

On the Number of Minimal 1-Steiner Trees

Boris Aronov* Marshall Bern† David Eppstein‡

Abstract

We count the number of nonisomorphic geometric minimum spanning trees formed by adding a single point to an n -point set in d -dimensional space, by relating it to a family of convex decompositions of space. The $O(n^d \log^{2d^2-d} n)$ bound that we obtain significantly improves previously known bounds and is tight to within a polylogarithmic factor.

The *minimum Steiner tree* of a point set is the tree of minimum total length, spanning the input points and possibly having some additional vertices. Georgakopoulos and Papadimitriou [1] define the *minimum 1-Steiner tree* as the minimum spanning tree whose set of vertices consists of the input points and exactly one extra point. They describe an algorithm for constructing this tree in the plane by enumerating all combinatorially distinct *minimal* 1-Steiner trees, that is, trees formed by fixing the position of the extra vertex and considering the minimum spanning tree of the resulting set of points. They give an $O(n^2)$ bound on the number of such trees and construct a matching lower bound. Independently, Monma and Suri [3] proved the same bounds in the plane and gave bounds of $O(n^{2d})$ and $\Omega(n^d)$ in any dimension d . We significantly improve Monma and Suri's upper bounds. We show that the maximum possible number of combinatorially distinct minimal 1-Steiner trees on n points in d -space is $O(n^d \log^{2d^2-d} n)$ for all dimensions d . Our bounds require—as do the bounds in previous papers—either an assumption of general position or a tie-breaking rule, so that the minimum spanning tree is unique; otherwise even in the plane there can be exponentially many distinct topologies.

We begin by stating a theorem of Kalai [2]:

Theorem 1. *In a family of n convex sets in d -space with the property that no $c + 1$ of the sets share a point, the number of k -tuples of sets that do have a point in common is bounded by $\sum_{i=0}^d \binom{c-d}{k-i} \binom{n+d-c}{i}$, whenever $0 \leq k \leq c$.*

Here $\binom{c-d}{k-i} = 0$ if $c - d < k - i$. So in particular, putting $k = c$ we have:

Corollary 1. *Given a family of convex decompositions of d -space with a total of n cells, the number of cells in the decomposition formed by overlaying them is less than $\binom{n}{d}$.*

A slightly weaker bound can be proved by an elementary argument that we now sketch. Assume that the convex decompositions are in general position with respect to each other;

*Computer Science Dept., Polytechnic University, Six MetroTech Center, Brooklyn, NY 11201.

†Xerox Palo Alto Research Center, 3333 Coyote Hill Rd., Palo Alto, CA 94304.

‡Dept. of Information and Computer Science, University of California, Irvine, CA 92717; research performed in part while visiting Xerox PARC.

otherwise, a small perturbation only increases the number of cells. Assume also that no edges are perpendicular to a vertical axis. First we count the number of cells in the overlay that are bounded below. The lowest point of a cell is the lowest point in the intersection of some collection of i cells, $i \leq d$, each from a different convex decomposition. Adding an “infinitely low” hyperplane, that is, lying below all finite cells, allows us to count all other cells in the same way. Altogether, we can conclude that there are at most $\sum_{i=0}^d \binom{n+1}{i}$ cells in the overlay.

Let $S(P)$ be the number of distinct labeled minimum spanning tree topologies that can arise from the addition of one point to a set P of n points in d -space. Let $S_d(n) = \max_{|P|=n} S(P)$. We now use the corollary above to bound $S_d(n)$.

Let T be the minimum spanning tree of point set P , $|P| = n$. When new point x is added to P , the new minimum spanning tree $T' = T'(x)$ has certain additional edges adjacent to x . The tree might also lack certain edges from T ; missing edges can be identified by adding new edges to T one by one, and removing the longest edge on each cycle thus created. The removed edges are completely determined by the set of the added edges. Below we construct a decomposition of space so that in each cell C of the decomposition the added edges come from a known set $X(C)$ of size $O(1)$. Once such a set is fixed, we build the 1-Steiner tree for a “generic” point $x \in C$ from T by repeatedly adding an edge $e \in X(C)$ to the current tree and deleting the longest edge in the cycle thus formed. Since there are $O(1)$ edges in the candidate list, it is easily verified that only $O(1)$ comparisons whose outcome depends on the exact position of x in C are needed and $O(1)$ distinct outcomes are possible. In particular, cell C in our decomposition gives rise to only $O(1)$ combinatorial types for T' .

Our overall strategy is as follows: (1) Construct a decomposition of space, and associate with each cell in the decomposition a small set of input points. The associated set of points is such that, for any point x in the cell, the set contains that point’s neighbors $N(x)$ in tree T' . (2) Refine this partition so that for points x within a cell a *constant-size* superset of $N(x)$ is fixed. (3) Use the above observation to deduce that each cell in the refined partition corresponds to $O(1)$ possible combinatorial types of T' .

Our construction depends on the next lemma, which gives a necessary condition for a point to belong to neighbor set $N(x)$. Define a *frame* to be a triangulation of the unit $(d-1)$ -sphere. In the plane, a frame is a partition of the unit circle into circular arcs, and in three dimensions, it is a partition of a sphere into triangles with arcs of great circles for sides. Given a point x and a frame F_d , we divide space into cones with common apex x and simplicial cross sections. The cones correspond to the simplices (“triangles”) of the frame: point y is in the cone corresponding to simplex s exactly when the unit vector parallel to xy is contained in s . Each cone is formed by intersecting d halfspaces that meet at x .

Lemma 1 (Yao [5]). *For any dimension d there exists a frame F_d , with κ_d simplices, having the following property. Let (x, y) be an edge in the minimum spanning tree of a point set. Point y is contained in some F_d -cone with apex x ; let K be that cone. Then y must be the closest input point to x in K . ■*

We now describe a method for constructing the initial subdivision of space, in which each cell corresponds to a small number of minimum spanning tree topologies. We start by building a family of convex decompositions and then apply Corollary 1. If x is newly added to the

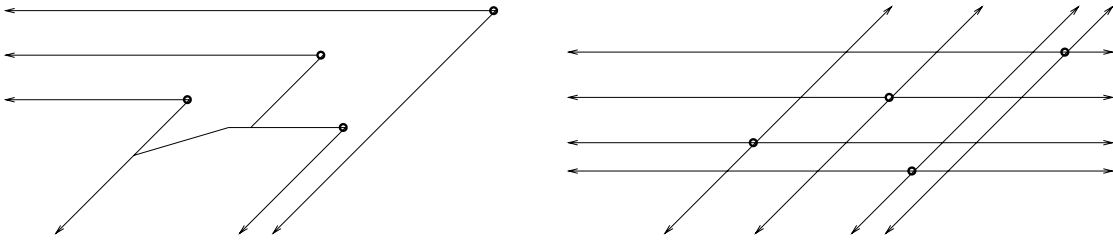


Figure 1. (a) Cone-Voronoi diagram for the simplex (arc) $[0, \pi/4]$; (b) Boxes for the same simplex.

point set, there can be $O(1)$ new edges in T' incident to x , namely, the edges (x, y) where y is the closest point to x in each of the $O(1)$ cones determined by frame F_d . We construct κ_d “cone-Voronoi diagrams”. Point x is in cell C_y of diagram i if y is the nearest input point to x among the points in the cone with apex x determined by the i th simplex. One such diagram is depicted in Figure 1(a). Superimposing the κ_d diagrams gives an arrangement in which, for each cell, the κ_d possible edges specified by Yao’s lemma are all determined.

In the plane, this construction can be used to prove that $O(n^2)$ topologies are possible. In higher dimensions, each cone-Voronoi diagram has $O(n^2)$ facets, giving an $O(n^{2d})$ bound [3]. We cannot improve this directly because the cells in each diagram are not convex. Instead, we construct a more complicated structure to which Corollary 1 can be applied directly.

Fix a single simplex of frame F_d . A cone corresponding to the simplex is defined by d hyperplanes. We pass d hyperplanes through each of the n input points, one parallel to each of the defining hyperplanes of the simplex. This divides space into $(n + 1)^d$ regions, in the form of parallelotopes (or *boxes*). Figure 1(b) depicts this subdivision in the plane. For any point x added as the new point in a 1-Steiner tree, the set of input points in the cone with apex x is determined only by the box containing x , and not by the exact location of x .

Number each parallel set of hyperplanes from 1 to n , in order of the hyperplanes’ intersections with a line passing through them, with numbers increasing in the direction in which a cone corresponding to the simplex is unbounded (left-to-right, bottom-to-top in Figure 1(b)). Associate numbers 0 and $n + 1$ with “hyperplanes at infinity” on either side of the n hyperplanes. Each box is now determined by $2d$ indices, in d pairs corresponding to the d hyperplane directions. If each pair consists of two adjacent hyperplanes, then the box is a smallest—or “atomic”—box; nonadjacent pairs give larger boxes that are unions of atomic boxes. We say a box is *regular* if each pair is of the form $(i2^j, (i + 1)2^j)$ for some i and j . Regular boxes can be grouped by “size”, e.g., 1×1 , 2×1 , 1×2 , etc. If we assume, without loss of generality, that $n + 1$ is a power of 2, each generic point is contained in the same number of boxes, and is contained in $c_d(n) = O(\log^d n)$ regular boxes, one of each size. For each possible additional point x , we can find a nonoverlapping set of $O(\log^d n)$ regular boxes, each of which is contained in the cone with apex x and which together contain all the input points in the cone. (This construction generalizes the familiar one that realizes a ray as a union of $O(\log n)$ basic segments in a segment tree [4].)

We say box b dominates box b' if each lower index for b is at least as large as the corresponding higher index for b' . For example, a box defined by pairs $(4, 6)$ and $(8, 16)$ dominates a box with pairs $(0, 4)$ and $(4, 5)$. Let the *domination cone* of box b be the union of all boxes dominated by b .

We construct the Euclidean Voronoi diagram for the points in each regular box and truncate the diagram to lie within the box's domination cone. We turn the truncated Voronoi diagram into a convex decomposition of space by cutting the complement of the domination cone into $O(d)$ convex pieces (for example, by extending one by one the hyperplanes bounding the domination cone). Each input point is included in (on the boundary or in the interior of) $2^d c_d(n) = O(\log^d n)$ regular boxes, so there are $O(n \log^d n)$ cells in all diagrams.

We repeat this construction for each one of the $\kappa_d = O(1)$ simplices of F_d and thus obtain a family of convex decompositions of space with a total of $O(n \log^d n)$ cells. We now classify each (generic) point of space according to the set of cells containing it (that is, its cell in the overlay). The number of such classes by Corollary 1 is

$$(O(n \log^d n))^d = O(n^d \log^{d^2} n).$$

For each cell C of the overlay, we can now fix a set of $O(\log^d n)$ "candidate" input points with the following property: for any point $x \in C$ and any simplex of F_d , one of the candidates is the closest input point to x in the cone with apex x corresponding to that simplex. To form the set of candidates for cell C , we consider each simplex of F_d in turn, and find the nonoverlapping set of $O(\log^d n)$ regular boxes that exactly cover the input points in a cone with apex in C . (Since C is contained in one atomic box, the choice of apex does not matter.) We then add to the candidate set exactly those input points (at most one per box) whose Voronoi cells contain C . Thus we have constructed the initial decomposition of space, referred to in step (1) above.

Actually, for any given simplex of F_d , not all of the $O(\log^d n)$ candidates are necessary. If one regular box lies in the domination cone of another regular box, then the candidate in the second box is further from x than the candidate in the first box. (A sphere centered at x through the furthest corner of the first box does not intersect the second box.) This observation reduces the size of the set of candidates for a given simplex to $O(\log^{d-1} n)$, because each nonempty "maximally dominated" box must have some lower index that is at least as small as the corresponding index for any other nonempty "maximally dominated" box. Constructing the Euclidean Voronoi diagram of the $O(\log^{d-1} n)$ candidate points partitions cell C according to the identity of the closest point. Overlaying the κ_d diagrams corresponding to different simplices produces a partition of C in each cell of which the closest point in any cone is fixed; this is the refinement of step (2).

The number of resulting sub-cells is $(O(\log^{d-1} n))^d = O(\log^{d^2-d} n)$ by another application of Corollary 1, giving a total of $O(n^d \log^{2d^2-d} n)$ cells overall. For each of the resulting cells C' , the neighbors of x in T' come from a fixed set of cardinality $\kappa_d = O(1)$. By previous discussion, C' must give rise to $O(1)$ distinct combinatorial types of the minimal 1-Steiner tree. Thus we have shown:

Theorem 2. $S_d(n)$ is $O(n^d \log^{2d^2-d} n)$.

We mention again that this bound, apart from the polylogarithmic factor, is tight. Clearly, there is some room for improvement. In fact, it is conceivable that $S_d(n)$ is $\Theta(n^d)$.

In conclusion, we thank Janos Pach for providing a crucial pointer.

References

- [1] G. Georgakopoulos and C. Papadimitriou. The 1-Steiner tree problem. *J. Algorithms* 8 (1987) 122–130.
- [2] G. Kalai. Intersection patterns of convex sets. *Israel J. Math.* 48 (1984) 161–174.
- [3] C. Monma and S. Suri. Transitions in geometric spanning trees. *7th ACM Symp. Computational Geometry* (1991) 239–249.
- [4] F.P. Preparata and M.I. Shamos. *Computational Geometry: An Introduction*. Springer-Verlag, 1985.
- [5] A.C. Yao. On constructing minimum spanning trees in k -dimensional space and related problems. *SIAM J. Comput.* 11 (1982) 721–736.