Journal of Combinatorial Optimization, 9, 243–262, 2005

On Approximation Ratios of Minimum-Energy Multicast Routing in Wireless Networks

HONGXU CAI hxcai00@mails.tsinghua.edu.cn
YINGCHAO ZHAO yczhao@mails.tsinghua.edu.cn
State Key Lab of Intelligent Technology & System, Department of Computer Science and Technology, Tsinghua University, China. Supported by Natural Science Foundation of China (60223004, 603210022)

Received July 10, 2004; Revised March 14, 2005; Accepted March 14, 2005

Abstract. In the broadcasting of ad hoc wireless networks, energy conservation is a critical issue. Three heuristic algorithms were proposed in Wieselthier et al. (2000) for finding approximate minimum-energy broadcast routings: MST(minimum spanning tree), SPT(shortest-path tree), and BIP(broadcasting incremental power). Wan et al. (2001) characterized their performance in terms of approximation ratios. This paper points out some mistakes in the result of Wan et al. (2001), and proves that the upper bound of sum of squares of lengths of the edges in Euclidean MST in unit disk can be improved to 10.86, thus improves the approximation ratios of MST and BIP algorithm.

Keywords: ad hoc wireless networks, broadcasting, MST

1. Introduction

Ad hoc wireless network has not any wired backbone infrastructure, and the nodes in it use antennas to transmit and receive signals. The signals can be transmitted either in single-hop or in multi-hop through intermediate relaying nodes, and transmission range of the sender is decided by the power of the sender. However, the power supplied to the nodes in a wireless network is from batteries only. Thus, we need to design efficient power assignment schemes to make the lifetime of the network as long as possible while keeping the connectivity of the network.

In the power-attenuation model (Rappaport, 1996), the signal power falls as the function $\frac{1}{r^k}$, where r is the distance between the sender and the receiver, and k is a constant between 2 and 4. In the general case, we assume that all the receivers have the same power threshold for detecting signals which is normalized to one. Then the power consumption between the sender and receiver with distance equal to r is r^2 if we choose $k = 2$. Finding a route to minimize the total power consumption is referred to as the $Minimum - Enerygy$ $Routing\ Problem$. We can use several models to design the routing scheme, for example MST(minimum spanning tree), SPT (shortest-path tree) and BIP(broadcasting incremental power) in Wieselthier et al. (2000). In this paper, we use the model of MST, and the problem can be described as follows.

Consider a unit disk with center O. Let P be a finite set of points in unit disk O. The question is that what is the value of

$$c = \sup_P \min_T \sum_{e \in T} \|e\|^2, \tag{1}$$

where T is over all spanning trees on $P \cup \{O\}$ and e is over all edges of T.

This problem above comes from the study of multicast in wireless network. Analysis of several routing algorithms is based on establishing upper bound for this max-min value. Wan etc. gave the first upper bound of 14.51 in Wan et al. (2001), and later in the corresponding erratum they corrected some mistake and gave the bound of 12.141 which is the best known result. We follow their way and improve the bound to 10.86.

Our proof of the bound is in some sense an extension of the basic idea of Wan et al. (2001), which made use of disjoint-diamonds covering on the unit disk. They obtained their result by comparing the area of the unit disk with the total area of the deliberately constructing diamonds. And we adjust their diamonds into some more complicated shape (actually they are diamonds with changeable extra areas), thus compute the area more precisely.

The paper is organized as follows. We give some preliminaries and notations in Section 2, and introduce the previous results on this problem in Section 3. In Section 4 we analyze the upper bound of sum of squares of lengths of the edges in EMST and give the proof in detail. In Section 5, we summarize the result and highlight some future research directions.

2. Preliminaries and notations

We first introduce some notations that will be used.

$\|AB\|$: the length of line segment AB.
$\triangle ABC$: the triangle ABC.
$\blacktriangle ABC$: the area inside $\triangle ABC$.
$\angle ABC$: the angle between the two rays BA and BC. Also referred to as the region inside the angle.
$D(A_1A_2)$: the rhombus whose vertices are A_1 and A_2 with sides of length $\frac{\sqrt{3}}{3}\|A_1A_2\|$(see figure 1(a)). Also called the diamond determined by edge A_1A_2.
$B(A, r)$: the set of points whose distances from A are less than r.

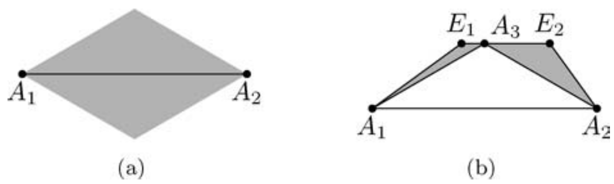


Figure 1. Illustration of (a) $D(A_1A_2)$ and (b) $E_{\alpha,\beta}(A_1A_2)$.

$L(A_1A_2)$: $B(A_1, \|A_1A_2\|) \cap B(A_2, \|A_1A_2\|)$.
$E_{\alpha,\beta}(A_1A_2)$: Let A_3 be the vertex of $D(A_1A_2)$ such that the points A_1,A_2,A_3 are in anticlockwise order, line $E_1A_3E_2$ is parallel to line A_1A_2, with $\angle E_1A_1A_3 = \alpha$ and $\angle E_2A_2A_3 = \beta$.

Then $\blacktriangle A_1A_3E_1 \cup \blacktriangle A_2A_3E_2$ is denoted by $E_{\alpha,\beta}(A_1A_2)$ (see figure 1(b)).

Now let's take a look at (1). Note that for any point set P, by Prim's algorithm for constructing minimum spanning tree, we could find that $\sum_{e\in T} \|e\|^2$ achieves its minimum value if and only if T is an Euclidean minimum spanning tree(EMST) of P. Therefore, c could also be written by

$$c = \sup_{P\subset B(O,1)} \sum_{e\in \mathrm{EMST}(P\cup\{O\})} \|e\|^2.$$

In other words, what we want to know is the upper bound of sum of squares of lengths of the edges in Euclidean MST in unit disk.

The EMSTs have many nice properties (Wan et al., 2001). For example:

- The length of every edge is no more than 1 (for EMSTs in the unit disk).
- Let A_1A_2 be any edge in EMST(P), then $L(A_1A_2)$ doesn't contain any other points in P.
- Let A_1A_2 and B_1B_2 be any two edges in EMST(P), then B_1 and B_2 are either both outside $B(A_1, \|A_1A_2\|)$ or both outside $B(A_2, \|A_1A_2\|)$.
- every edge in EMST(P) is an edge of P's Delaunay triangulation.

And in Wan et al. (2001), we know an important property of EMSTs:

Theorem 1. *In EMST, the two diamonds determined by any two edges are disjoint.*

There is another lemma in Wan et al. (2001) which we will also use in this paper, stated as follows.

Lemma 1. *Let A_1, A_2 and B be any three points in the plane, $\exists i \in \{1, 2\}$, s.t. $\|A_iB\| > \|A_1A_2\|$. Let points A_1'(A_2' respectively) be any point on the opposite side of A_1B(A_2B respectively) from A_2(A_1 respectively) such that $\angle A_1BA_1'$($\angle A_2BA_2'$ respectively) $= \frac{\pi}{6}$ (see figure 2). Then $D(A_1A_2) \subseteq \angle A_1'BA_2'$.*

3. Incorrectness in previous result

From Theorem 1, Wan et al. (2001) first analyzed the relationship between the total area of the diamonds determined by edges in EMST and sum of squares of lengths of the edges in EMST, and gave an estimation 14.51 of the upper bound of $\sum_{e\in \mathrm{EMST}} \|e\|^2$. Then by estimating the sticking-out area more precisely, they proved a better upper bound 12. However, some parts of the derivation of this new bound is wrong.

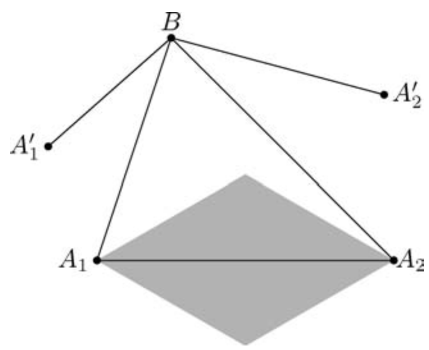


Figure 2. Illustration for Lemma 1.

The mistake is in their Lemma 11, stated as follows:
Let

$$S(\alpha) = \frac{1}{2}\sin\alpha + \frac{\sqrt{3}}{6}(1 - \cos\alpha) - \frac{\alpha}{2}.$$

Then for any $\alpha, \beta \in (0, \frac{\pi}{3})$,

1. if $\alpha + \beta \leq \frac{\pi}{3}$, $S(\alpha) + S(\beta) \leq S(\alpha + \beta)$;
2. if $\alpha + \beta \geq \frac{\pi}{3}$, $S(\alpha) + S(\beta) \leq S(\alpha + \beta - \frac{\pi}{3}) + S(\frac{\pi}{3})$.

Indeed, the second inequality does NOT hold! The mistake comes from the last expression in their proof, which missed a negative sign. So the inequality just holds exactly on the opposite direction.

Based on this lemma, they proved that if $\alpha_i \in (0, \frac{\pi}{3}]$ and $\sum_{i=1}^{k} \alpha_i \leq 2\pi$, then the sticking-out area $\sum_{i=1}^{k} S(\alpha_i) \leq 2\sqrt{3} - \pi$.

But actually this statement is false. For a counter example, we choose $\alpha_i = \frac{\pi}{4}, i = 1, \ldots, 8$, then

$$\sum_{i=1}^{k} S(\alpha_i) = 0.4625 > 2\sqrt{3} - \pi = 0.3225.$$

Clearly, the upper bound of sum of squares of lengths of the edges in EMST in unit disk couldn't be proved to be 12 in this way.

Wan also discovered their mistake. They fixed the bug and presented a slightly larger upper bound of 12.141. However, there are still enough rooms to improve the upper bound. We will introduce another method in later sections, and get our new bound 10.86.

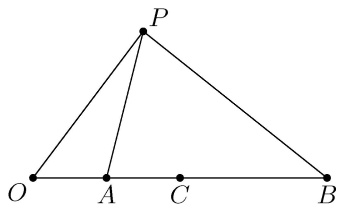


Figure 3. Illustration for Lemma 2.

4. Upper bound of sum of squares of lengths of the edges in EMST

Lemma 2. *Let O, A, C and B be collinear in the order, with $\|OC\|^2 = \|OA\| \cdot \|OB\|$ (see figure 3). Then for any point P on the plane, we have*

$$\frac{\|PA\|}{\|PB\|} \begin{cases} < \dfrac{\|CA\|}{\|CB\|} & \textit{if and only if } \|OP\| < \|OC\|, \\ = \dfrac{\|CA\|}{\|CB\|} & \textit{if and only if } \|OP\| = \|OC\|, \\ > \dfrac{\|CA\|}{\|CB\|} & \textit{if and only if } \|OP\| > \|OC\|. \end{cases}$$

Proof: It's easy to verify the correctness of the conclusion if P is on straight line AB. And if not, we consider the three cases respectively:

If $\|OP\| = \|OC\|$, then

$$\|OP\|^2 = \|OA\| \cdot \|OB\|$$

Therefore,

$$\triangle OAP \sim \triangle OPB$$

We have

$$\angle BPC = \angle OCP - \angle OBP = \angle OPC - \angle OPA = \angle APC$$

which implies

$$\frac{\|PA\|}{\|PB\|} = \frac{\|CA\|}{\|CB\|}$$

If $\|OP\| < \|OC\|$, let P' be on the ray AP such that $\|OP'\| = \|OC\|$ and B' be on the ray AB such that PB is parallel to $P'B'$, then $\|AP'\| > \|AP\|$ and $\angle ACP'$ is acute (see

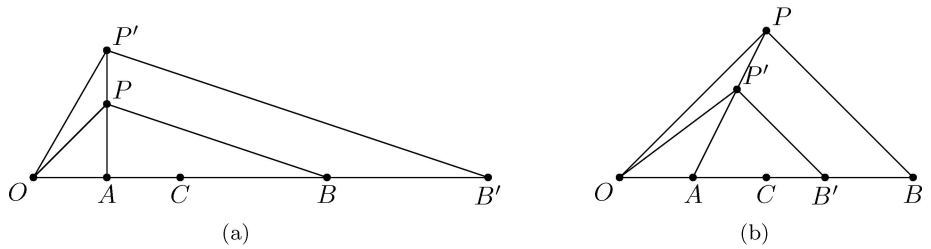


Figure 4. The case that (a) $\|OP\| < \|OC\|$ and (b) $\|OP\| > \|OC\|$.

figure 4(a)). So we have

$$\angle PBB' > \angle P'CB > \angle P'CA > \angle P'B'A$$

That implies $\|P'B'\| > \|P'B\|$, so

$$\frac{\|PA\|}{\|PB\|} = \frac{\|P'A\|}{\|P'B'\|} < \frac{\|P'A\|}{\|P'B\|} = \frac{\|CA\|}{\|CB\|}$$

If $\|OP\| > \|OC\|$, let P' be on the ray AP such that $\|OP'\| = \|OC\|$ and let B' be on the ray AB such that PB is parallel to $P'B'$, then $\|AP'\| < \|AP\|$ and $\angle ACP'$ is acute (see figure 4(b)). So we have

$$\angle PB'B > \angle P'CB > \angle P'CA > \angle P'BA$$

That implies $\|P'B\| > \|P'B'\|$, so

$$\frac{\|PA\|}{\|PB\|} = \frac{\|P'A\|}{\|P'B'\|} > \frac{\|P'A\|}{\|P'B\|} = \frac{\|CA\|}{\|CB\|}$$

The proof is complete. □

Lemma 3. *Let A_1A_2 and B_1B_2 be any two edges in EMST satisfying $\|A_1B_i\| \geq \|A_1A_2\|$, $i = 1, 2$ (see figure 5), then*

$$D(B_1B_2) \cap B\left(A_1, \frac{\sqrt{3}}{3}\|A_1A_2\|\right) = \emptyset$$

Proof: We will prove that any point on the boundary of $D(B_1B_2)$ is at least $\frac{\sqrt{3}}{3}\|A_1A_2\|$ away from A_1. Without loss of generality, assume that $\|A_1B_1\| \geq \|A_1B_2\|$. Then by the properties of EMST (Wan et al., 2001), we know that $\|A_1B_1\| \geq \|B_1B_2\|$. Let B_3 be any vertex of $D(B_1B_2)$ other than B_1 and B_2. Let P be any point on either line segment B_1B_3

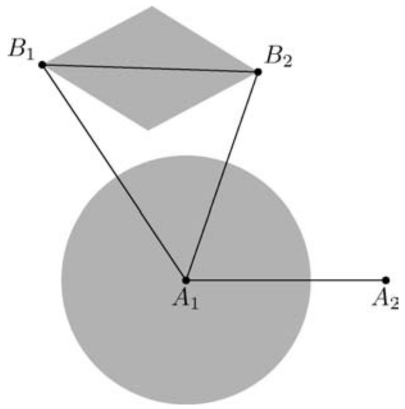


Figure 5. Illustration for Lemma 3.

or line segment B_2B_3, and D be the point on line segment B_2P such that $\frac{\|B_2D\|}{\|DP\|} = \sqrt{3}$. Let O be the point on the extending line of B_2P such that $\|PO\| = \frac{1}{2}\|B_2P\|$. Then it's easy to verify that

$$\|OD\|^2 = \|OP\| \cdot \|OB_2\|$$

Case 1: if P is on line segment B_1B_3, let C be the point on the same side of line B_1B_2 as B_3 such that $\triangle B_1B_2C$ is equilateral (see figure 6(a)). Notice that

$$\frac{\|PC\|}{\|B_2C\|} \geq \frac{\|B_3C\|}{\|B_2C\|} = \frac{\sqrt{3}}{3} = \frac{\|PD\|}{\|B_2D\|}.$$

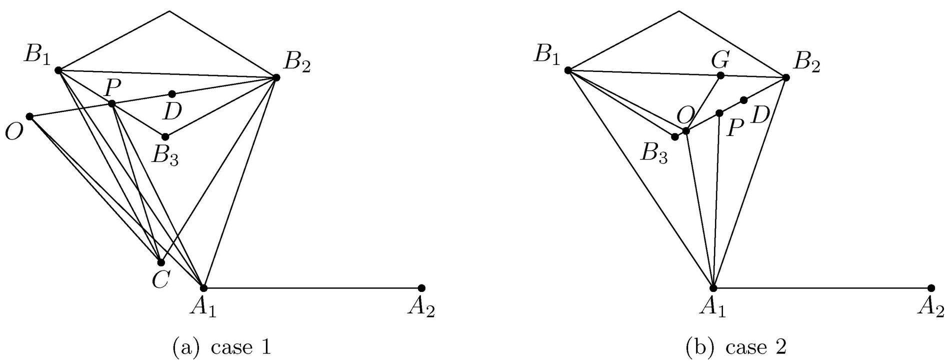


Figure 6. Two cases in Lemma 3.

Then from Lemma 2, we have $\|OC\| \geq \|OD\|$.

In addition, because $\angle OCA_1 \geq \angle B_1CA_1 \geq \angle B_1A_1C \geq \angle OA_1C$, we have

$$\|OA_1\| \geq \|OC\| \geq \|OD\|.$$

Case 2: if P is on line segment B_2B_3, let G be the point on line segment B_1B_2 such that $\|B_2G\| = \|OG\|$ (see figure 6(b)). Then $\angle B_1GO = 2\angle B_1B_2O = \frac{\pi}{3}$, so

$$\angle B_1OG = \frac{2\pi}{3} - \angle GB_1O \geq \frac{\pi}{3} = \angle B_1GO.$$

In other words, $\|GB_1\| \geq \|OB_1\|$. Therefore,

$$\begin{aligned}\|OA_1\| &\geq \|A_1B_1\| - \|OB_1\| \geq \|B_1B_2\| - \|OB_1\| \geq \|B_1B_2\| - \|GB_1\| \\ &= \frac{\sqrt{3}}{3}\|OB_2\| = \|OD\|.\end{aligned}$$

We can see that in both cases, $\|OA_1\| \geq \|OD\|$. From Lemma 2, we know that

$$\|A_1P\| \geq \frac{\sqrt{3}}{3}\|A_1B_2\| \geq \frac{\sqrt{3}}{3}\|A_1A_2\|$$

In total, the distances from A_1 to any boundary points of $D(B_1B_2)$ are no smaller than $\frac{\sqrt{3}}{3}\|A_1A_2\|$. Clearly, $D(B_1B_2)$ is outside $B(A_1, \frac{\sqrt{3}}{3}\|A_1A_2\|)$. □

Corollary 1. *Let A_1A_2 and B_1B_2 be any two edges in EMST, then $\exists i \in \{1, 2\}$, s.t.*

$$D(B_1B_2) \cap B\left(A_i, \frac{\sqrt{3}}{3}\|A_1A_2\|\right) = \emptyset.$$

Proof: Only to notice that B_1 and B_2 are either both outside $B(A_1, \|A_1A_2\|)$ or both outside $B(A_2, \|A_1A_2\|)$. □

Lemma 4. *Let A_1A_2, B_1B_2 and C_1C_2 be any three edges in EMST(P) (see figure 7), then at least one of following three statements holds*:

(1) $\|A_1B_i\| \geq \|A_1A_2\|, i = 1, 2$
(2) $\|A_2C_i\| \geq \|A_1A_2\|, i = 1, 2$
(3) $\|B_iC_j\| \geq \|A_1A_2\|, i, j = 1, 2$

Proof: If none of these three statements holds, then $\exists i, j, k, l \in \{1, 2\}$, s.t. $\|A_1B_i\| < \|A_1A_2\|$, $\|A_2C_j\| < \|A_1A_2\|$, and $\|B_kC_l\| < \|A_1A_2\|$.

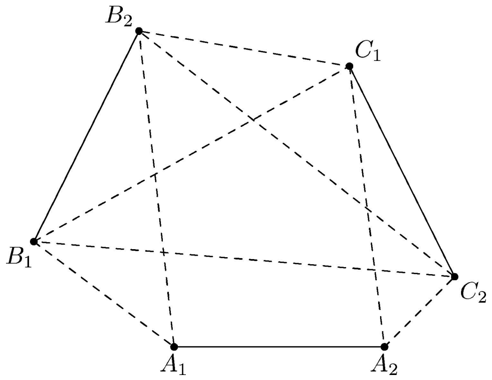


Figure 7. Illustration for Lemma 4.

Let T be the Euclidean minimum spanning tree of point set P, then graph $T \setminus \{A_1A_2\}$ has two connected components. Clearly, B_1 and B_2 are in the same connected component, C_1 and C_2 are in the same connected component, but A_1 and A_2 are in different connected components. So we have three cases to consider:

Case 1: One connected component contains A_1, and the other contains A_2, B_1, B_2, C_1, C_2. Then $(T \setminus \{A_1A_2\}) \cup \{A_1B_i\}$ contains $|P| - 1$ edges and no cycles, thus is a spanning tree of P, and has a total length less than T. This causes a contradiction.

Case 2: One connected component contains A_1, B_1, B_2, and the other contains A_2, C_1, C_2. Then $(T \setminus \{A_1A_2\}) \cup \{B_kC_l\}$ is a spanning tree of P and has a total length less than T. A contradiction.

Case 3: One connected component contains A_1, B_1, B_2, C_1, C_2, and the other contains A_2. Then $(T \setminus \{A_1A_2\}) \cup \{A_2C_j\}$ is a spanning tree of P and has a total length less than T. A contradiction.

In total, every case causes a contradiction. This completes the proof. □

Now we come to the main lemma in this paper.

Lemma 5. *Let A_1A_2, B_1B_2 and C_1C_2 be any three edges in EMST. Let $\alpha_0 = \sup\{x \in [0, \frac{\pi}{6}] : E_{x,0}(A_1A_2) \cap D(B_1B_2) = \emptyset\}$ and $\beta_0 = \sup\{y \in [0, \frac{\pi}{6}] : E_{0,y}(A_1A_2) \cap D(C_1C_2) = \emptyset\}$. Then*

$$\alpha_0 + \beta_0 \geq \frac{\pi}{6}.$$

Proof: If $\alpha_0 + \beta_0 < \frac{\pi}{6}$, we will derive a contradiction.

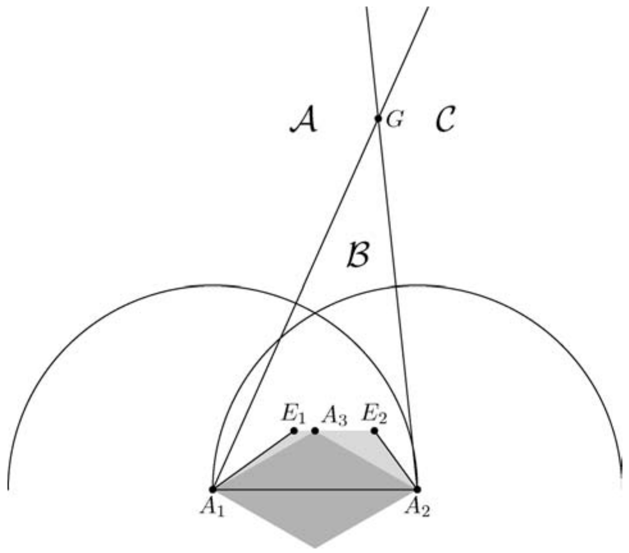


Figure 8. The area partition.

Choose $\alpha, \beta \in [0, \frac{\pi}{6}]$, s.t. $\alpha > \alpha_0$, $\beta > \beta_0$, and $\alpha + \beta = \frac{\pi}{6}$, then

$$E_{\alpha,0}(A_1A_2) \cap D(B_1B_2) \neq \emptyset$$

and

$$E_{0,\beta}(A_1A_2) \cap D(C_1C_2) \neq \emptyset.$$

Let G be the point on the same side of A_1A_2 as A_3 such that $\angle GA_1A_2 = \frac{\pi}{3} + \alpha$ and $\angle GA_2A_1 = \frac{\pi}{3} + \beta$. We partition the area where B_1,B_2,C_1 or C_2 can possibly be in three regions: $\mathcal{A}$,$\mathcal{B}$ and $\mathcal{C}$, where $\mathcal{A}$ is the region on the left side of line A_1G, $\mathcal{B}$ is the region inside $\triangle A_1A_2G$, and $\mathcal{C}$ is the region on the right side of line A_2G (see figure 8).

Let a denote $\|A_1A_2\|$. Now we consider every case according to which region each point lies in(note that $\mathcal{A}$ and $\mathcal{C}$ share some area so we can treat a point in this area either in $\mathcal{A}$ or in $\mathcal{C}$), and prove that none of these cases would happen. Without loss of generality, assume $\angle B_1A_1A_2 \geq \angle B_2A_1A_2$ and $\angle C_1A_2A_1 \leq \angle C_2A_2A_1$.

Case 1: B_2 lies in region $\mathcal{A}$. Then B_1 also lies in $\mathcal{A}$. Note that $\angle E_1A_1G = \frac{\pi}{6}$, by Lemma 1, $D(B_1B_2)$ is on the left side of line A_1E_1, which contradicts $E_{\alpha,\beta}(A_1A_2) \cap D(B_1B_2) \neq \emptyset$. Similarly, there would also be a contradiction if C_1 lies in region $\mathcal{C}$.

Case 2: B_2 lies in region $\mathcal{B}$. There are two subcases.

Case 2.1: At least one of C_1 and C_2 lie in region $\mathcal{B}$. Without loss of generality, assume C_1 lies in $\mathcal{B}$. Let point H_1 and H_2 respectively be the point on line segment GA_1 and GA_2, such that $\|A_2H_1\| = \|A_1H_2\| = a$.

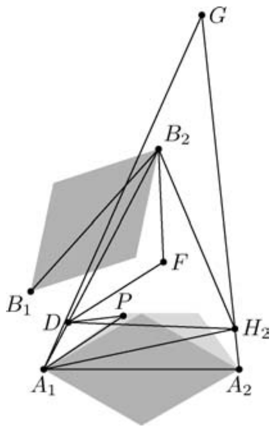


Figure 9. Illustration for case 2.1.

Now if $\|B_2H_2\| \geq a$, let D be the point on line segment A_1B_2 such that $\|B_2D\| = \|B_2H_2\|$. Let F be the vertex of $D(B_2D)$ on the same side of B_2D as A_2. Let P be any point inside $E_{\alpha,0}(A_1A_2)$ (see figure 9).

Since $\|DH_2\| \leq a \leq \|DB_2\| = \|H_2B_2\|$, we have $\angle B_2DH_2 \geq \frac{\pi}{3}$, therefore

$$\angle A_1DH_2 = \pi - \angle H_2DB_2 \leq \pi - \frac{\pi}{3} = \frac{2\pi}{3}.$$

On the other hand, P must be in $D(A_1H_2)$ since

$$\begin{aligned}\angle PA_1H_2 &= \angle GA_1A_2 - \angle GA_1P - \angle A_2A_1H_2 \\ &< \left(\alpha + \frac{\pi}{3}\right) - \frac{\pi}{6} - \left[\pi - 2\left(\beta + \frac{\pi}{3}\right)\right] = \beta < \frac{\pi}{6}.\end{aligned}$$

We have

$$\angle A_1PH_2 > \frac{2\pi}{3} \geq \angle A_1DH_2.$$

In other words, D is outside the circumcircle of $\triangle A_1PH_2$, which means $\angle PDH_2 < \angle PA_1H_2 < \frac{\pi}{6}$. Now,

$$\angle B_2DP \geq \angle B_2DH_2 - \angle PDH_2 > \frac{\pi}{3} - \frac{\pi}{6} = \frac{\pi}{6}.$$

Note the arbitrariness of P, $E_{\alpha,0}(A_1A_2)$ is outside $\angle DFB_2$.

By Lemma 4, $\|B_1B_2\| \leq \|B_2C_1\| \leq \|B_2H_2\| = \|B_2D\|$. Because $\angle B_1B_2F \geq \angle DB_2F$, it's easy to see that $D(B_1B_2) \subset \angle DFB_2$, which implies $E_{\alpha,0}(A_1A_2) \cap D(B_1B_2) = \emptyset$. This is a contradiction. Therefore, $\|B_2H_2\| < a$.

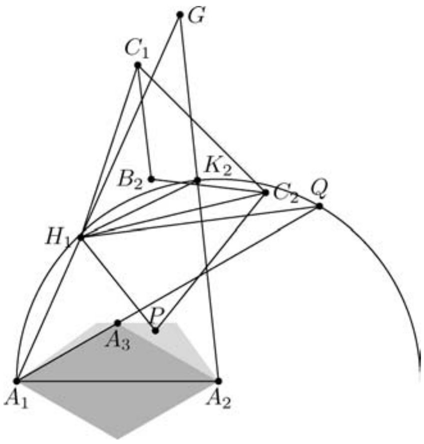


Figure 10. Illustration for case 2.2.

By the same reason we have $\|C_1H_1\| < a$.

Furthermore, from Lemma 3, $\exists i, j$, s.t. $\|A_1B_i\| < a$ and $\|A_2C_j\| < a$. So by the properties of EMST, B_1 and B_2 must be both outside $B(A_2, a)$, and C_1 and C_2 must be both outside $B(A_1, a)$. Finally, $B_2 \in \mathcal{B} \cap B(H_2, a) \setminus B(A_2, a)$ and $C_1 \in \mathcal{B} \cap B(H_1, a) \setminus B(A_1, a)$. Now, by simple analysis we can find that wherever B_2 or C_1 be, $\|B_2C_1\| < a$ always hold. That contradicts Lemma 4.

Case 2.2: Now C_1 must be in region $\mathcal{A}$ and C_2 must be in region $\mathcal{C}$(because they can't be both in $\mathcal{A}$ or both in $\mathcal{C}$).

Let K_2 be the point on line segment A_2G such that $\|A_2K_2\| = a$ (see figure 10).

Because A_1,H_1 and K_2 are all on $\odot A_2$, we have

$$\angle GH_1K_2 = \frac{\angle GA_2A_1}{2} \geq \frac{\pi}{6}.$$

And by Lemma 5, $C_2 \in B(A_2, a)$, so C_2 is on the right side of line H_1K_2. We have $\angle C_1H_1C_2 \geq \angle GH_1K_2$. Furthermore, B_2 is below line C_1C_2, for otherwise by the fact that $D(B_1B_2)$ and $D(C_1C_2)$ are disjoint, $D(B_1B_2)$ would be outside quadrilateral $A_1C_1C_2A_2$, thereby having no intersection with $E_{\alpha,0}(A_1A_2)$, which makes a contradiction. So B_2 is inside $\triangle C_1H_1C_2$, and

$$\angle C_1B_2C_2 \geq \angle C_1H_1C_2 \geq \angle GH_1K_2 \geq \frac{\pi}{6}.$$

By Lemma 4, $\|B_2C_1\| \geq \|C_1C_2\|$ and $\|B_2C_2\| \geq \|C_1C_2\|$, so $\angle C_1B_2C_2$ is the smallest angle in $\triangle B_2C_1C_2$. Now we have

$$\angle C_1C_2H_1 \geq \angle C_1C_2B_2 \geq \angle C_1B_2C_2 \geq \frac{\pi}{6}.$$

Let P be any point inside $E_{0,\beta}(A_1A_2)$. Let A_3 be the vertex of $D(A_1A_2)$ on the upper side of A_1A_2 and Q be the intersection point of the extending line of A_1A_3 and $\odot A_2$. Then $\angle GH_1P \geq \angle GH_1Q = \frac{\pi}{3}$. Therefore,

$$\angle GH_1C_2 \leq C_1H_1C_2 \leq C_1B_2C_2 \leq \frac{\pi}{3} \leq \angle GH_1P.$$

So P is below line H_1C_2, and we have $\angle C_1C_2P \geq \angle C_1C_2H_1 \geq \frac{\pi}{6}$. This implies $E_{0,\beta}(A_1A_2) \cap D(C_1C_2) = \emptyset$. A contradiction.

Case 3: B_2 lies in region $\mathcal{C}$, then B_1 must lie in $\mathcal{A}$ or $\mathcal{B}$.

Case 3.1: C_1 lies in region $\mathcal{B}$. This case is the same with case 2 by symmetry.
Case 3.2: C_1 is in region $\mathcal{A}$, then C_2 must lie in $\mathcal{B}$ or $\mathcal{C}$.

Now either of following two statements must hold, for otherwise line segment B_1B_2 and C_1C_2 would have intersection, which contradicts the properties of EMST.

(1) line segment B_1B_2 is outside quadrilateral $A_1C_1C_2A_2$, or
(2) line segment C_1C_2 is outside quadrilateral $A_1B_1B_2A_2$.

Without loss of generality, assume that statement 1 holds. Then by the fact that $D(B_1B_2)$ and $D(C_1C_2)$ are disjoint, $D(B_1B_2)$ would be outside quadrilateral $A_1C_1C_2A_2$, thereby having no intersection with $E_{\alpha,0}(A_1A_2)$. This is a contradiction.

In conclusion, $\alpha_0 + \beta_0 \geq \frac{\pi}{6}$, and $E_{\alpha_0,\beta_0}(A_1A_2) \cap D(B_1B_2) = \emptyset$, $E_{\alpha_0,\beta_0}(A_1A_2) \cap D(C_1C_2) = \emptyset$. This completes the proof. □

Theorem 2. *Let A_1A_2 be an edge in EMST, then $\exists \alpha_0, \beta_0 \geq 0$, with $\alpha_0 + \beta_0 \geq \frac{\pi}{6}$, such that $E_{\alpha_0,\beta_0}(A_1A_2)$ is disjoint from every diamond determined by any edges in EMST.*

Proof: Let

$$\mathcal{S} = \{(B_1B_2) \in \text{EMST} : E_{\frac{\pi}{6},0}(A_1A_2) \cap D(B_1B_2) \neq \emptyset\}$$

and

$$\mathcal{T} = \{(C_1C_2) \in \text{EMST} : E_{0,\frac{\pi}{6}}(A_1A_2) \cap D(C_1C_2) \neq \emptyset\}.$$

Then from Corollary 1, $\mathcal{S} \cap \mathcal{T} = \emptyset$. Let

$$\alpha_0 = \min_{(B_1B_2)\in\mathcal{S}} \left\{ \sup \left\{ x \in \left[0, \frac{\pi}{6}\right] : E_{x,0}(A_1A_2) \cap D(B_1B_2) = \emptyset \right\} \right\}$$

and

$$\beta_0 = \min_{(C_1C_2)\in\mathcal{T}} \left\{ \sup \left\{ y \in \left[0, \frac{\pi}{6}\right] : E_{0,y}(A_1A_2) \cap D(C_1C_2) = \emptyset \right\} \right\}.$$

Then $E_{\alpha_0,\beta_0}(A_1A_2)$ is disjoint from every diamond determined by any edges in EMST. And from Lemma 5,

$$\alpha_0 + \beta_0 \geq \frac{\pi}{6}.$$

□

By Theorem 2, we can assign an α_1, β_1 and α_2, β_2 value for each edge A_1A_2 in EMST such that $E(A_1A_2) \triangleq E_{\alpha_1,\beta_1}(A_1A_2) \cup E_{\alpha_2,\beta_2}(A_2A_1)$ is disjoint from diamonds determined by any other edges in EMST(We call it the extra area of edge A_1A_2). The following theorem states that the extra area of all the edges in EMST would not overlap more than twice in any place.

Theorem 3. *Let A_1A_2, B_1B_2 and C_1C_2 be any three edges in EMST, then*

$$E(A_1A_2) \cap E(B_1B_2) \cap E(C_1C_2) = \emptyset.$$

Proof: Assume $\exists P \in E(A_1A_2) \cap E(B_1B_2) \cap E(C_1C_2)$, we will derive a contradiction.

Let A_i, B_i and C_i, $i = 1, \ldots, 4$ respectively denote the four vertices of $D(A_1A_2)$, $D(B_1B_2)$ and $D(C_1C_2)$. It's easy to see that there exists a side of each of $D(A_1A_2)$, $D(B_1B_2)$ and $D(C_1C_2)$, without loss of generality denoted by A_1A_3, B_1B_3 and C_1C_3, s.t. $\triangle A_1PA_3$, $\triangle B_1PB_3$ and $\triangle C_1PC_3$ is respectively contained in $E(A_1A_2)$, $E(B_1B_2)$ and $E(C_1C_2)$.

Because every extra area is disjoint from every diamond, we can prove that $\angle A_1PA_3$, $\angle B_1PB_3$ and $\angle C_1PC_3$ are all disjoint. In fact, if any two of them are not disjoint (assume they are $\angle A_1PA_3$ and $\angle B_1PB_3$), then one of following two statements must hold.

(1) A_1 or A_3 are inside $\triangle B_1PB_3$, or
(2) B_1 or B_3 are inside $\triangle A_1PA_3$.

Without loss of generality, assume statement 1 holds. Then $D(A_1A_2) \cap E(B_1B_2) \neq \emptyset$, which is a contradiction.

Therefore, $\angle A_1PA_3 + \angle B_1PB_3 + \angle C_1PC_3$ is no more than 2π. But clearly they are all greater than $\frac{2\pi}{3}$. This is a contradiction, which completes the proof. □

Now we can describe how to estimate the upper bound of $\sum_{e\in\mathrm{EMST}(P)} \|e\|^2$ using Theorems 1–3.

Let P be any finite point set in unit disk. Construct diamonds and extra areas for every edge in EMST(P) according to previous statements. Notice that for some edge e, $D(e)$ or $E(e)$ may exceed the unit disk. We denote these out-of-disk areas of $D(e)$ and $E(e)$ by

$Ex(D(e))$ and $Ex(E(e))$ respectively. Because all diamonds are disjoint, the total inside-disk area of the diamonds equals $\sum_{e\in \text{EMST}(P)} Area(D(e)) - Ex(D(e))$. And according to Theorem 3, all the extra areas cannot overlap more than twice at any location. So the total inside-disk area of the extra areas is at least $\frac{1}{2}\sum_{e\in \text{EMST}(P)} Area(E(e)) - Ex(E(e))$.

For an edge $e \in \text{EMST}(P)$, let the sticking-out area of e be

$$T(e) = Ex(D(e)) + \frac{Ex(E(e))}{2}, \tag{2}$$

Because the area of the unit disk is π, it's easy to see that

$$\pi \geq \sum_{e\in \text{EMST}(P)} Area(D(e)) + \frac{1}{2}\sum_{e\in \text{EMST}(P)} Area(E(e)) - \sum_{e\in \text{EMST}(P)} T(e).$$

For an edge $e \in \text{EMST}(P)$, it's easy to compute that $Area(D(e)) = \frac{\sqrt{3}}{6}\|e\|^2$, and $Area(E(e)) \geq 2Area(E_{\frac{\pi}{6},0}(e)) = \frac{\sqrt{3}}{18}\|e\|^2$. Therefore,

$$\sum_{e\in \text{EMST}(P)} \frac{\sqrt{3}}{6}\|e\|^2 + \frac{1}{2}\sum_{e\in \text{EMST}(P)} \frac{\sqrt{3}}{18}\|e\|^2 - \sum_{e\in \text{EMST}(P)} T(e) \leq \pi.$$

In other words,

$$\sum_{e\in \text{EMST}(P)} \|e\|^2 \leq \frac{\pi + \sum_{e\in \text{EMST}(P)} T(e)}{\frac{7\sqrt{3}}{36}}. \tag{3}$$

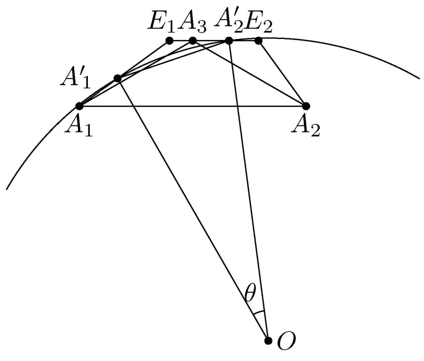
This is how we will estimate the upper bound of sum of squares of lengths of the edges in EMST in unit disk. As what we mentioned when describing the mistakes in Wan et al. (2001), how to effectively calculate $\sum_{e\in \text{EMST}(P)} T(e)$ plays an important role in this estimation. So we need one more lemma.

Lemma 6.

$$\sum_{e\in \text{EMST}(P)} T(e) \leq 0.5166$$

Proof: Let A_1A_2 be an edge in EMST which has positive sticking-out area. It's easy to see that the boundary of $E(A_1A_2)$ would have two intersection points with the boundary of the unit disk(the unit circle). Assume they are A_1' and A_2' in clock-wise order on the unit circle. Let $\theta_{A_1A_2} = \angle A_1' O A_2'$ where O is the center of the unit disk, then $\|A_1'A_2'\| = 2\sin\frac{\theta}{2}$. And it's easy to see that $\theta_{A_1A_2} \leq \frac{\pi}{3}$ since $\|A_1A_2\| \leq 1$. Define function

$$S(x) = \frac{\sin x}{2} + \left(\frac{\sqrt{3}}{4} - \frac{1}{12}\right)(1 - \cos x) - \frac{x}{2}. \tag{4}$$



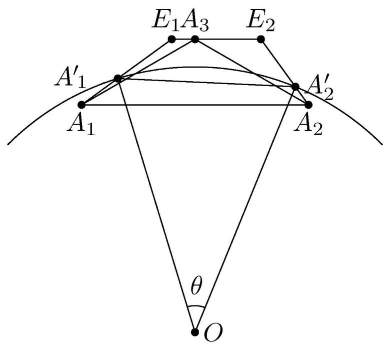


Figure 11. Illustration for Lemma 6.

Claim 1: $T(A_1A_2) \leq S(\theta_{A_1A_2})$.

Clearly only one side of $D(A_1A_2) \cup E(A_1A_2)$ from A_1A_2 could possibly stick out the unit disk. Without loss of generality, assume the sticking-out area is on the same side of A_1A_2 as A_3. There are two cases to consider:

Case 1: A_1' and A_2' are respectively on line segment A_1E_1(or A_2E_2) and line segment E_1E_2. Without loss of generality, assume that A_1' is on line segment A_1E_1 and A_2' is on line segment E_1E_2 (see figure 11(a)). Then

$$\begin{aligned} Ex(D(A_1A_2)) &\leq Ex(D(A_1'A_2')) \\ &= \frac{\sqrt{3}}{12}\|A_1'A_2'\|^2 + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2} \\ &= \frac{\sqrt{3}}{6}(1-\cos\theta_{A_1A_2}) + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2}, \end{aligned}$$

and because $\angle A_1'E_1A_2' \geq \frac{2\pi}{3}$, we have

$$\begin{aligned} Ex(D(A_1A_2) \cup E(A_1A_2)) &= Area(\triangle A_1'E_1A_2') + Area(\triangle A_1'A_2'O) \\ &\quad - Area(Sector A_1'A_2'O) \\ &\leq \frac{1}{4}\tan\frac{\pi}{6}\cdot\|A_1'A_2'\|^2 + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2} \\ &= \frac{\sqrt{3}}{6}(1-\cos\theta_{A_1A_2}) + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2}. \end{aligned}$$

Case 2: A_1' and A_2' are respectively on line segment A_1E_1 and line segment A_2E_2 (see figure 11(b)). Then

$$Ex(D(A_1A_2)) \le Ex(D(A_1'A_2')) = \frac{\sqrt{3}}{6}(1-\cos\theta_{A_1A_2}) + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2},$$

and

$$\begin{aligned} Ex(D(A_1A_2) \cup E(A_1A_2)) &= Area(Quad A_1'E_1E_2A_2') + Area(\triangle A_1'A_2'O) \\ &\quad - Area(Sector A_1'A_2'O) \\ &\le \frac{1}{2}Area(D(A_1'A_2') \cup E(A_1'A_2')) + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2} \\ &= \frac{2\sqrt{3}-1}{6}(1-\cos\theta_{A_1A_2}) + \frac{\sin\theta_{A_1A_2}}{2} - \frac{\theta_{A_1A_2}}{2}. \end{aligned}$$

Therefore, in both cases, we have

$$\begin{aligned} T(A_1A_2) &= Ex(D(A_1A_2)) + \frac{Ex(E(A_1A_2))}{2} \\ &= \frac{1}{2}[Ex(D(A_1A_2)) + Ex(D(A_1A_2) \cup E(A_1A_2))] \\ &\le \frac{\sin\theta_{A_1A_2}}{2} + \left(\frac{\sqrt{3}}{4} - \frac{1}{12}\right)(1-\cos\theta_{A_1A_2}) - \frac{\theta_{A_1A_2}}{2}. \end{aligned}$$

Thus Claim 1 holds.

Claim 2: Suppose $e_1, \ldots, e_k$ are all edges in EMST whose diamonds or extra areas stick out the unit disk, then

$$\sum_{i=1}^{k} \theta_{e_i} \le 2\pi.$$

Assume A_1A_2 and B_1B_2 are two edges in EMST which have positive sticking-out areas. A_1',A_2'(B_1',B_2' respectively) are the intersection points of the unit circle and the boundary of $E(A_1A_2)$($E(B_1B_2)$ respectively). It's easy to see that arc $A_1'A_2'$(the part of unit circle between A_1' and A_2')$\subset D(A_1A_2) \cup E(A_1A_2)$ and $B_1'B_2' \subset D(B_1B_2) \cup E(B_1B_2)$.

Then we can prove that arc $A_1'A_2'$ and arc $B_1'B_2'$ are disjoint. In fact, if there exists point P which is both on arc $A_1'A_2'$ and arc $B_1'B_2'$, then $P \in E(A_1A_2) \cap E(B_1B_2)$. By the similar reason as in the proof of Theorem 3, $\angle A_1PA_2$ and $\angle B_1PB_2$ are disjoint. Because $\exists i, j \in \{1, 2\}$, s.t. $\angle A_1PA_2 \cup \angle B_1PB_2 \subset \angle A_iPB_j$, we have $\angle A_1PA_2 + \angle B_1PB_2 \le \angle A_iPB_j$.

Since P is on the unit circle, and A_i and B_j are on or inside the unit circle, $\angle A_iPB_j < \pi$. But on the other hand, clearly $\angle A_1PA_2$ and $\angle B_1PB_2$ are both $\ge \frac{\pi}{2}$, so

$$\angle A_1PA_2 + \angle B_1PB_2 \ge \pi > \angle A_iPB_j.$$

This is a contradiction. Thus Claim 2 holds.

Now our task is equivalent to estimating the upper bound of following summation.

$$\sum_{i=1}^{k} S(x_i),$$
$$\text{s.t. } \sum_{i=1}^{k} x_i \le 2\pi,$$

where each $x_i \in \left(0, \frac{\pi}{3}\right]$.

To estimate this upper bound, we first analyze the convexity of function $S(x)$. Differentiate $S(x)$ two times, we can find that its second derivative satisfies

$$S''(x)\begin{cases} > 0, & \text{if } 0 < x < c, \\ \le 0, & \text{if } c \le x \le \frac{\pi}{3}, \end{cases}$$

where $c = \arctan \frac{3\sqrt{3}-1}{6} \approx 0.6103$.

In other words, $S(x)$ is convex if $0 < x < c$ and concave if $c \le x \le \frac{\pi}{3}$. By the properties of convex functions, we can prove following results:

1. for any $x_1, x_2 \in (0, c)$,

$$S(x_1) + S(x_2) \le \begin{cases} S(x_1 + x_2), & \text{if } x_1 + x_2 < c, \\ S(c) + S(x_1 + x_2 - c), & \text{if } x_1 + x_2 \ge c. \end{cases}$$

2. for any $x_1, \dots, x_m \in [c, \frac{\pi}{3}]$,

$$S(x_1) + \cdots + S(x_m) \le mS\left(\frac{x_1 + \cdots + x_m}{m}\right).$$

It is easy to see that Result 2 holds. To prove Result 1, without loss of generality, we assume $0 \le x_1 \le x_2 \le c$. Then if $x_1 + x_2 < c$, we have

$$\frac{S(x_1 + x_2) - S(x_2)}{x_1} \ge S'(x_2) \ge S'(x_1) \ge \frac{S(x_1) - S(0)}{x_1} = \frac{S(x_1)}{x_1}.$$

And if $x_1 + x_2 \ge c$, we have

$$\frac{S(c) - S(x_2)}{c - x_2} \ge S'(x_2) \ge S'(x_1) \ge \frac{S(x_1) - S(x_1 + x_2 - c)}{c - x_2}.$$

Thus Result 1 holds.

Now for $x_i \in (0, \frac{\pi}{3}]$, $i = 1, \ldots, k$, s.t. $\sum_{i=1}^{k} x_i \leq 2\pi$. By repeatedly applying the two inequalities in Result 1, the number of x_i such that $x_i < c$ can be reduced to at most 1. So without loss of generality, we can assume $x_1 < c$ and $x_2, \ldots, x_n \geq c$. Then by Result 2,

$$\sum_{i=1}^{k} S(x_i) \leq S(x_1) + (k-1)S\left(\frac{x_2 + \cdots + x_k}{k-1}\right).$$

Let $x_0 = \frac{x_2+\cdots+x_k}{k-1}$, we have

$$\sum_{i=1}^{k} S(x_i) \leq S(x_1) + (x_2 + \cdots + x_n)\frac{S(x_0)}{x_0} \leq S(x_1) + (2\pi - x_1)\frac{S(x_0)}{x_0}.$$

By solving the equation $\frac{d}{dx}\frac{S(x)}{x} = 0$, we can find that

$$\min_{x \in [c, \pi/3]} \frac{S(x)}{x} = \frac{S(0.914)}{0.914} = 0.0822.$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{k} S(x_i) &\leq S(x_1) + 0.0822(2\pi - x_1)\\ &\leq S(0) + 0.0822(2\pi - 0)\\ &= 0.5166.\end{aligned}$$

The second inequality holds because the right hand side of the first inequality is an increasing function of x_1 on the interval $(0, c)$ (this can be verified by differentiating it). □

Theorem 4.

$$6 \leq c \leq 10.86.$$

Proof: From (3) and Lemma 6, we have

$$\sum_{e \in \text{EMST}(P)} \|e\|^2 \leq \frac{\pi + 0.5166}{\frac{7\sqrt{3}}{36}} = 10.86.$$

The lower bound 6 is achieved by letting P be the six vertices of the regular hexagon of side length 1. □

5. Future works

In many parts of our proof, there are still rooms to improve the upper bound. But the proof may become too intricate, and the improvement we shall get may be rather small. Finding other approaches may be a better way.

So far, people still cannot construct any instances of P that has a $\sum_{e \in \mathrm{EMST}(P)} \|e\|^2$ value of more than 6. So, it is very likely that $c = 6$. Is it possible to prove this? There is a max-min lemma included in the proof of Gilbert-Pollak conjecture, applicable to this max-min problem (Du and Hwang, 1992). However, the help is quite limited due to the circle constraint.

We can also consider that how to generalize the problem to higher dimensional spaces.

References

D.-Z. Du and F.K. Hwang, "The state of art on Steiner ratio problems," in *Computing in Euclidean Geometry*, D.-Z. Du and F.K. Hwang (eds.), World Scientific, 1992.

F.P. Preparata and M.I. Shamos, *Computational Geometry: an Introduction*, Springer-Verlag, 1985.

T.S. Rappaport, *Wireless Communications Principles and Practices*, Prentice Hall, 1996.

P.J. Wan, G. Calinescu, X.Y. Li, and O. Frieder, "Minimum-energy broadcast routing in static ad hoc wireless networks," in *IEEE INFOCOM '2001*, 2001.

J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," in *IEEE INFOCOM '2000*, 2000.