JOURNAL OF COMBINATORIAL THEORY, Series A **74**, 115–130 (1996)
ARTICLE NO. 0040

Disproofs of Generalized Gilbert–Pollak Conjecture on the Steiner Ratio in Three or More Dimensions

DING-ZHU DU

Computer Science Department, University of Minnesota, Minneapolis, Minnesota 55455.
Support in part by the National Science Foundation under Grant CCR-9208913

AND

WARREN D. SMITH

NEC, 4 Independence Way, Princeton, New Jersey 08540

Received December 4, 1992

The Gilbert–Pollak conjecture, posed in 1968, was the most important conjecture in the area of "Steiner trees." The "Steiner minimal tree" (SMT) of a point set P is the shortest network of "wires" which will suffice to "electrically" interconnect P. The "minimum spanning tree" (MST) is the shortest such network when only *intersite line segments* are permitted. The generalized GP conjecture stated that $\rho_d = \inf_{P \subset \mathbf{R}^d} (l_{\mathrm{SMT}}(P)/l_{\mathrm{MST}}(P))$ was achieved when P was the vertices of a regular d-simplex. It was showed previously that the conjecture is true for $d=2$ and false for $3 \leqslant d \leqslant 9$. We settle remaining cases completely in this paper. Indeed, we show that any point set achieving ρ_d must have cardinality growing at least exponentially with d. The real question now is: What are the true minimal-ρ point sets? This paper introduces the "d-dimensional sausage" point sets, which may have a lit to do with the answer.

1. INTRODUCTION

The conjecture of Gilbert–Pollak [5] was the most important conjecture in the area of "Steiner trees." The "Steiner minimal tree" (SMT) of a point set P is the shortest network of "wires" which will suffice to interconnect P. The "minimum spanning tree" (MST) is the shortest such network when only *intersite line segments* are permitted. The Gilbert and Pollak [5] conjectured that

$$\rho_2 = \inf_{P \subset \mathbf{R}^d} \frac{l\,\mathrm{SMT}(P)}{l\,\mathrm{MST}(P)}$$

0097-3165/96 $18.00

was achieved when P was the vertices of an equilateral triangle. In the same paper, they also suggested a generalization of their conjecture that

$$\rho_d = \inf_{P \subset \mathbf{R}^d} \frac{l\,\mathrm{SMT}(P)}{l\,\mathrm{MST}(P)}$$

was achieved when P was the vertices of a regular d-simplex. This generalization was also referred as the Gilbert–Pollak conjecture in some later references [7, 8, 12]. To distinguish their original serious conjecture and their suggested conjecture,[1] we will use the terminologies, the Gilbert–Pollak conjecture and the generalized Gilbert–Pollak conjecture, respectively. The Gilbert–Pollak conjecture was proved by Du and Hwang [4]. The generalized Gilbert–Pollak conjecture in the cases $3 \leqslant d \leqslant 9$ was disproved by Smith [12]. We settle remaining cases completely in this paper.

The purpose of this paper is to present disproofs in all dimensions $\geqslant 3$. We[2] will first (Section 3) present an ultrasimplified disproof. The reader who is solely interested in convincing himself that the generalized Gilbert–Pollak conjecture is false when $d \geqslant 3$, may stop here. Next (Section 5) we will sketch a second disproof, whose advantage is that it brings to center stage the remarkable "d-dimensional sausage" point set.

Finally (Section 7), we present a third disproof based on a nonconstructive "packing principle." This one's advantage is that it shows that any N-point set in $\mathbf{R}^d$ having the minimal Steiner ratio must have cardinality N growing at least exponentially as a function of d. Both the second and the third disproofs work well for large d and leave small d for computational verification or more careful calculation.

The attention now (Section 8) shifts to determining just what point set P *is* the one with minimal ρ_d for each $d \geqslant 2$. The d-sausage could be the answer. Section 6 presents tables of ρ values showing that it is the current record-holder in all dimensions. A forthcoming study by WDS and J. MacGregor Smith will present extensive heuristic evidence to support the belief that the 3-sausage is optimal. There is reason to suspect, however (Section 8), that in sufficiently high dimensions d, the d-sausage is *not* optimal. (Section 8 summarizes what is and is not known about ρ_d.) Thus, the situation is rather mysterious.

[1] The original statement is as follows: "One might further conjecture, by analogy with the situation in the plane, that the corners of the regular simplex in D-dimensional Euclidean space provide a minimum of L_s/L_m." ($L_s = lSMT$ and $L_m = lMST$.)

[2] Different disproofs were found by both authors, working independently. We then decided to write a joint paper.

2. Some Well-Known Facts and Some Notation

Any Steiner tree is in fact a tree made of line segments and has no angle $<120°$. (*Reason.* otherwise, an infinitesimal perturbation would shorten the tree.) An easy theorem, which was proven in more generality by Rankin [9], asserts that no more than three points on a d-sphere can exist, such that the angular separation between any two of them is $\geqslant 120°$. Hence every vertex of a Steiner three has valency $\leqslant 3$.

We will let $\rho(P)$ denote the Steiner ratio (length of the Steiner minimal tree divided by the length of the minimum spanning tree) of a point set P. ρ_d is the infimum of $\rho(P)$ over all point sets P in $\mathbf{R}^d$. We use "MST" to stand for minimum spanning tree and "SMT" to stand for Steiner (minimal) tree. ▌ denotes the end of, or absence of, a proof. (Proofs of easy lemmas will be omitted.) All "simplices" mentioned from Section 4 on will be regular.

3. Ultra-simplified Disproof

Here is the simplest disproof of the generalized Gilbert–Pollak conjecture we have been able to devise.

Theorem 1. *The regular d-simplex cannot have the minimal Steiner ratio if $d \geqslant 3$.*

Proof. Consider the set P of $1+d+d^2$ points in $\mathbf{R}^{d+1}$ whose Cartesian coordinates are all permutations of $(1, -1, 0, 0, ..., 0)$, plus the additional point $(0, 0, ..., 0)$. Observe:

1. These points actually lie in the d-space arising from the hyperplane $\sum_{i=1}^{d+1} x_i = 0$.
2. The distance between any two of these points is $\geqslant \sqrt{2}$.
3. The points form $d+1$ regular d-simplices, each of which share the common vertex $(0, 0, ..., 0)$ but no other vertices. (The ith of these $d+1$ regular simplices just arise from the points having a 1 in coordinate i.)
4. Each of these regular simplices has edge length $\sqrt{2}$.

Now clearly $\rho(P) \leqslant \rho$ (d-simplex) by observations 2, 4. But by observation 3, the fact that $d+1 \geqslant 4$, and the fact that any tree with a point of valency $\geqslant 4$ cannot be minimal, we see $\rho(P) < \rho$ (d-simplex) if $d \geqslant 3$. ▌

4. The d-Sausage

The disproof above showed that an angle was $<120°$. The proof of this was simply by use of Rankin's theorem; no attempt was made to force angles to be as small as possible and thus gain the largest possible decrease in ρ. (Indeed, the point set we used probably never even minimizes ρ to small perturbations.)

The d-dimensional point set which we call the "d-sausage" *is* the product of such a minimization attempt.[3] This point set may be described as follows.

1. Start with a unit (diameter) ball in d-space.

2. Successively add unit balls so that the Nth ball you add, is always touching the $\min(d, N-1)$ most recently added balls.

This procedure uniquely (up to congruence) defines an infinite sequence of interior-disjoint numbered d-balls. The centers of these balls form a discrete point set, which we call the d-sausage. Indeed it is most convenient to consider the doubly infinite sausage in which there is a ball corresponding to very integer, both positive and negative (although this makes no difference to the Steiner ratio). The first N points of the d-sausage will be called the "N-point d-sausage."

There are two convenient mental images of the d-sausage. The first is to view it as the vertices of a simplicial complex made of d-simplices pasted together at common faces. Namely, points numbers $m, m+1, m+2, ..., m+d$ for a regular d-simplex for any integer m, and two such simplices are adjacent at a common $(d-1)$-face iff their m values are adjacent.

The second mental image is that the points lie in consecutive order equally spaced along a d-dimensional "helix." That is, the d-sausage has an infinite, transitive, discrete group of symmetry isometries. An isometry of d-space exists that will map sausage point m to sausage point $m+k$ for all m simultaneously where k is any integer, and it consists of a translation by k times a particular fixed "axial" vector, plus k times a particular "twist" (rotation of d-space preserving the axial vector). The complete linear transformation—axial translation and twist—arising when $k=1$ is readily uniquely determined from the coordinates of points $0, 1, 2, ..., d+1$. Suitable helices would arise by generalizing to allow k to be a real number by means of suitable (nonunique) notions of noninteger powers of a linear transformation.

It is an immediate consequence of this second view.

Lemma 2. *The construction above can be continued indefinitely and hence the d-sausage does exist.*

[3] The 6-point 3-sausage was first discovered by J. MacGregor Smith. The generalization to any values of N and d is due to WDS.

It also follows that

LEMMA 3. *Every sausage point is an extreme point of the d-sausage's convex hull.*

Clearly the length of the MST of the N-point d-sausage is $N-1$. The Steiner ratio of the d-sausage will be defined as the infimum of the Steiner ratios of the N-point d-sausages as $N \to \infty$. Since ρ for the N point sausage is obviously at least as large as ρ for the $(2N-1)$-point sausage (consider gluing two SMTs for an N-point sausage together at a common sausage point), one realizes that this is not only the infimum, but in fact it is the limit.

5. THE d-SAUSAGE'S STEINER RATIO IS LESS THAN THE d-SIMPLEX'S

We will prove

THEOREM 4. *The Steiner ratio of the* $(2d+1)$-*point d-sausage is less than the Steiner ratio of the regular d-simplex* (*which is the* $(d+1)$-*point d-sausage*) *if* $d \geqslant 3$.

Remark 5. In fact a stronger statement apparent from our argument is that the Steiner ratio of the $(2N-1)$-point d-sausage is less than the Steiner ratio of the N-point one, for each integer N, $N > d \geqslant 10$.

It is immediately obvious that

$$\rho(d\text{-sausage}) \leqslant \rho(d\text{-simplex}) \qquad \text{if} \quad d \geqslant 1.$$

This arises from the "first mental image" of the sausage, by consideration of a network T made of the SMTs of the simplices whose vertices are $\{1, 2, ..., d+1\}$, $\{d+1, d+2, ..., 2d+1\}$, $\{2d+1, 2d+2, ..., 3d+1\}$, et cetera.

In fact, equality holds when $d=1$ and $d=2$. (When $d=2$ this is a consequence of the Du-Hwang theorem (verification of the GP conjecture when $d=2$).)

To prove strict inequality when $d \geqslant 3$, we will show that T is improvable, because at sausage points numbered 1 mod d, which have valence 2 in T, the angle Ψ_d formed by the incident line segments of T is $<120°$. Note that there are many possible equal-length versions of T, arising from applying any of $(d+1)!$ possible isometries to each of these simplices.

Let β_d denote the angle, as viewed from sausage point $d+1$, between the centers of the simplices whose vertices are sausage points $1, 2, ..., d+1$, and $d+1, d+2, ..., 2d+1$.

Frank Morgan pointed out the following to us.

LEMMA 6. *At least one of the equal length version of* T *must involve an angle, at sausage point* $d+1$, *whose cosine exceeds* $\cos\beta_d$.

Proof. The $(2d+1)$-point d-sausage consists of two simplices S, S' with a point p (which we have been calling "sausage point $d+1$") in common. Let $\hat{a}$ be a unit vector from p towards the center of S. Let $\hat{b}_1$ be a unit vector of some edge from p of the SMT of S. Let $\{\hat{b}_i: i=1...m\}$ be its images under the m isometries of $S\cap T$ fixing p. Let primes indicate similar objects of S'. Then

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{b}_i\cdot\hat{b}'_j=\left(\sum_{i=1}^{m}\hat{b}'_i\right)\cdot\left(\sum_{j=1}^{m}\hat{b}'_j\right)=(s\hat{a})\cdot(s'\hat{a}')=ss'\cos\beta_d$$

with s, s' positive and at most m. ∎

So if β_d is less than 120°, our theorem is proven.

Although we do not know a closed formula for β_d, the first few values arise from $\cos\beta_d=f(d)/(d^d\binom{d+1}{2})$, where the values of the integer $f(d)$ are

d	f(d)
1	−1
2	−6
3	−34
4	−40
5	5829
6	225568
7	7221500
8	229477248
9	5781285995
10	264454496768.

We have calculated the first 500 values by use of the following convenient explicit coordinates for the points of the d-sausage: The first $d+1$ points are columns of the $(d+1)\times(d+1)$ identity matrix in $\mathbf{R}^{d+1}$. (These, of course, lie in the hyperplane whose coordinate sum is 1.) The remaining points $\mathbf{p}_i$ are got by the recurrence

$$\mathbf{p}_i=\frac{2}{d}\sum_{j=1}^{d}\mathbf{p}_{i-j}-\mathbf{p}_{i-d-1}.$$

(We are now using ball diameters of $\sqrt{2}$ instead of 1; for convenience. Hopefully the reader will have no trouble adjusting to this irrelevant scale factor.)

For all d with $3 \leqslant d \leqslant 500$, we found $\cos \beta_d > -\frac{1}{2}$. In fact, as is suggested by the fact that $\cos \beta_{160} \approx 0.960637$,

LEMMA 7. *When* $d \to \infty$, β_d *approaches* 0.

Remark 8. Indeed, $\beta_d \approx \sqrt{12.6/d}$ radians when d is large. (This number "12.6" is merely an extrapolation from numerical results. It may not be precise and we do not claim to have proven it.)

Proof Sketch. Use the convenient coordinates mentioned above. The distance from a vertex to the simplex center is $(d^2 - d + 1)/d^2 \to 1$. Constructing the sausage by a sequence of reflections of simplices in their faces, we see that the distance $L_{d,k}$ between simplex centers m and $m+k$ is the total length of a k-step walk, each of whose steps is $\leqslant 2/d$ long (i.e., since $1/d$ upper bounds the distance from a simplex center to a face center). The turning angles between the walk steps (regarded as line segments) are the same as the dihedral angle of a d-simplex, that is $\arccos(1/d)$, which approaches a right angle.

Were it the case that these were exactly right angles, every step being in a direction orthogonal to all previous ones, then it would be the case that $L_{d,k} = L_{d,1}\sqrt{k}$, which is the intuition behind this proof—matters being slightly more complicated since the angles are not exactly right angles. In fact,

$$L_{d,k} = |(I + R + R^2 + \cdots + R^{k-1})\,\mathbf{v}|,$$

where $\mathbf{v}$ is the vector representing the first walk step, and R is a certain orthogonal matrix whose effect is to shift one's sense of direction in between walk-steps. (Incidentally, $R - I$ is noninvertible so the usual trick for summing these series does not apply.) Imagine you are walking through $\mathbf{R}^d$ carrying a $(d-1)$-simplex, oriented normally to your direction of motion. R has the effect of turning your direction of motion an angle $\arccos(1/d)$ away from the last vertex of our simplex, and then shifting your sense of direction in the directions orthogonal to your motion, by mapping vertex i of our $(d-1)$-simplex to vertex $i+1 \bmod d$. Now, if it were the case (which it is not) that R were the circulant matrix "Q" with 1's on the superdiagonal and a 1 in the bottom-left entry, and 0's elsewhere, then $L_{d,k} = L_{d,1}\sqrt{k}$ for $k = 1 \ldots d$, and this would be precisely the all-right-angle case we had just considered. (Once $k > d$, the growth eventually becomes linear in k. This regime will not be of interest to us.) In fact, R *is* very similar in structure to Q. The angles we are rotating through are not right angles but in fact are arc sec d and arc sec$(d-1)$, which, when d is large, are very close to right angles. As a result, in a suitable basis R is a small perturbation of Q. In fact, one may see that the L_2 norm $|R-Q|$ of the perturbation is $O(1/d)$. As a result $L_{d,k}=\sqrt{k}\,L_{d,1}+O(k/d)\,L_{d,1}$ so that the perturbation has no effect on the fact that $L_{d,d}=O(1/\sqrt{d})$.

This proves $\beta_d=O(1/\sqrt{d})$, since β_d is the apex angle of an isosceles triangle with legs ~ 1 and base $L_{d,d}$. ∎

These two lemmas show that the GP conjecture is false in all sufficiently large dimensions. In fact, by exercising slightly more care (i.e., using explicit numbers instead of big ohs) one may see that Lemma 7 shows falsity when $d \geqslant 100$. The range $3 \leqslant d \leqslant 500$ is covered by our explicit calculations of β_d.

6. Explicit Upper Bounds on the Steiner Ratio of d-Sausages

We have computed the following upper bounds $g(d)$ on $\rho(d\text{-sausage})$:

d	g(d)
1	1
2	0.866025403784438647=sqrt(3)/2
3	0.784190373377122247
4	0.743985617828134
5	0.721810674855881
6	0.708536716660975
7	0.700120875002678
8	0.694558735568743
9	0.690767953841645
10	0.688124343497604
11	0.686248729433044
12	0.684900899944705
13	0.683923344683512
14	0.683209788374833
15	0.6826867992
16	0.68230263
17	0.68205
18	0.68183

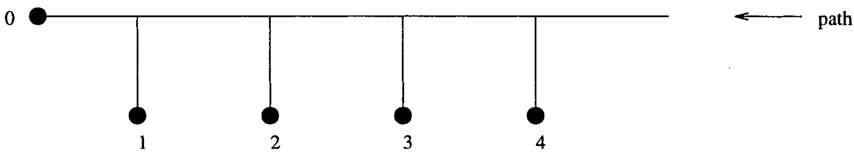

FIG. 1. Postulated "path"-type topology for d-sausage SMT.

These upper bounds[4] $g(d)$ are the lengths of certain Steiner trees for the d-sausage which have the "path-topology" given in Fig. 1, where the numbers denote the sausage points in their natural order.

The best Steiner tree with path-topology was found by using the well-known [5] concave-up behavior of the SMT-length as a function of Steiner (or any) point coordinates with the SMT topology held fixed, to see that there must be a unique local (and global) length minimizing configuration which has the same symmetry group as the d-sausage itself. Hence the geometry of the SMT within each individual sausage-simplex must be the same (i.e., the SMT is periodic, with the "unit cell" being a simplex). In that case $\rho(d\text{-sausage})$ may be found by solving a $(d-1)$-dimensional convex optimization problem, which we wrote a special-purpose computer program to do.

Specifically, inside a "unit cell" simplex there is exactly one Steiner point, attached to exactly one vertex of the simplex and to the "same" points on the two cell-faces. One need only solve for the $d-1$ barycentric coordinates of the point on the cell face which minimizes the length of this 3-terminal Steiner tree (the point on the other cell face has the same barycentric coordinates, only cyclically shifted), to determine everything. The formula for the length L of the SMT of three points forming a triangle with sides a, b, c, $a \leqslant b \leqslant c$, and area Q is

$$L = \min(\sqrt{(a^2+b^2+c^2)/2 + 2\sqrt{3}\,Q}, a+b)$$

and this length-function, of course, is a concave-up function of the unknown barycentric coordinates. This concavity makes the minimization problem easy despite its $(d-1)$-dimensionality.

For example when $d=4$, the barycentric coordinates at the minimum are $\approx (0.162, 0.338, 0.338, 0.162)$. If the unit cell simplex is taken to have vertices which are the rows of the 5×5 identity matrix, then the intersection of the 4-sausage's SMT with the unit cells is in fact the SMT of the three

[4] We believe the numerical errors in $g(d)$ are a unit last place. Regardless of possible numerically inadequate optimization, the values given are still valid upper bounds to unit last place accuracy.

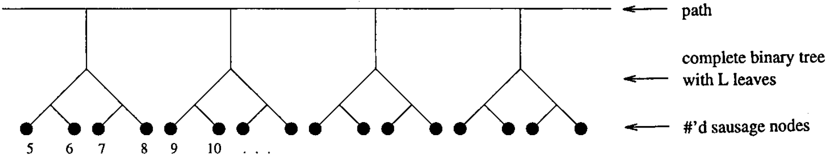

FIG. 2. Postulated "path of trees" topology for d-sausage SMT.

points with coordinates[5] (0.162, 0.338, 0.338, 0.162, 0), (0, 0, 1, 0, 0), and (0, 0.162, 0.338, 0.338, 0.162). By (6) we see that this SMT has length ≈ 1.052. Meanwhile the MST of this 4-sausage has length $\sqrt{2} \approx 1.414$ per sausage point, for a ratio $g(4) \approx 0.744$.

The postulation of the path-topology of Fig. 1 is supported by the fact that it is true when $d=1$ and true WLOG when $d=2$, and the path topology is correct for the N-point sausages when $d=3$ and $N=4, 5, \ldots, 13$. Even if and when this postulation is not true (more on that below), it still leads to valid upper bounds, of course.

We have the following monotonicity lemma.[6]

LEMMA 9. *ρ(d-sausage) is a strictly decreasing function of d.*

Proof Sketch. The $(d-1)$-sausage may be embedded inside the d-sausage by "folding" $(d-1)$-space inside d-space much as you would fold a sheet of paper. The folding process leads to an interconnecting tree for the d-sausage which is not longer than the shortest such tree for the $(d-1)$-sausage; in fact this tree must be improvable due to the fact that it contains bent line segments, which it must do unless the SMT consisted of a union of SMTs each lying inside a component simplex, a possibility which we have previously ruled out. ∎

By applying Wynn's epsilon extrapolation [14] to $g(1 \ldots 14)$, we find that

$$\lim_{d \to \infty} g(d) \approx 0.6812 \pm 0.0001.$$

Meanwhile it is know [1] that the d-simplex has the Steiner ratio which tends to $C = \sqrt{3}/(4-\sqrt{2}) \approx 0.66984$ or below, and that the 80-simplex

[5] For your information, the exact values of the numbers "≈ 0.162" and "≈ 0.338" are the root r of $90x^4 + 24x^3 + 5x^2 - 8x + 1 = 0$ and $1/2 - r$, respectively.

[6] Meanwhile, one may show $a\rho(a\text{-simplex}) + b\rho(b\text{-simplex}) < (a+b)\,\rho([a+b]\text{-simplex})$ if $a \geqslant b \geqslant 1$. Also, obviously ρ_d is a nonincreasing function of d. Finally, by a proof exactly similar to the lemma's, $g(d)$ is a strictly decreasing function.

has $\rho \leqslant 0.677754$. The fact that $0.6812 > 0.6777 > 0.6698$ indicates that, although the path-topology of Fig. 1 is almost certainly correct when $d=3$, it probably is *not* optimal in sufficiently large dimensions of d, in particular, $d=80$.

We conjecture that the correct topology is in the "path-of-trees" topology family pictured in Fig. 2. (The path-topology of Fig. 1 is the special case of Fig. 2, arising when $L=1$.)

7. Sets with the Minimal Steiner Ratio Have Exponentially Large Cardinality

The fundamental idea which made the proof of Theorem 1 work, is encapsulated by the following theorem.

Theorem 10 (Packing principle). *Given a d-dimensional point set P, $d \geqslant 3$, construct the union U of cones of (diametric) apex angle 60° whose common apex is a fixed point $A \in P$, and whose axial rays are rays* **AB**, *$B \in P$. If at least four interior-disjoint copies of U can be placed in d-space with common apex A, and $|P| < \infty$, then $\rho_d < \rho(P)$.*

Proof. Four suitable copes of P placed with common point A, form a set P_4 of cardinality $4\,|P|-3$, which will have the Steiner ratio strictly smaller than $\rho(P)$ if $|P| < \infty$ since the Steiner tree formed of four copies of the SMT of P has a 4-valent vertex A and is, hence, improvable. The 60° angular "buffer zones" prevent $l\,\mathrm{MST}(P_4) < 4l\,\mathrm{MST}(P)$. ∎

One may conjecture that a minimal-ρ point set $\mathbf{R}^d$ for any $d \geqslant 3$ must have infinite cardinality. We have no inkling of how to prove this conjecture, but we can show the following remarkable consequence of the packing principle.

Theorem 11. *An N-point set $P \subset \mathbf{R}^d$ which achieves the minimal Steiner ratio ρ_d, must have cardinality N obeying*

$$N \geqslant \left\lceil \tfrac{1}{2}\sqrt{VB(\pi/3,\, d)} \right\rceil + 1,$$

where

$$VB(\theta,\, d) = \frac{2I_{d-2}(\pi/2)}{I_{d-2}(\theta)}$$

and

$$I_m(x) = \int_0^x (\sin u)^m \, du.$$

Remark 12. The functions $I_m(x)$ and $VB(\theta, m)$ are discussed in the WDS's Ph.D. thesis [10, Section III.C.3.1-2]. In particular, full asymptotic expansions, derived using Laplace's method, are given there for m large. For our present purposes it will suffice to say that

$$I_0(x) = x, \qquad I_1(x) = 1 - \cos x, \qquad I_2(x) = \frac{x - \sin x \cos x}{2}$$

and for $m > 2$ one may recover $I_m(x)$ by use of the integration by parts formula

$$I_m(x) = \frac{m-1}{m} I_{m-2}(x) - \frac{(\cos x)(\sin x)^{m-1}}{m}.$$

When d is large and θ, $0 < \theta < \pi/2$, is fixed, we have

$$VB(\theta, d) \sim \cos\theta \sqrt{2\pi\, d}\, (\csc\theta)^{d-1}.$$

Thus Theorem 11 proves that N grows at least exponentially as a function of d. Note $\csc(\pi/3) = 2/\sqrt{3}$; in fact,

$$N > (3/4)^{d/4} \qquad \text{when} \quad d \geqslant 2.$$

Proof. Definition. A point set P on the unit sphere in $\mathbf{R}^d$ is "K-unpackable" if it is impossible to find K (possibly transformed by multiplication by an orthogonal matrix) copies of P on the d-sphere so that each copy is at least distance 60° away from each other copy.

Thus, for example, if the angular "covering radius" of P is $<60°$, that is, if every point of the sphere is covered by at least one spherical cap centered at a point of P and having angular radius 60°, then certainly P is 2-unpackable. The reverse implication is not necessarily true, however.

For another example (and we intend to use this one!) if P is 4-unpackable, then any set $P \cup OP$, were O is some orthogonal matrix such that $\operatorname{dist}(P, OP) \geqslant 60°$ (distances being measured angularly on the sphere) must be 2-unpackable. ▌

The packing principle may now be restated as follows: *The radial projection of a finite point set in* $\mathbf{R}^d$ *achieving* ρ_d *onto a sphere centered at any one of its points, must be* 4-*unpackable.*

Now, we will show the following lemma, from which the theorem follows.

LEMMA 13. *Any 2-unpackable point set on a sphere in* $\mathbf{R}^d$ *must have cardinality* N *with*

$$N^2 \geqslant VB(\pi/3, d).$$

Proof. We will be implicitly using the Haar measure on the space O_d of $d \times d$ orthogonal matrices. This measure (as is well known) is uniquely specified by the facts that its total integrated mass is 1 and it is invariant under group operation (multiplication by an orthogonal matrix).

Now a mass $1/VB(\pi/3, d)$ of the matrices in O_d will map a specified point i of P to within angular distance $\leqslant 60°$ of a specified other point j (which could be the same point) of P. This is because $VB(\theta, d)$ is the surface area of a sphere in $\mathbf{R}^d$ of unit radius, divided by the $(d-1)$-volume of a spherical cap of angular radius θ drawn on this sphere. This set of matrices will be called "excluded" by point pair (i, j). Since there are N^2 pairs of points, even if all the excluded sets of O_d were disjoint, a non-excluded matrix would still have to exist if $N^2 < VB(\pi/3, d)$.

If a matrix O exists that is not excluded by any point pair, then P is not 2-unpackable since $P \cup OP$ is a 2-packing. ∎

It is an interesting question just what the minimal cardinality of K-unpackable sets are, as a function of K and d and possibly when "60°" is changed to some other angle.

In low dimensions the packing principle is rather weak. Suppose one conjectures that any 2-unpackable set on a sphere in $\mathbf{R}^3$ must have cardinality $\geqslant 5$. (The 5-point set arising from the radial central projection of the triangular bipyramid is 2-unpackable.) If this is true, it would follow that any 4-unpackable set must have cardinality $\geqslant 3$, and in fact the set arising from an equilateral triangle inscribed in the sphere's equator, is apparently 4-unpackable. Even if this is all true, then the packing principle when $d = 3$ leads to nothing stronger than the obvious fact that $N \geqslant 4$ for any ρ-minimizing N-point set in $\mathbf{R}^3$.

However, when d becomes large, the exponential growth in Theorem 11 eventually kicks in to yield impressive results. Thus to compute some numbers in (11)[7]

[7] As can be seen, Theorem 11 first shows $N \geqslant d + 2$, disproving the Gilbert–Pollak conjecture in a third way, when $d = 50$. The particular formula (11) is not strong enough to disprove GP when $d < 50$, due to the rather weak (Haar measure and 4-unpackability) argument we used, but the fundamental idea behind it (the packing-principle), of course, disproves GP in every dimension $d \geqslant 3$.

d	N is at least
49	49
50	53
100	2218
200	3481911
500	10233465928731809
100	50354832002265644106698068172260

8. What Is Known about ρ_d? Open Questions

In this section we will survey the known upper and lower bounds on ρ_d and state two remaining open questions and some speculation.

Chung and Gilbert [1] showed that $\limsup_{d\to\infty} \rho(d\text{-simplex}) \leqslant C = \sqrt{3}/(4-\sqrt{2}) \approx 0.66984$ and, in fact, it is likely [12] that equality holds. Certainly the "path of trees" topologies are capable of approaching C, and it may also be shown that $\rho(d\text{-octahedron}) \leqslant C + \varepsilon_d$, where $\varepsilon_d \to 0$ as $d \to \infty$. We do not know whether the d-sausages or the d-octahedra can get below C.

The proofs of Theorem 4 (and Theorem 1, if one is careful) show $\rho(d\text{-simplex}) - \rho(d\text{-sausage}) = \Omega(1/d)$ but cannot hope to show anything stronger.

Open Question 1. Is $\rho_d > C$ for all $d \geqslant 1$?

It is trivial to see [5] that in any metric space, $\rho \geqslant \frac{1}{2}$.

Graham and Hwang [6] showed that $\rho \geqslant 1/\sqrt{3} \approx 0.57735$ in any Euclidean space.

Du [3] found the improved bound $\rho \geqslant l$, where $l = (2 + z - \sqrt{z^2 + z + 1})/\sqrt{3} \approx 0.6158277481$, and $z \approx 0.148663719631161$ is the unique positive real root of

$$128z^6 + 456z^5 + 783z^4 + 764z^3 + 408z^2 + 108z - 28.$$

As regards finite point sets, we remark that the 6-point 3-sausage has $\rho \approx 0.808065$. This compares with the $(4d+1)$-point set from the proof of Theorem 1 (when $d = 3$) at ≈ 0.809325, the regular tetrahedron (≈ 0.813052) and the regular octahedron (≈ 0.811197). All numerical facts about $\leqslant 13$-point sets mentioned arise from use of the computer program described in Smith [12].

In fact, the N-point 3-sausages are the current record-ρ N-point sets in 3-space for every $N \geqslant 4$ with the exception of $N=5$. (The best known 5-point set in 3-space is the regular tetrahedron plus another point infinitesimally near to one of the tetrahedron vertices.)

Open Question 2. For which d, if any, is it the case that $\rho_d = \rho(d\text{-sausage})$?

Certainly $d=1$ and $d=2$ work. J. MacGregor Smith and WDS [13] will soon publish a study providing extensive experimental evidence (but no proof!) that it is also true when $d=3$. The sausages hold the current records in every dimensions. Could it be that the sausages are optimal in all dimensions?

Although this question is unresolved in any dimension $\geqslant 3$, the packing principle method makes it seem plausible—although it does no prove!—that the d-sausage is *not* optimal in all sufficiently large dimensions. Indeed, this is plausible when $d \geqslant 15$. To explain this remark: in large dimensions d it becomes easy to find out d-sausages whose "balls" are interior-disjoint, except that all four sausages have exactly one ball in common. (By "easy" we mean that random rotations will work with high probability. Cf. the proof of Lemma 13.) In this case, the MST of the resulting 4-sausage object is the same as the union of the MSTs of the sausages, but the SMT is not the union of the four sausage's SMTs (since this union contains a point of valency 4) but is in fact *shorter*. This almost proves, but in fact does not prove (because $\infty - 1 = \infty$), that the d-sausage is nonoptimal in each suitable dimension.

Still, even if this almost-proof were to turn into a proof, the sausage would still be absolutely safe from it when $d=3$ and $d=4$, and quite safe, although not absolutely safe, when $d=5$. This is since four suitable sausages can only exist in dimensions d in which $8d$ disjoint unit balls can "kiss" one. The maximal kissing number when $d=3$ is 12, and when $d=4$ is 24 or 25 (widely believed to be 24 and arising in a unique configuration) and when $d=5$ is 40–46 (widely believed to be 40 and arising in a unique configuration). For these facts about kissing numbers, see [2].

Finally, a suitable infinitesimal perturbation of the d-sausage will show that the "greedy trees" (GT) of Smith and Shor do not necessarily achieve a better ratio than the MST, even for the current record-ρ point sets. This was recently observed by DZD (in a manuscript submitted to *Algorithmica*) when $d=2$, but his technique generalizes[8] to $d \geqslant 2$. Could it be that the d-sausage yields the minimal ratios for *both* SMT/MST and SMT/GT?

[8] The conjecture that greedy trees will force a better approximation than ρ_d, d fixed, may still be salvageable by altering the definition of "greed" to allow bounded "lookahead."

Acknowledgments

Frank Morgan and the second author independently simplified an earlier version of the first disproof. Frank Morgan also simplified Section 5 by proving Lemma 6. Both authors wish to thank him for his insightful comments.

References

1. F. R. K. Chung and E. N. Gilbert, Steiner trees for the regular simplex, *Bull. Inst. Math. Acad. Sinica* **4**, No. 2 (Dec. 1976), 313–325.
2. J. H. Conway and N. J. A. Sloane, "Sphere Packings, Lattices, and Groups," Springer-Verlag, New York, 1988.
3. D.-Z. Du, On Steiner ratio conjectures, *in* "Topological Network Design," Vol. 33, pp. 437–451, Annals of Operations Research (J. M. Smith and P. Winter, Eds.), Battzer, Basel.
4. D.-Z. Du and F. K. Hwang, State of the art on Steiner ratio problems, *in* "Computing in Euclidean Geometry," pp. 163–192 (D. Z. Du and F. K. Hwang, Eds.), World Scientific, Singapore, 1992.
5. E. N. Gilbert and H. O. Pollak, Steiner minimal trees, *SIAM J. Appl. Math.* **16** (1968), 1–29.
6. R. L. Graham and F. K. Hwang, A remark on Steiner minimal trees I, *Bull. Inst. Math. Acad. Sinica* **4** (1976), 177–182.
7. F. K. Hwang and D. S. Richards, Steiner tree problems, *Networks* **22** (1992), 55–89.
8. F. K. Hwang, D. S. Richards, and P. Winter, "The Steiner Tree Problems," p. 78, North-Holland, Amsterdam, 1992.
9. R. A. Rankin, The closest packing of spherical caps in *n* dimensions, *Proc. Glasgow Math. Assoc.* **2** (1955), 145–146.
10. W. D. Smith, "Studies in Computational Geometry Motivated by Mesh Generation," Ph.D. thesis, September 1988, Applied Math Dept., Princeton University; available from University Microfilms, Inc., Ann Arbor, MI.
11. W. D. Smith and P. W. Shor, Steiner tree problems, *Algorithmica* (1992), 329–332.
12. W. D. Smith, How to find Steiner minimal trees in Euclidean *d*-space, *Algorithmica* **7** (1992), 137–177.
13. J. M. Smith and W. D. Smith, On the Steiner ratio in 3D, manuscript.
14. P. Wynn, *Math. Comp.* **16** (1961), 23–29.