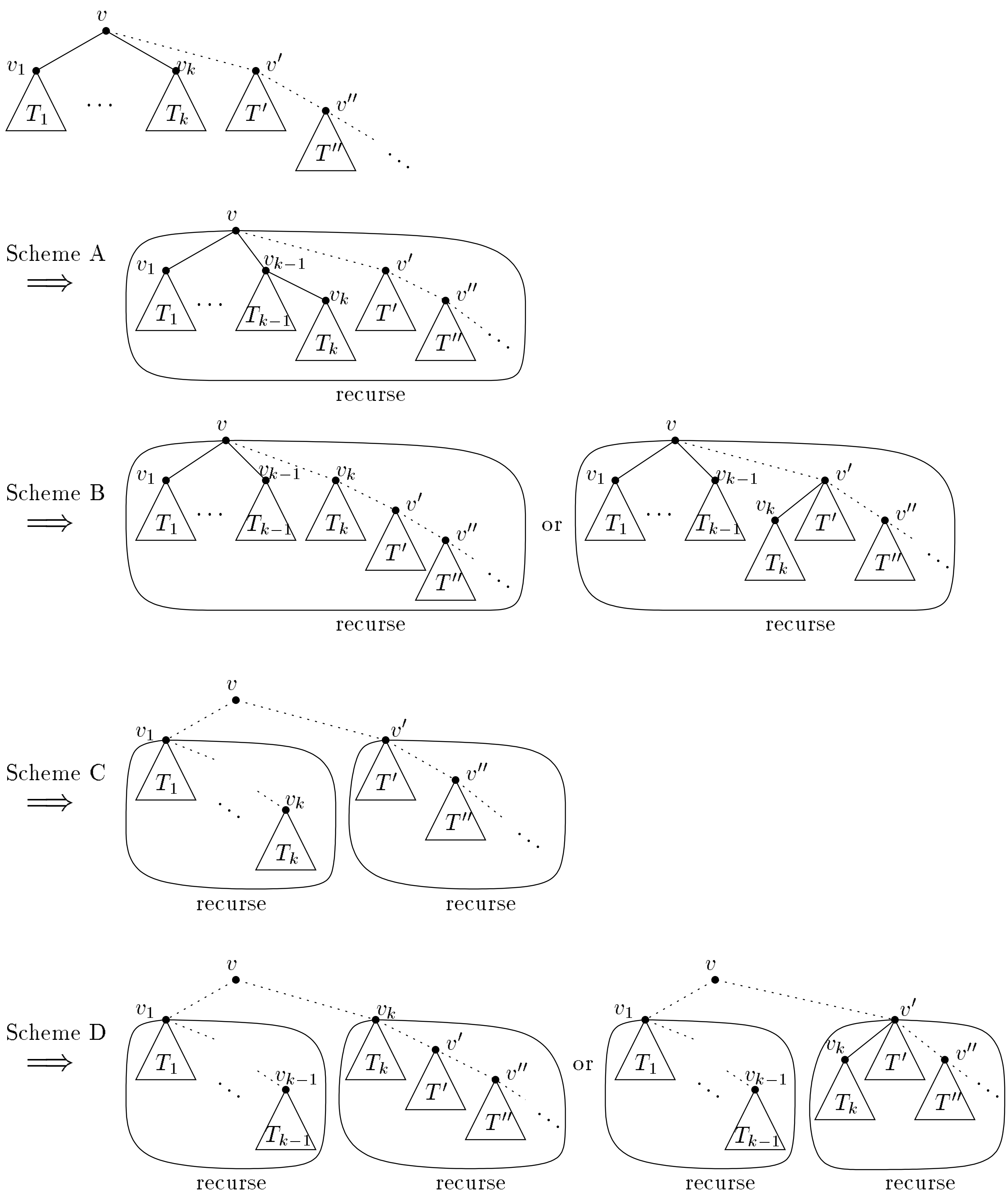


Figure 6: The new degree-3 algorithm in picture form.

can be implemented in $O(n^2)$ time. In practice, the degrees probably do not grow by much, and the running time should be closer to linear.

In fact, a near-linear worst-case time bound can be guaranteed by a variant of the algorithm. In this variant, we insist that all nodes in T have degree at most 7. When a degree-8 node w is created (because the linkage of v_k to v_{k-1} in Scheme A or the linkage of v_k to v' in Scheme B or D), we "fix" w as follows. Let $w_1, \ldots, w_6$ be the children of w, sorted by angle, excluding the child with the largest subtree size and the child w' along the path of charged edges. (Subtree sizes can be maintained easily at all nodes.) One of these six angles is at most 60°; without loss of generality, say this angle is $\angle w_1 w w_2$, with $|ww_1| \le |ww_2|$. Replace the uncharged edge ww_2 with the uncharged edge $w_1 w_2$. Recursively fix w_1 if it now has degree 8. Since the subtree size at w_1 is at most 7/8 times the subtree size at w, the running time of this vertex-fixing operation is $O(\log n)$. Because we have maintained that $k \le 6$, the algorithm requires $O(n)$ operations and a total of $O(n \log n)$ time.

4.2 Preliminaries for the proof

The analysis of the degree-3 algorithm requires a different set of tools. At some point we make use of an alternative triangle inequality (Lemma 4.1), which refines Lemma 3.1 when $\theta < 90°$, provided that z is the largest side. Lemma 4.2 deals with an expression on triangle sides that occurs already in our definition of cost. Finally, Lemma 4.3 comes in handy in bounding linear expressions.

Lemma 4.1 *If a triangle has sides x, y, z with $x \le y \le z$, and the angle opposite z is θ, then*

$$z \le G(\theta)x + (F(\theta) - G(\theta))y, \quad \text{where } F(\theta) := 2\sin(\theta/2), \;\; G(\theta) := 1/(F(\theta)+1).$$

Proof: $z \ge y$ implies that $x \ge 2y\cos\theta$. The above is an equality when $x = 2y\cos\theta$, because of the identity $2G(\theta)\cos\theta + F(\theta) - G(\theta) = 1$. We also have equality when $x = y$. Since z is a convex function of x for a fixed y, the inequality holds for all x between $2y\cos\theta$ and y. □

Lemma 4.2 *If a triangle has sides x, y, z, and the angle opposite z is θ, then*

$$\min\{x + z - y, \, \tau z\} \le H(\theta)x,$$

where

$$H(\theta) := \max\left\{\frac{\tau}{\tau-1}\left[1 - J(\theta) + \sqrt{J(\theta)^2 - 1}\right], \; \tau\right\}, \qquad J(\theta) := \frac{1 - (\tau-1)^2\cos\theta}{1 - (\tau-1)^2}.$$

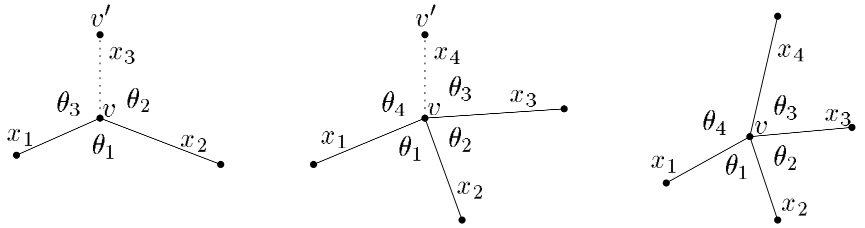
Proof: W.l.o.g., say $x = 1$. We want to upper-bound the following expression by $H(\theta)$:

$$\min\left\{1 + \sqrt{y^2 + 1 - 2y\cos\theta} - y, \; \tau\sqrt{y^2 + 1 - 2y\cos\theta}\right\}. \tag{3}$$

The expression is decreasing for a sufficiently large y. Since both min terms are convex, the maximum of (3) occurs either when $y = 0$ (with value τ), or when the two min terms coincide, i.e., when

$$y^2 + 1 - 2y\cos\theta = (1-y)^2/(\tau-1)^2 \qquad (y \le 1).$$

This quadratic equation can be rewritten as $y^2 - 2J(\theta)y + 1 = 0$, which has solution $y = J(\theta) - \sqrt{J(\theta)^2 - 1}$. At this y, (3) evaluates to $[\tau/(\tau-1)](1-y)$, as desired. □



(a) $k = 2$ (b) $k = 3$ (c) $k = 4$

Figure 7: Notation for the degree-3 analysis.

Lemma 4.3 *If* $0 \le x_1 \le x_2 \le \cdots \le x_m$, *then*

$$a_1 x_1 + \cdots + a_m x_m \;\le\; \max\left\{a_m,\, \tfrac{1}{2}(a_{m-1} + a_m),\, \ldots,\, \tfrac{1}{m}(a_1 + \cdots + a_m)\right\}(x_1 + \cdots + x_m).$$

Proof: The inequality trivially holds for $x_1 = x_2 = \cdots = x_m = 0$. When $x_{m-j+1}, \ldots, x_m$ increase simultaneously by a common amount δ for a given j, the L.H.S. increases by $\delta(a_{m-j+1} + \cdots + a_m)$, while the R.H.S. increases by at least $\frac{1}{j}(a_{m-j+1} + \cdots + a_m)\delta j$. Therefore, the inequality remains true as long as $0 \le x_1 \le \cdots \le x_m$. □

4.3 The analysis

Suppose that Scheme A is not applicable for any permutation. Then $|v_i v_j| > |vv_i|, |vv_j|$ for all i, j. In particular, this implies that each angle $\angle v_i v v_j$ must exceed $60°$. In addition, it validates the subsequent applications of Lemma 4.1.

Suppose further that Scheme B is not applicable. This implies that the angle $\angle v_i v v'$ for each i must exceed $72°$, because of Lemma 4.2, since $H(72°) = \tau$.

Thus, $k \le 4$. The $k = 1$ case is trivial. We consider the $k = 2$, $k = 3$, and $k = 4$ cases separately. Before plunging into the details, some words of caution: The present proof is less elegant than the previous proof, because the algorithm is now more involved and there is less symmetry. On the other hand, when required, we only take the simplest convex combinations, i.e., averages, in contrast to the fancier ones used in Lemma 3.2. The $k = 2$ case turns out to be the "critical" case, and so we can afford to be looser in our estimates for $k = 3$ and $k = 4$—otherwise, a complete analysis, especially for the $k = 4$ case, would be even more daunting.

- **Case** $k = 2$. Say v_a, v_b, v' are the children of v in T. Let $x_1 = |vv_a|$, $x_2 = |vv_b|$, $\theta_1 = \angle v_a v v_b$, $\theta_2 = v_b v v'$, and $\theta_3 = v' v v_a$, as in Figure 7(a). W.l.o.g., say $x_1 \le x_2$.

  We want to show that Scheme C or D under some permutation has cost at most $\tau(x_1 + x_2)$.

  - Subcase $\theta_1 \le 128.6°$. By Lemma 3.1, Scheme C yields cost bounded by

    $$F(\theta_1)x_1 + x_2 \;\le\; \max\left\{1, \tfrac{1}{2}(F(\theta_1) + 1)\right\}(x_1 + x_2),$$

    using Lemma 4.3. The above coefficient is at most $\frac{1}{2}(F(128.6°) + 1) < 1.402$.
  - Subcase $\theta_3 \le 115.7°$. By Lemma 4.2, Scheme D yields cost bounded by

    $$x_2 + H(\theta_3)x_1 \;\le\; \max\left\{1, \tfrac{1}{2}(H(\theta_3) + 1)\right\}(x_1 + x_2),$$

    using Lemma 4.3. The above coefficient is at most $\frac{1}{2}(H(115.7°) + 1) < 1.402$.

– Subcase $\theta_1 \geq 128.6°$, $\theta_3 \geq 115.7°$. By Lemma 4.2, Scheme D yields costs at most

$$x_2 + H(\theta_3)x_1, \quad x_1 + H(\theta_2)x_2.$$

The minimum is upper-bounded by the average $\frac{1}{2}(H(\theta_3)+1)x_1 + \frac{1}{2}(H(\theta_2)+1)x_2$, which by Lemma 4.3 is at most

$$\max\left\{\tfrac{1}{2}(H(\theta_2)+1),\ \tfrac{1}{4}(H(\theta_2)+H(\theta_3)+2)\right\}(x_1+x_2).$$

Because $\theta_2 \leq 115.7°$ and $\theta_2+\theta_3 \leq 231.4°$, both max terms are again at most $\frac{1}{2}(H(115.7°)+1) < 1.402$.

- **Case** $k = 3$. This case is the lengthiest.

  Say v_a, v_b, v_c, v' are the children of v in T sorted by angle. Let $x_1 = |vv_a|$, $x_2 = |vv_b|$, $x_3 = |vv_c|$, $\theta_1 = \angle v_a v v_c$, $\theta_2 = \angle v_b v v_c$, $\theta_3 = \angle v_c v v'$, and $\theta_4 = \angle v' v v_a$, as in Figure 7(b). W.l.o.g., say $x_1 \leq x_3$.

  We want to show that Scheme C or D under some permutation has cost at most $\tau(x_1+x_2+x_3)$.

  – Subcase $x_1 \leq x_3 \leq x_2$. By Lemmas 3.1 and 4.2, Scheme D yields costs at most

  $$F(\theta_1)x_1 + x_2 + H(\theta_3)x_3, \quad F(\theta_2)x_3 + x_2 + H(\theta_4)x_1.$$

  Taking the average and applying Lemma 4.3 bound the minimum cost by

  $$\max\left\{1,\ \tfrac{1}{4}(F(\theta_2)+H(\theta_3)+2),\ \tfrac{1}{6}(F(\theta_1)+F(\theta_2)+H(\theta_3)+H(\theta_4)+2)\right\}(x_1+x_3+x_2).$$

  Since $\theta_2+\theta_3 \leq 360° - 60° - 72° = 228°$, it can be confirmed numerically that the second max term is maximized near $\theta_2 \approx 125°$, $\theta_3 \approx 103°$ and has value below 1.372. Since $\theta_1+\theta_2+\theta_3+\theta_4 = 360°$, it can confirmed numerically that the third max term is maximized near $\theta_1, \theta_2 \approx 100°$, $\theta_3, \theta_4 \approx 80°$ and has value below 1.341.

  – Subcase $x_1 \leq x_2 \leq x_3$.

    * *Subsubcase* $\theta_3 \leq 115°$. By Lemmas 3.1 and 4.2, Scheme D yields costs at most

    $$F(\theta_1)x_1 + x_2 + H(\theta_3)x_3, \quad F(\theta_2)x_2 + x_3 + H(\theta_4)x_1.$$

    Taking the average and applying Lemma 4.3 bound the minimum cost by

    $$\max\Big\{\tfrac{1}{2}(H(\theta_3)+1),\ \tfrac{1}{4}(F(\theta_2)+H(\theta_3)+2),$$
    $$\tfrac{1}{6}(F(\theta_1)+F(\theta_2)+H(\theta_3)+H(\theta_4)+2)\Big\}(x_1+x_2+x_3).$$

    The first max term is at most $\frac{1}{2}(H(115°)+1) < 1.399$. As in the previous subcase, the second and third max terms are bounded by 1.372 and 1.341.

    * *Subsubcase* $\theta_3 \geq 115°$. By Lemma 3.1, Scheme C yields cost at most

    $$F(\theta_1)x_1+F(\theta_2)x_2+x_3 \ \leq\ \max\left\{1,\ \tfrac{1}{2}(F(\theta_2)+1),\ \tfrac{1}{3}(F(\theta_1)+F(\theta_2)+1)\right\}(x_1+x_2+x_3),$$

    using Lemma 4.3. Since $\theta_2 \leq 360° - 115° - 72° - 60° = 113°$, the second max term is at most $\frac{1}{2}(F(113°)+1) < 1.334$. Since $\theta_1+\theta_2 \leq 360° - 115° - 72° = 173°$, the third max term is at most $\frac{1}{3}(2F(86.5°)+1) < 1.247$.

- SUBCASE $x_2 \le x_1 \le x_3$.
  * *Subsubcase* $\theta_3 \le 115°$, $\theta_3+\theta_4 \le 195°$. By Lemmas 3.1 and 4.2, Scheme D yields costs at most
  $$F(\theta_1)x_2 + x_1 + H(\theta_3)x_3, \quad F(\theta_2)x_2 + x_3 + H(\theta_4)x_1.$$
  Taking the average and applying Lemma 4.3 bound the minimum cost by
  $$\begin{aligned} \max\Big\{&\tfrac{1}{2}(H(\theta_3)+1),\ \tfrac{1}{4}(H(\theta_3)+H(\theta_4)+2),\\ &\tfrac{1}{6}(F(\theta_1)+F(\theta_2)+H(\theta_3)+H(\theta_4)+2)\Big\}\,(x_2+x_1+x_3).\end{aligned}$$
  The first max term is at most $\frac{1}{2}(H(115°)+1) < 1.399$. Since $\theta_3+\theta_4 \le 195°$, it can be verified that the second max term is bounded by 1.333. As in the earlier subcases, the third max term is bounded by 1.341.
  * *Subsubcase* $\theta_3 \ge 115°$, $\theta_3+\theta_4 \le 195°$. Here, just take the upper bound
  $$F(\theta_2)x_2+x_3+H(\theta_4)x_1 \ \le\ \max\Big\{1,\ \tfrac{1}{2}(H(\theta_4)+1),\ \tfrac{1}{3}(F(\theta_2)+H(\theta_4)+1)\Big\}\,(x_2+x_1+x_3),$$
  using Lemma 4.3. Since $\theta_4 \le 80°$, the second max term is at most $\frac{1}{2}(H(80°)+1) < 1.245$. Since $\theta_2 \le 360° - 115° - 72° - 60° = 113°$, the third max term is at most $\frac{1}{3}(F(113°)+H(80°)+1) < 1.386$.
  * *Subsubcase* $\theta_3+\theta_4 \ge 195°$. By Lemma 4.1, Scheme C yields cost at most
  $$\begin{aligned} &x_1 + [G(\theta_1)x_2 + (F(\theta_1)-G(\theta_1))x_1] + [G(\theta_2)x_2 + (F(\theta_2)-G(\theta_2))x_3]\\ \le\ &\max\Big\{F(\theta_2)-G(\theta_2),\ \tfrac{1}{2}(F(\theta_1)-G(\theta_1)+F(\theta_2)-G(\theta_2)+1),\\ &\tfrac{1}{3}(F(\theta_1)+F(\theta_2)+1)\Big\}\,(x_2+x_1+x_3),\end{aligned}$$
  using Lemma 4.3. Since $\theta_2 \le 360° - 195° - 60° = 105°$, the first max term is at most $F(105°) - G(105°) < 1.201$. Since $\theta_1+\theta_2 \le 360° - 195° = 165°$, it can be verified that the second max term is at most $F(82.5°) - G(82.5°) + \frac{1}{2} < 1.388$ and the third max term is at most $\frac{1}{3}(2F(82.5°)+1) < 1.213$.

- **Case** $k = 4$. Having experienced the $k = 3$ case, we are happy to report, counter to intuition, that $k = 4$ case can be disposed of more quickly. As it turns out, Scheme C alone is enough to provide the desired bound here.

  Let v_a, v_b, v_c, v_d be the children of v in T sorted by angle, excluding v'. Let $x_1 = |vv_a|$, $x_2 = |vv_b|$, $x_3 = |vv_c|$, $x_4 = |vv_d|$, $\theta_1 = \angle v_a v v_b$, $\theta_2 = \angle v_b v v_c$, $\theta_3 = \angle v_c v v_d$, and $\theta_4 = \angle v_d v v_a$, as in Figure 7(c). W.l.o.g., say $x_4 \ge x_1, x_2, x_3$.

  We know that one of these four angles, depending on the placement of v' around v, exceeds $2(72°) = 144°$. W.l.o.g., say $\max\{\theta_2, \theta_4\} \ge 144°$. Note that $\theta_1, \theta_3 \le 360° - 144° - 60° - 60° = 96°$, and so $G(\theta_1), G(\theta_3) > 0.402$.

  We want to show that Scheme C under some permutation has cost at most $\tau(x_1+x_2+x_3+x_4)$.

  - SUBCASE $x_1 \le x_2$. By Lemmas 3.1 and 4.1, we get cost bounded by
  $$\begin{aligned} &x_1 + [0.4x_1 + (F(\theta_1)-0.4)x_2] + [x_2+x_3] + [0.4x_3 + (F(\theta_3)-0.4)x_4]\\ \le\ &1.4(x_1+x_3)\ +\ \max\Big\{F(\theta_3)-0.4,\ \tfrac{1}{2}(F(\theta_1)+F(\theta_3)+0.2)\Big\}\,(x_2+x_4),\end{aligned}$$

using Lemma 4.3. The first max term is at most $F(96°) - 0.4 < 1.087$. Since $\theta_1 + \theta_3 \le 360° - 144° - 60° = 156°$, the second max term is at most $F(78°) + 0.1 < 1.359$.

– SUBCASE $x_2 \le x_1$. By Lemmas 3.1 and 4.1, we get cost bounded by

$$\begin{aligned} & x_1 + [0.4x_2 + (F(\theta_1) - 0.4)x_1] + [x_2 + x_3] + [0.4x_3 + (F(\theta_3) - 0.4)x_4] \\ \le\ & 1.4(x_2 + x_3) \ +\ \max\left\{F(\theta_3) - 0.4,\ \tfrac{1}{2}(F(\theta_1) + F(\theta_3) + 0.2)\right\}(x_1 + x_4), \end{aligned}$$

using Lemma 4.3. Again, the max terms are bounded by 1.087 and 1.359.

A ratio of 1.402 has thus been established.

Remarks: The constant is tight in following sense: if $\tau = 1.401$ instead, there exists placements of vertices v, v_1, v_2, v' for $k = 2$, such that neither Schemes C nor D can yield cost less than $\tau(|vv_1| + |vv_2|)$, under the present definition of cost.

More precisely, we can replace 1.402 by a root τ of the simultaneous equations $\tau = \frac{1}{2}(F(\theta_1)+1) = \frac{1}{2}(H(\theta_3)+1)$ and $\theta_1 + 2\theta_3 = 360°$, by inspection of our $k = 2$ proof. Surprisingly, this $\tau = 1.401038\ldots$ is algebraic: calculations reveal that it satisfies the quartic equation

$$8\tau^4 - 44\tau^3 + 85\tau^2 - 64\tau + 13 = 0.$$

5 Better degree-3 spanning trees in arbitrary dimensions

We close with a preliminary discussion on higher-dimensional spanning trees. Khuller et al. [11] showed that a nontrivial result can be obtained already for the degree-3 case: the same algorithm described in Section 2 always produces a spanning tree of weight $(5/3)w(T)$. This is a consequence of the following fact: given arbitrary points $v, v_1, \ldots, v_k \in \mathbb{R}^d$, there exists a path that starts at v and visits $v_1, \ldots, v_k$ in some order, with weight at most $\frac{5}{3}\sum_{i=1}^{k}|vv_i|$. We observe that by a slightly more careful analysis, the constant 5/3 can be reduced to $\frac{2}{3}\sqrt{6} < 1.633$.

The proof is based on Khuller et al.'s but uses a different geometric inequality:

Lemma 5.1 *Given points $v, v_0, v_1, v_2, v_3, v_4 \in \mathbb{R}^d$ such that $|vv_0| \le |vv_1| \le \cdots \le |vv_4|$,*

$$\begin{aligned} \min\{ & |v_0v_1| + |v_1v_2| + |v_2v_3| + |v_3v_4|, \\ & |v_0v_2| + |v_2v_3| + |v_3v_1| + |v_1v_4|, \\ & |v_0v_3| + |v_3v_1| + |v_1v_2| + |v_2v_4|\} \ \le\ \tfrac{2}{3}\sqrt{6}\,(|vv_1| + |vv_2| + |vv_3| + |vv_4|). \end{aligned}$$

Proof: By bounding the minimum with the average, it suffices to prove that

$$\begin{aligned} & |v_0v_1| + |v_0v_2| + |v_0v_3| + |v_4v_1| + |v_4v_2| + |v_4v_3| + 2|v_1v_2| + 2|v_2v_3| + 2|v_3v_1| \\ & \qquad \le\ 2\sqrt{6}\,(|vv_1| + |vv_2| + |vv_3| + |vv_4|). \end{aligned} \tag{4}$$

First consider the case when $|vv_0| \le |vv_1| = |vv_2| = |vv_3| = |vv_4| = r$. The expression

$$|v_0v_1| + |v_0v_2| + |v_0v_3| + |v_1v_2| + |v_2v_3| + |v_3v_1|$$

is bounded by $4\sqrt{6}r$ (or, in geometric terms, the total edge length of a tetrahedron inside a fixed sphere is maximized when the tetrahedron is regular); Khuller et al. cited Lillington [13], although it is not hard to establish this fact algebraically, by the Cauchy-Schwarz inequality. Similarly, $|v_4v_1| + |v_4v_2| + |v_4v_3| + |v_1v_2| + |v_2v_3| + |v_3v_1| \le 4\sqrt{6}r$. Therefore, (4) is true, since the L.H.S. is at most $8\sqrt{6}r$ and the R.H.S. is exactly $8\sqrt{6}r$.

Now, suppose v_2, v_3, v_4 are moved radially outward so that $|vv_2|, |vv_3|, |vv_4|$ all increase by αr. Then the L.H.S. of (4) increases by at most $7\alpha r + \alpha|v_4v_2| + \alpha|v_4v_3| + 2\alpha|v_2v_3|$; here, the $|v_iv_j|$'s refer to the old distances. The expression

$$|v_2v_3| + |v_3v_4| + |v_4v_2|$$

is bounded by $3\sqrt{3}r$ (or, in geometric terms, the perimeter of a triangle inside a fixed circle is maximized when the triangle is equilateral). Therefore, the amount of change to the L.H.S. is bounded by $(9 + 3\sqrt{3})\alpha r \le 14.2\alpha r$, but the amount of change to the R.H.S. is $6\sqrt{6}\alpha r > 14.6\alpha r$. So, (4) still holds.

Now, suppose $|vv_3|, |vv_4|$ increase further by δ. Then the L.H.S. of (4) increases by at most 9δ, while the R.H.S. increases by $4\sqrt{6}\delta > 9.7\delta$. So, (4) is still true.

Finally, suppose $|vv_4|$ increases further yet by ε. Then the L.H.S. of (4) increases by at most 3ε, while the R.H.S. increases by $2\sqrt{6}\varepsilon > 4.8\varepsilon$. We conclude that (4) holds for all values of $|vv_0| \le |vv_1| \le |vv_2| \le |vv_3| \le |vv_4|$. □

Sort $v_1, \ldots, v_k$ by distances, so that $|vv_1| \le \cdots \le |vv_k|$. The main idea is to divide the sequence $v_1, \ldots, v_k$ into blocks of 4 and rearrange each block separately (Khuller et al. in contrast divide into blocks of 3). Lemma 5.1 implies that there is a path from v_{k-4j-4} to v_{k-4j} via $v_{k-4j-3}, \ldots, v_{k-4j-1}$, of weight at most $\frac{2}{3}\sqrt{6}\sum_{i=k-4j-3}^{k-4j} |vv_i|$, for all $j = 0, \ldots, \lfloor (k-5)/4 \rfloor$. Special cases of Lemma 5.1 imply that for $\ell \in \{1, \ldots, 4\}$, there is a path from v to v_ℓ via $v_1, \ldots, v_{\ell-1}$, of weight at most $\frac{2}{3}\sqrt{6}\sum_{i=1}^{\ell} |vv_i|$. Our result follows immediately by concatenating these paths.

As Khuller et al. have observed, this algorithm can be implemented in linear time for any dimension d (possibly nonconstant) if the MST is given. No near-linear time bound is known for the MST in three and higher dimensions, but a $(1+\varepsilon)$-factor approximation to the MST (which is good enough for our purpose) can be constructed in $O(n \log n)$ time for any fixed $\varepsilon > 0$ if d is a constant [8].

6 Conclusions

The obvious open problem is to improve the upper bounds further by designing better algorithms. Currently, the only published lower bounds [11] are $\frac{1}{4}(\sqrt{2}+3) > 1.103$ and $\frac{1}{5}(F(72°)+4) > 1.035$ respectively for the worst-case ratio of the minimum degree-3 and degree-4 spanning tree to the MST in the plane (achieved by the center plus vertices of a square and a regular pentagon respectively); Fekete et al. [9] boldly conjectured that these lower bounds are tight. If we insist that a designated root has degree at most 2 and 3 respectively, then the lower bounds increase to $\frac{1}{3}(\sqrt{3}+2) > 1.244$ and $\frac{1}{4}(\sqrt{2}+3) > 1.103$ (this time, by the center plus vertices of an equilateral triangle and a square).

We hope that our work would inspire more progress on the determination of these fascinating constants.

Acknowledgements

I thank Alex Golynski for asking whether the 1.5 bound for planar degree-3 trees is improvable, and I thank the referees for making some helpful suggestions on the presentation of the algorithms.

References

[1] S. Arora. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *J. ACM*, 45:753–782, 1998.

[2] S. Arora. Approximation schemes for NP-hard geometric optimization problems: a survey. *Math. Programming*, 97:27–42, 2003.

[3] S. Arora and K. C. Chang. Approximation schemes for degree-restricted MST and red-blue separation problem. In *Proc. 13th Int. Colloq. Automata, Languages, and Programming*, Lect. Notes in Comput. Sci., vol. 2719, Springer-Verlag, pages 176–188, 2003.

[4] M. Bern and D. Eppstein. Approximation algorithms for geometric problems. In D. S. Hochbaum (ed.), *Approximation Algorithms for NP-Hard Problems*, PWS Publishing Co., Boston, MA, pages 296–345, 1997.

[5] N. Christofides. Worst-case analysis of a new heuristic for the traveling salesman problem. Tech. Report 388, Graduate School of Industrial Administration, Carnegie Mellon University, 1975.

[6] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms* (2nd ed.). McGraw-Hill, 2001.

[7] D.-Z. Du and F. K. Hwang. A proof of the Gilbert-Pollak conjecture on the Steiner ratio. *Algorithmica*, 7:121–136, 1992.

[8] D. Eppstein. Spanning trees and spanners. In J.-R. Sack and J. Urrutia (eds.), *Handbook of Computational Geometry*, Elsevier Science Publishers B.V. North-Holland, Amsterdam, pages 425–461, 2000.

[9] S. P. Fekete, S. Khuller, M. Klemmstein, B. Raghavachari, and N. Young. A network-flow technique for finding low-weight bounded-degree trees. *J. Algorithms*, 24:310–324, 1997.

[10] S. P. Fekete and H. Meijer. On minimum stars and maximum matchings. *Discrete Comput. Geom.*, 23:389–407, 2000.

[11] S. Khuller, B. Raghavachari, and N. Young. Low degree spanning trees of small weight. *SIAM J. Comput.*, 25:355–368, 1996.

[12] J. Könemann and R. Ravi. A matter of degree: improved approximation algorithms for degree-bounded minimum spanning trees. In *Proc. 32nd ACM Sympos. Theory Comput.*, pages 537–546, 2000.

[13] J. N. Lillington. Some extremal properties of convex sets. *Math. Proc. Cambridge Philosophical Society*, 77:515–524, 1975.

[14] J. S. B. Mitchell. Guillotine subdivisions approximate polygonal subdivisions: a simple polynomial-time approximation scheme for geometric TSP, k-MST, and related problems. *SIAM J. Comput.*, 28:1298–1309, 1999.

[15] C. Monma and S. Suri. Transitions in geometric minimum spanning trees. *Discrete Comput. Geom.*, 8:265–293, 1992.

[16] C. H. Papadimitriou and U. V. Vazirani. On two geometric problems related to the traveling salesman problem. *J. Algorithms*, 5:231–246, 1984.

[17] B. Raghavachari. Algorithms for finding low-degree structures. In D. S. Hochbaum (ed.), *Approximation Algorithms for NP-Hard Problems*, PWS Publishing Co., Boston, MA, pages 266-295, 1997.

[18] R. Ravi, M. V. Marathe, S. S. Ravi, D. J. Rosenkrantz, and H. B. Hunt III. Approximation algorithms for degree-constrained minimum-cost network-design problems. *Algorithmica*, 31:58–78, 2001.

[19] G. Robins and J. S. Salowe. Low-degree minimum spanning trees. *Discrete Comput. Geom.*, 14:151–165, 1995.

[20] D. J. Rosenkrantz, R. E. Stearns, and P. M. Lewis. An analysis of several heuristics for the traveling salesman problem. *SIAM J. Comput.*, 6:563–581, 1977.