Approximate minimum spanning ellipse in the streaming model

Asish Mukhopadhyay*and Eugene Greene†
and Animesh Sarker‡
and Tom Switzer

Abstract

In [11] Zarrabi-Zadeh and Chan proposed an algorithm to compute an approximate minimum spanning ball in the streaming model of computation and showed by an elegant analysis that the radius of the approximate ball is to within 3/2 of the exact one. Inspired by this result, in this paper we consider the extension of this result to computing approximate minimum spanning ellipses. The ball algorithm is simple to the point of being trivial, but the extension of the algorithm to ellipses is non-trivial. Surprisingly, the approximation ratio of the areas is not bounded and we provide an elegant proof of this.

1 Introduction

Of late geometric computation in the streaming model has received a lot of attention from researchers in the Computational Geometry community. See [4, 1, 5, 12, 6] for a sampling of results in this area. In this model, we assume there is limited workspace to store data steaming by. It is a particularly attractive model for solving computational problems requiring massive amounts of data, which has become rather commonplace in the context of powerful current technologies.

In [11] Zarrabi-Zadeh and Chan proposed a very simple algorithm for computing an approximate minimum spanning spanning ball of a set of n points $P = \{p_1, p_2, p_3, ..., p_n\}$ in the streaming model of computation and showed by an elegant analysis that the radius of the approximate ball is to within 3/2 of the exact one. The main assumptions in their algorithm are that we are allowed to store only the center and radius of the current ball.

Motivated by this result, in this paper we consider the extension of this result to computing approximate minimum spanning ellipses (mse, for short), under the same assumptions. The ball algorithm is simple to the point of being trivial, but the extension of the algorithm to ellipses is non-trivial. Surprisingly, the approximation ratio of the areas is not bounded and we provide an elegant proof of this.

The problem of computing a minimum spanning ellipse of a planar set of points has been studied quite a bit. Silverman and Titterington [8] proposed an $O(n^6)$ algorithm for this problem, which was improved to $O(n^4)$ by Post [7] by a prune-and-search technique. Much later, Dyer [3] proposed an $O(n)$ time deterministic algorithm for this (indeed in $O(n)$ time for any fixed dimension d). Welzl [10] proposed an $O(n)$ expected time algorithm.

In the next section we describe our algorithm for computing an approximate mse and in the following section we provide an analysis of this algorithm. We conclude in the last section.

*E-mail: asishm@cs.uwindsor.ca
†E-mail: egreene08@hotmail.com
‡E-mail: sarke1a@uwindsor.ca

2 Streaming problem

At the i-th iteration of the algorithm, we are given a point p_i, and the current approximate mse, E, spanning the set $P_i = \{p_1, p_2, \ldots, p_{i-1}\}$. The parameters that provide a complete description of E are knwon to us. Given the constraint that the set $P_i = \{p_1, p_2, \ldots, p_{i-1}\}$ is no longer available to us, we would like to compute an mse that spans E and p_i, which becomes the new approxinate mse.

This is a special case of the problem of computing a minimum spanning ellipsoid that spans a given set of ellipsoids of full volume. Therefore it can be set up as a convex programming optimization problem [2]. In this paper we take a different route, suited to this simpler case.

There is a very simple streaming algorithm for balls [11] that picks the first input point as the center, and makes the ball just big enough to contain the farthest subsequent point. This gives an approximate spanning ball whose radius is within 2 of the exact.

We can apply the same idea to ellipses by using the first two points as foci, and making the ellipse just big enough to cover all subsequent points. If the points of S lie very close to a line segment, then the area of the actual minimum ellipse will be very close to zero. The first two points in the stream could both be very close to one endpoint of the line segment, and so the approximate mse will have area close to that of the circle with the line segment as a radius. Thus the approximation ratio of the areas can be made arbitrarily large.

The Zarrabi-Chan algorithm [11] finds an approximate minimum spanning ball by computing the smallest ball that contains the new point and the previous one. The final approximate minimum spanning ball has a radius that is within 3/2 of the optimal one. In this paper we show how to extend this idea to ellipses. Though this extension is non-trivial, we show, and this is surprising, that the area of the approximate mse can become unbounded with respect to the exact ellipse just as the naive algorithm suggested in the previous paragraph.

3 Approximation algorithm for minimum spanning ellipses

We begin with a brief sketch of the main ingredients of our algorithm.

On the first input point, the approximate mse, E_A, is set to be this point. On the second input point, E_A is set to the line segment joining this to the first one. As long as subsequent points are on the supporting line, l, of this segment, E_A continues to be a line segment joining two extreme points on l. On the first input point that is not incident on l, E_A is computed as the non-degenerate ellipse defined by the new point and the endpoints of the current segment. So far, E_A is thus identical with the exact ellipse. For a subsequent input point, p, that does not lie inside E_A, we solve the problem of finding an ellipse of minimum area that spans E_A and p.

We do this by finding an elliptic transformation (nomenclature due to Post [7]), T, that transforms E_A to a unit circle; the same transformation is applied to p. Then, in the transformed plane, we solve the easier problem of finding an mse of the unit circle and $T(p)$. Finally, we apply an inverse transformation $T^{-1}(.)$ to the mse in the transformed plane to find the desired mse in the original plane. That the transformed ellipse in the original plane is of minimum area is due to the fact that the elliptic transformation T preserves relative areas. We will prove this.

The details are in the next two subsections.

3.1 Finding T

Our T must have the following two necessary properties:

P1 For any ellipse E, both $T(E)$ as well as its inverse image $T^{-1}(E)$ is an ellipse;

P2 T preserves relativity of ellipse areas, so that $area(E_1) \leq area(E_2) \Leftrightarrow area(T(E_1)) \leq area(T(E_2))$.

It's obvious that if T is a rotation or translation it satisfies the above properties. The following lemma shows that this is also true when T is a scaling in any direction (in particular the x-axis, say)

Lemma 1. *Let $p = (x, y)$ and $\alpha > 0$. Then $T(p) = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ is an x-scaling that satisfies $P1$ and $P2$.*

Proof: Let E be an ellipse with center at $p_0 = (x_0, y_0)$, given by the matrix equation

$$1 = [p - p_0]^T A [p - p_0]$$

where $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is a positive definite matrix ($det(A) > 0$).

Let $p' = T(p)$ for some p on E and $p'_0 = T(p_0)$. Since T^{-1} is well-defined, the matrix equation of the transformed ellipse is

$$\left[T^{-1}p' - T^{-1}p'_0\right]^T A \left[T^{-1}p' - T^{-1}p'_0\right] = 1$$

or, on a little simplification

$$[p' - p'_0]^T T^{-1}AT^{-1} [p' - p'_0] = 1,$$

where

$$\begin{aligned} T^{-1}AT^{-1} &= \begin{bmatrix} \frac{1}{\alpha} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{a}{\alpha} & b \\ \frac{b}{\alpha} & c \end{bmatrix} \\ &= \begin{bmatrix} \frac{a}{\alpha^2} & \frac{b}{\alpha} \\ \frac{b}{\alpha} & c \end{bmatrix} \end{aligned}$$

Hence $T(E)$ is an ellipse.

Now, it is well-known that the area of E is given by $\frac{\pi}{\sqrt{\det(A)}} = \frac{\pi}{\sqrt{ac-b^2}}$. Therefore the area of $T(E) = \frac{\pi * \alpha}{\sqrt{\det(A)}}$. From this and the fact that $T^{-1}(E)$ is an ellipse because T^{-1} is an x-scaling, it follows that T satisfies $P1$ and $P2$. □

Let $2r_1$ and $2r_2$ be the lengths of the principal axes of an ellipse, centered at (x_0, y_0), and ϕ the angle that the principal axis of length $2r_1$ makes with the x-axis (see Fig. 1).

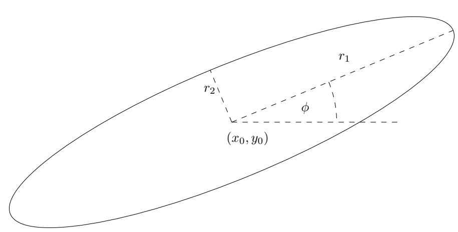

Figure 1: An ellipse

Then,

Lemma 2. *The transformation $T = S^y(1/r_2)S^x(1/r_1)R(-\phi)[(x, y) - (x_0, y_0)]^t$, where S^y is an y-scaling by $1/r_2$, S^x is an x-scaling by $1/r_1$, $R(-\phi)$ is a rotation by $-\phi$, and $(x, y) - (x_0, y_0)$ is a translation, takes the ellipse E into a unit circle.*

Proof: Since rotations and translations satisfy P_1 and P_2, and by Lemma 1, both S^x and S^y does as well, T satisfies them too. Also, from the choice of the scalings it is clear that $T(E)$ is a unit circle. $\square$

We can further simplify the problem by applying a rotation R to T that carries $p' = T(p)$ onto the positive x-axis in the transformed plane so that $T^r(p') = (d, 0)$. This understood, we will use p' in place of $R(p')$ hereafter.

So the problem of finding the smallest approximate mse E_A, containing the current ellipse E and a point p outside it, is reduced to that of finding the smallest ellipse E'_A containing the unit circle and the point $(d, 0)$ on the positive x-axis.

We discuss this next.

3.2 Finding E'_A

Lemma 3. *Let E and $\overline{E}$ be two distinct ellipses that are the reflections of each other in the x-axis. Then the area of the (strict) convex combination $\lambda E + (1-\lambda)\overline{E}$ (where $0 < \lambda < 1$) is smaller than the areas of E and $\overline{E}$.*

Proof: Let $F(x, y) = a(x-x_0)^2+2b(x-x_0)(y-y_0)+c(y-y_0)^2-1 = 0$ be the qudratic form of E. Negating y and y_0 in $F(x, y)$, gives the quadratic form of $\overline{E}$: $\overline{F}(x, y) = a(x-x_0)^2-2b(x-x_0)(y-y_0)+c(y-y_0)^2-1 = 0$. Thus $\lambda F(x, y)+(1-\lambda)\overline{F}(x, y) = a(x-x_0)^2+2b(2\lambda-1)(x-x_0)(y-y_0)+c(y-y_0)^2-1 = 0$ is the quadratic form of the ellipse corresponding to the convex combination $\lambda E + (1-\lambda)\overline{E}$. Now $0 < \lambda < 1$, so $-1 < 2\lambda - 1 < 1$, and hence $ac - b^2(2\lambda-1)^2 > ac - b^2 = ac - (-b)^2$. Thus the area of the convex combination is smaller than that of E (and $\overline{E}$). $\square$

Lemma 4. *The smallest ellipse E'_A containing p' and the unit circle will have a principal axis oriented along the x-axis.*

Proof: Suppose that the principal axes of E'_A are not parallel to the x- and y-axes and $0 < \lambda < 1$. Since p' is on the x-axis and the unit circle is symmetric with respect to it, there must be another ellipse $\overline{E'_A}$ with the same area, that is the reflection of E'_A in the x-axis. The convex combination $\lambda E'_A + (1-\lambda)\overline{E'_A}$ also spans the unit circle and p'. By Lemma 3, the area of $\lambda E'_A + (1-\lambda)\overline{E'_A}$ is strictly smaller than E'_A (or $\overline{E'_A}$). Therefore, E'_A cannot be the mse. $\square$

Let C be the smallest circle that contains p' and the unit circle (Fig. 2).

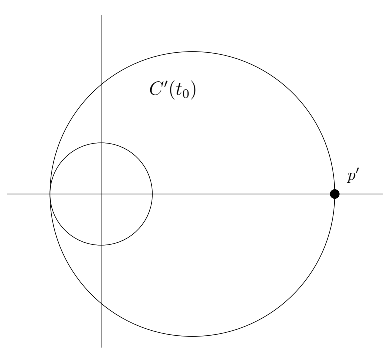

Figure 2: A circle containing and tangent to the unit circle, and passing through p'

Let us stretch (shrink) C along the x- and y-axes into an ellipse C', always maintaining the following three invariants:

1. C' contains the unit circle;

2. C' is tangent to the unit circle at at least one point;

3. C' passes through p'.

See Figure 3. E'_A will have these properties as well. being the C' with smallest area.

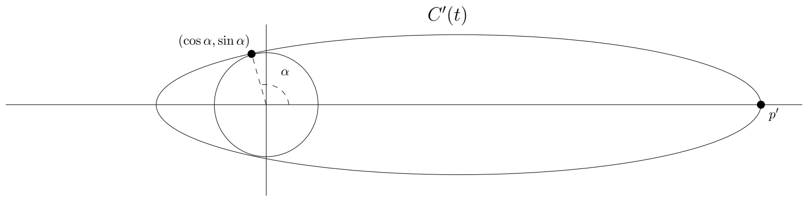

Figure 3: An ellipse containing and tangent to the unit circle, and passing through p'

Now we can shrink C along the y-axis, and keep the point of tangency at $(-1, 0)$, up to a certain point until we start violating the containment constraint. At this point, we can start stretching the ellipse along the x-axis.

Let us use a continuous parameter t as an argument to an ellipse, obtained by the stretching/shrinking process, that satisfies the above invariants. Set $C' = C(t_0)$.

Assume that t decreases as we stretch along the y-axis, and increases as we first shrink along the y-axis and then stretch along the x-axis.

If we start t at t_0, and decrease t, the minor axis of $C'(t)$ remains constant, but the major axis increases in length. So for $t < t_0 \Rightarrow area(C'(t)) > area(C'(t_0))$.

Consider another value t_2, such that the center of $C'(t_2)$ is the origin. See Fig. 4. Then $C'(t_2)$ is tangent to the unit circle at $(0, 1)$ and $(0, -1)$. If we increase t beyond this value, the major and minor axes of $C'_{i+1}(t)$ will increase in length. So for $t > t_2 \Rightarrow area(C'(t)) > area(C'(t_2))$.

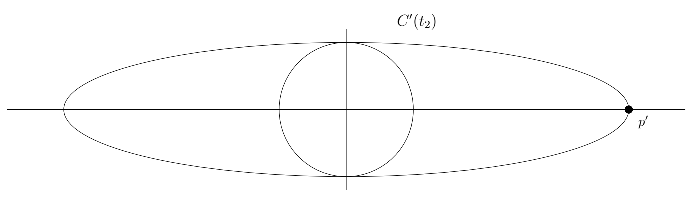

Figure 4: $C'(t)$ corresponding to $t = t_2$

Thus there $\exists t' \in [t_0, t_2]$ such that $E'_A = C(t')$. This suggests that we should be able to write the area of $C'(t)$ in terms of just t, and then minimize that expression to find E'_A.

There is yet another restriction we can put on t. As we increase t beyond t_0, there is some value t_1 such that we can no longer shrink $C'(t)$ along the y-axis while driving it through p' and $(-1, 0)$. See Fig. 5. Thus for $t \in [t_0, t_1)$, $area(C'(t)) > area(C'(t_1))$. Thus for the minimum area ellipse, E'_A we can focus on the range $[t_1, t_2]$.

The curvature of $C'(t)$ at $(r_1 \cos\theta, r_2 \sin\theta)$ is $\frac{r_1 r_2}{({r_2}^2 \cos^2\theta + {r_1}^2 \sin^2\theta)^{\frac{3}{2}}}$ [9]. So the curvature at $(-1, 0)$ is $\frac{r_1}{{r_2}^2}$. If $t \in [t_0, t_1]$, then the major axis has length $d + 1$. Setting this curvature to 1 to match that of the unit circle (which happens when $t = t_1$), we get $r_2 = \sqrt{\frac{d+1}{2}}$. The area of $C'(t_1)$ is then $\pi \left(\frac{d+1}{2}\right)^{\frac{3}{2}}$.

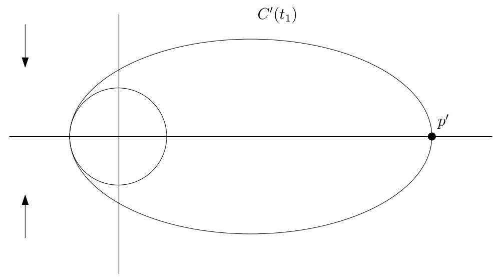

Figure 5: $C'(t)$ corresponding to $t = t_1$

The equation of $C'(t)$ is $a(x - x_0)^2 + cy^2 = 1$. Since p' is on it,

$$a(d - x_0)^2 = 1 \tag{1}$$

Let α be the angle made by the x-axis, the origin, and the point at which $C'(t)$ touches the upper half of the unit circle. (See Figure 3.) The tangent to the circle at $(\cos\alpha, \sin\alpha)$ is

$$x\cos\alpha + y\sin\alpha = 1 \tag{2}$$

The tangent to the ellipse at $(\cos\alpha, \sin\alpha)$ is

$$\begin{aligned} a(x - x_0)(\cos\alpha - x_0) + cy\sin\alpha &= 1 \\ xa(\cos\alpha - x_0) + yc\sin\alpha &= 1 + ax_0(\cos\alpha - x_0) \end{aligned} \tag{3}$$

If $\alpha = \frac{\pi}{2}$, then the tangent is

$$y = 1$$

from the circle, and

$$-ax_0x + cy = 1 - ax_0^2$$

from the ellipse. Then it must be that $x_0 = 0$ and $c = 1$. This verifies that the area of $C'(t_2)$ is πd. So we are looking for a minimum area at most $min(\pi d, \pi\left(\frac{d+1}{2}\right)^{\frac{3}{2}})$.

Assume from now on that $\alpha \in (\frac{\pi}{2}, \pi)$. Since (2) and (3) are the same line, we know $\frac{\cos\alpha}{a(\cos\alpha - x_0)} = \frac{1}{c} = \frac{1}{1+ax_0(\cos\alpha - x_0)}$. So

$$\begin{aligned} a(\cos\alpha - x_0) &= \cos\alpha + ax_0\cos\alpha(\cos\alpha - x_0) \\ a(\cos\alpha - x_0 - x_0\cos\alpha(\cos\alpha - x_0)) &= \cos\alpha \\ a(\cos\alpha - x_0)(1 - x_0\cos\alpha) &= \cos\alpha \\ a &= \frac{\cos\alpha}{(\cos\alpha - x_0)(1 - x_0\cos\alpha)} \end{aligned}$$

and

$$\begin{aligned} c &= \frac{\cos\alpha - x_0}{\cos\alpha}\frac{\cos\alpha}{(\cos\alpha - x_0)(1 - x_0\cos\alpha)} \\ &= \frac{1}{1 - x_0\cos\alpha} \end{aligned}$$

Using (1), we know that

$$\begin{aligned} (d - x_0)^2 &= (\cos\alpha - x_0)(\sec\alpha - x_0) \\ d^2 - 2dx_0 &= 1 - x_0(\cos\alpha + \sec\alpha) \\ x_0(\cos\alpha + \sec\alpha - 2d) &= 1 - d^2 \\ x_0 &= \frac{d^2 - 1}{2d - \cos\alpha - \sec\alpha} \end{aligned}$$

We can therefore write the area of $C'(t)$ in terms of just $\beta = \cos\alpha$:

$$\begin{aligned}
area(C'(t)) &= \frac{\pi}{\sqrt{ac}} \\
&= \pi\sqrt{f(\beta)} \\
&= \pi\sqrt{\frac{(\beta - x_0)(1 - x_0\beta)^2}{\beta}} \\
&= \pi\sqrt{\frac{\left(\beta - \frac{d^2-1}{2d-\beta-\beta^{-1}}\right)\left(1 - \frac{d^2-1}{2d-\beta-\beta^{-1}}\beta\right)^2}{\beta}} \\
&= \pi\sqrt{\left(1 + \frac{d^2-1}{\beta^2 - 2d\beta + 1}\right)\left(1 + (d^2-1)\frac{\beta^2}{\beta^2 - 2d\beta + 1}\right)^2} \\
&= \pi\sqrt{\frac{(\beta - d)^2}{\beta^2 - 2d\beta + 1}\left(\frac{(d\beta - 1)^2}{\beta^2 - 2d\beta + 1}\right)^2} \\
&= \pi\sqrt{\frac{(\beta - d)^2(d\beta - 1)^4}{(\beta^2 - 2d\beta + 1)^3}}
\end{aligned}$$

To find the $\beta \in (-1, 0)$ that minimizes $f(\beta)$,

$$\begin{aligned}
f'(\beta) &= \frac{(2(\beta-d)(d\beta-1)^4 + (\beta-d)^2 4(d\beta-1)^3 d)(\beta^2-2d\beta+1)^3 - (\beta-d)^2(d\beta-1)^4 3(\beta^2-2d\beta+1)^2(2\beta-2d)}{(\beta^2-2d\beta+1)^6} \\
&= 2(\beta-d)(d\beta-1)^3\frac{((d\beta-1)+2d(\beta-d))(\beta^2-2d\beta+1) - 3(\beta-d)^2(d\beta-1)}{(\beta^2-2d\beta+1)^4}
\end{aligned}$$

Setting $f'(\beta) = 0$,

$$\begin{aligned}
0 &= ((d\beta-1)+2d(\beta-d))(\beta^2-2d\beta+1) - 3(\beta-d)^2(d\beta-1) \\
&= (3d\beta - 2d^2 - 1)(\beta^2 - 2d\beta + 1) - 3(\beta^2 - 2d\beta + d^2)(d\beta - 1) \\
&= 3d\beta^3 - 6d^2\beta^2 + 3d\beta - (2d^2+1)\beta^2 + 2(2d^3+d)\beta - 2d^2 - 1 - 3d\beta^3 + 6d^2\beta^2 - 3d^3\beta + 3\beta^2 - 6d\beta + 3d^2 \\
&= 2(1-d^2)\beta^2 + d(d^2-1)\beta + (d^2-1)
\end{aligned}$$

Now $d > 1$, since p' is outside the unit circle, so $1 - d^2 \neq 0$.

$$\begin{aligned}
&= 2\beta^2 - d\beta - 1 \\
\beta &= \frac{d \pm \sqrt{d^2+8}}{4}
\end{aligned}$$

We need $\beta < 0$, so

$$\beta = \frac{d - \sqrt{d^2+8}}{4} \tag{4}$$

If the area of the ellipse corresponding to this value of β is smaller than $min(\pi d, \pi\left(\frac{d+1}{2}\right)^{\frac{3}{2}})$, then E'_A is described by $a(x - x_0)^2 + cy^2 = 1$, where

$$x_0 = \frac{d^2-1}{2d - \beta - \beta^{-1}} \tag{5}$$

$$a = \frac{1}{(\beta - x_0)(\beta^{-1} - x_0)} \tag{6}$$

$$c = \frac{1}{1 - x_0\beta} \tag{7}$$

4 Approximation ratio of area

We first construct an example input sequence to show that the approximation ratio of the area of the approximate mse to the exact ellipse can be unbounded.

To construct the example we need the following important lemma.

Lemma 5. *The exact ellipse incident on $(0,k)$, $(0,-k)$ and $(3l,0)$ must also pass through $(-l,0)$. Moreover, the length of the minor axis of this exact ellipse is $\frac{4k}{\sqrt{3}}$.*

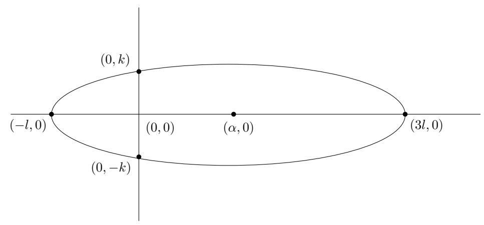

Figure 6: Exact ellipse through $(0,k)$, $(0,-k)$ and $(3l,0)$

Proof: Let $\dfrac{(x-\alpha)^2}{a^2}+\dfrac{y^2}{b^2}=1$ be the equation of the exact ellipse. Since it passes through $(3l,0)$, we have $a=3l-\alpha$. By substituting $(0,k)$ we get $b=\dfrac{ak}{\sqrt{a^2-\alpha^2}}$. The area of this ellipse $A=\pi ab=\pi\frac{(3l-\alpha)^2k}{\sqrt{(3l-\alpha)^2-\alpha^2}}$. By setting $\dfrac{dA}{d\alpha}=0$, we get $\alpha=l,\dfrac{3l}{2},3l$. But $\alpha=l$ gives the minimum area. The center of the exact ellipse is $(l,0)$, and $a=3l-\alpha=2l$. Therefore it passes through $(-l,0)$. Since $\alpha=l$ and $a=2l$, we get $b=\dfrac{ak}{\sqrt{a^2-\alpha^2}}=\dfrac{2lk}{\sqrt{4l^2-l^2}}=\dfrac{2k}{\sqrt{3}}$. Thus the length of the minor axis is $\dfrac{4k}{\sqrt{3}}$. □

In the following lemma we construct an example for which the ratio of the length of the minor axis of approximate mse to the length of the minor axis of exact ellipse can be made unbounded. We make crucial use of Lemma 5.

Lemma 6. *There exists an input point sequence for which the ratio of the minor axis of approximate mse to that of the exact ellipse becomes unbounded.*

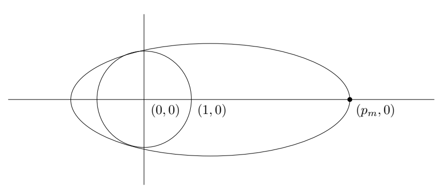

Figure 7: m-th transformed plane P_m

Proof: Consider three points $(0,-1)$, $(0,1)$ and $(-1,0)$. Let E_0 be the exact ellipse through these three points. Since the approximate mse through these 3 points is identical with the exact one, the ratio of their areas is 1 in this case. Now we start adding points on the positive x-axis.

If we add a point $(p_m,0)$ on the positive x-axis then the exact ellipse will pass through $(0,-1)$, $(0,1)$, and $(p_m,0)$. For any p_m, the length of the minor axis of the exact ellipse is $\frac{4}{\sqrt{3}}$ (Lemma 5).

We denote the original plane by P_0. The transformation S_0T_0 will transform E_0 to the unit circle where T_0 is a translation and S_0 is a scaling. Let $P_1=S_0T_0P_0$ be the transformed plane. We can choose a point $(p',0)$ on this transformed plane such that the minor axis of the spanning ellipse, which passes through $(p',0)$

and encloses the unit circle, is greater than 1. Suppose ST is the transformation which transforms this new ellipse to the unit circle where S is a scaling and T is a translation. Let $P_2 = STP_1$. Take $(p', 0)$ on P_2 and find the ellipse that passes through $(p', 0)$ and encloses the unit circle.

If we repeat this process n times, then each time we increase the minor axis of the ellipse same amount. Denote the approximate mse on P_n by E_n. The approximate mse on the original plane can be found by applying the transforamtion $(T^{-1}S^{-1})^n T_0^{-1} S_0^{-1} E_n$. Since S is a shrinking along y-axis, S^{-1} is an expansion along y-axis. Therefore we can expand the minor axis of approximate mse as much as we want by adding a sufficient number of points. Since the minor axis of the exact ellipse is fixed, we conclude that the ratio of the minor axis of approximate mse to that of the exact ellipse is unbounded. □

Note that the point $(p', 0)$ on the m-th transformed plane corresponds to $(T^{-1}S^{-1})^m T_0^{-1} S_0^{-1}(p', 0)$ on the original plane. □

Lemma 7. *The ratio of the length of major axis of approximate mse to the length of major axis of exact ellipse, constructed by the method described in Lemma 6, is greater than* $3/4$.

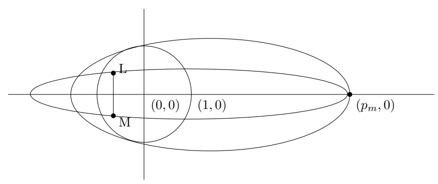

Figure 8: LM is the transformed line segment joining $(0, -1)$ and $(0, 1)$ on the original plane

Proof: Suppose $2a_A$ and $2a_E$ denote the length of the major axis of the approximate mse and the exact ellipse respectively. Figure 8 shows the m-th transformed plane P_m, and the line segment LM is the transformed line segment joining $(0, 1)$ and $(0, -1)$ on the original plane. By Lemma 5 the length of the major axis is $4/3$ of the distance of $(p_m, 0)$ from LM. Since LM must be enclosed by the unit circle, the maximum distance of point $(p_m, 0)$ from LM is $p_m + 1$. Hence $2a_E \leq \frac{4}{3}(p_m + 1)$. From Figure 8 it is clear that $2a_A \geq p_m + 1$. This implies that $\dfrac{2a_A}{2a_E} \geq \dfrac{p_m + 1}{(4/3)(p_m + 1)} \geq \dfrac{3}{4}$. □

Theorem 1. *There exists an input point sequence of points for which the approximation ratio of the ellipse areas becomes unbounded.*

Proof: Let a_A, b_A, and Area_A denote the semi-major axis, semi-minor axis and the area of the approximate mse respectively in the transformed plane, while a_E, b_E, and Area_E denote the same quantities respectively for the exact ellipse. By the method described in Lemma 6 we can construct an input point sequence for which $\dfrac{b_A}{b_E}$ is unbounded. By Lemma 7 we have $\dfrac{a_A}{a_E} \geq \dfrac{3}{4}$. Since relative areas are preserved by an elliptic transformation, the ratio $\dfrac{\text{Area}_A}{\text{Area}_E}$ is identical for the ellipses in the original plane. Now $\dfrac{\text{Area}_A}{\text{Area}_E} = \dfrac{\pi a_A b_A}{\pi a_E b_E} = \dfrac{a_A}{a_E}\dfrac{b_A}{b_E} \geq \dfrac{3}{4}\dfrac{b_A}{b_E}$. Since by Lemma 2, there exists an input sequence for which the ratio $\dfrac{b_A}{b_E}$ is unbounded, we conclude that $\frac{\text{Area}_A}{\text{Area}_E}$ is unbounded for this input sequence. □

Consider a sequence of ellipses described in Lemma 6. Suppose for E_0, the lengths of the axis parallel to x-axis and y-axis are r_x and r_y respectively. The scaling factors to transform E_0 to a unit circle are $1/r_x$ and $1/r_y$ along x-axis and along y-axis respectively. Suppose on the transformed planes $P_1, P_2, P_3, \cdots$, the minimum ellipses $E_1, E_2, E_3, \cdots$ that are enclosing the unit circle and the point $(p_m, 0)$ has the minor axis $1 + s$ and the major axis $p_m + 1 + t$. On the original plane the minor axis and the major axis of E_n are $r_x(p_m + 1 + t)^n$ and $r_y(1 + s)^n$ respectively. In the following theorem we will show that the approximation ratio does not depend on the eccentricity.

Theorem 2. *The approximation ratio does not depend on the eccentricity of the exact ellipse.*

Proof: Suppose the eccentricity of an exact ellipse e and a positive real number M are given. We will construct an example such that the approximation ratio is greater than or equal to M.

Choose an n such that $\frac{(1+s)^{n-1}}{p_m+1+t} > 2M$ and choose an a such that $(p_m+1+t)^{n-1} < a < (p_m+1+t)^n$. Then $b = a\sqrt{1-e^2}$. Let $k = \frac{\sqrt{3}b}{2}$.

Start with three points $(0,-k)$, $(0,k)$ and $(-1,0)$ and construct a sequence of ellipses described in Lemma 6. The lengths of the major axis and the minor axis of E_{n-1} on the original plane are $a_A = r'_x(p_m+1+t)^{n-1}$, $b_A = r'_y(1+s)^{n-1}$ respectively. The length of the major axis and the minor axis of the exact ellipse are $a_E = a$ and $b_E = \frac{2k}{\sqrt{3}}$. Therefore,

$$
\begin{aligned}
\frac{\text{Area}_A}{\text{Area}_E} &= \frac{\pi a_A b_A}{\pi a_E b_E} \\
&= \frac{a_A}{a_E}\frac{b_A}{b_E} \\
&= \frac{r'_x(p_m+1+t)^{n-1}}{a}\frac{r'_y(1+s)^{n-1}}{\frac{2k}{\sqrt{3}}} \\
&= \frac{r'_x(p_m+1+t)^{n}}{a}\frac{\sqrt{3}r'_y(1+s)^{n-1}}{2k(p_m+1+t)} \\
&= r'_x\frac{(p_m+1+t)^{n}}{a}\cdot\frac{\sqrt{3}}{2}\cdot\frac{r'_y}{k}\cdot\frac{(1+s)^{n-1}}{(p_m+1+t)} \\
&\geq \frac{1}{2}\cdot 1\cdot 1\cdot 1\cdot 2M \\
&= M
\end{aligned}
$$

case. □

In the next theorem we will show that for any given exact ellipse we can find a sequence of points for which the approximation ratio could be as large as we want.

Theorem 3. *For any given exact ellipse there is a sequence of points for which the approximation ratio is arbitrarily large.*

Proof: Suppose an exact ellipse is given on a plane P_{-2}. Let $M \in \mathbb{R}$. Using appropriate rotation and translation align the major axis and the minor axis of this exact ellipse along x-axis and y-axis respectively. Let denote this plane by P_{-1} and let e be the eccentricity of the exact ellipse.

Now determine a and b which are described in the proof of Theorem 2. Using the appropriate scaling factor transform the exact ellipse to the ellipse with the major axis and the minor axis being a and b respectively. Denote this plane by P_0. Determine a sequence of points described in the proof of Theorem 2. By transforming this sequence of points to P_{-2} plane we get the desired sequence. □

5 Directional boundedness

Fortunately, though the ratio of areas of the arrroximate to the exact ellipse can get unbounded, there exists a direction in which the widths of the ellipses are bounded. Below is a proof of this claim.

The idea is fairly straight forward. Given some minimum spanning ellipse (exact or approximated) of a point set, we need to be able to find 2 points in that point set whose distance between themselves in some direction is at least a constant fraction of the width of the ellipse in that direction. The following proof shows that we can always find 2 such points and a direction.

The proof hinges on the fact that the convex hull of the point set will always contain the center of the minimum ellipse, approximated or not. So, to start with, we prove an equivalent version of this; that any halfspace that contains the center of a minimum spanning ellipse of a point set, also contains a point from that set. We start with the case of the exact minimum spanning ellipse and then use induction to extend this to the approximated case.

Lemma 8. *If E is some exact minimum spanning ellipse that contains some point set Q, then any closed halfspace that contains the center, c, of E must also contain a point $p \in Q$.*

Proof. Assume that there exists some closed halfspace, H, such that $c \in H$, but $H \cap Q = \emptyset$. Without loss of generality, we can assume c lies on the boundary of H. Let $\epsilon > 0$ be equal to the distance from H to Q in the direction of the diameter through c conjugate to the bounding line of H . We can translate E by ϵ in the direction of the vector parallel to this conjugate diameter , while still ensuring $Q \subset E$. After the translation, Q no longer has any points on the boundary of E, so E cannot be the minimum ellipse that contains Q, a contradiction. □

Lemma 9. *Let $Q = \{q_1, q_2, ..., q_k\}$ be the set of points used to find E_0, the initial minimum spanning ellipse. Let E_n be the approximated ellipse after the $(n+k)$-th point, p_n has been seen. Any halfspace that contains the center c_n of the ellipse E_n also contains at least one point $p \in P = Q \cup \{p_1, p_2, ..., p_n\}$.*

Proof. From Lemma 8 we know that the base case ($n = 0$) is true. Let us assume that the $(n-1)$-th case was true. If a new point $p_n \notin E_{n-1}$ is seen, then a new minimum ellipse E_n is found that contains both E_{n-1} and p_n. If c_{n-1} is the center of E_{n-1} and $\overline{p_n c_{n-1}}$ is the line segment whose endpoints are p_n and c_{n-1}, then, from the construction of E_n we see that $c_n \in \overline{p_n c_{n-1}}$, where c_n is the center of E_n. Any halfspace H that has a non-empty intersection with $\overline{p_n c_{n-1}}$ (ie. $H \cap \overline{p_n c_{n-1}} \neq \emptyset$), must contain either p_n, c_{n-1} or both. If H contains p_n then clearly it contains a point from P, since $p_n \in P$. Further more, if H contains the c_{n-1}, center of E_{n-1}, then it must contain some point from $P - \{p_n\} \subseteq P$. Thus, any halfspace that contains $c_n \in \overline{p_n c_{n-1}}$ must also contain a point $p \in P$. □

With these 2 Lemmas, we can now give the main proof.

Definition 1. *Let $\Psi_{\vec{v}}(P)$ be a real-valued function that returns the width of a point set P in direction $\vec{v}$.*

Theorem 4. *If E is a minimum area ellipse (exact or approximated) that contains the point set P, then there must exist 2 points $\{p_i, p_j\} \subseteq P$ and a direction $\vec{n}$, such that $\Psi_{\vec{n}}(E) \leq 2\Psi_{\vec{n}}(\{p_i, p_j\})$.*

Proof. Since E is the minimum ellipse (exact or approximated) containing P, we can find a point $p_i \in P$ that lies on the boundary of E. Let c be the center of E and $\vec{n}$ be the vector normal to E at the point p_i. Let H be the halfspace whose normal is $\vec{n}$, does not contain p_i and whose boundary contains c. From Lemma 9, we can find another point $p_j \in P \cap H$. Clearly, $\Psi_{\vec{n}}(\{p_i, p_j\}) \geq \Psi_{\vec{n}}(\{p_i, c\})$. Since $2\Psi_{\vec{n}}(\{p_i, c\}) = \Psi_{\vec{n}}(E) \leq 2\Psi_{\vec{n}}(\{q_i, q_j\})$, the points p_i and p_j fufill the requirements; that is, in direction $\vec{n}$, the ellipse E has a bounded width by a constant factor 2. □

Theorem 5. *There exists a direction at which the approximation ratio is bounded.*

Proof. Suppose E_E is the exact minimum ellipse and E_A is the approximate ellipse containing P. By theorem 4, we get $\Psi_{\vec{n}}(E_A) \leq 2\Psi_{\vec{n}}(\{p_i, p_j\})$. Clearly we can see that $\Psi_{\vec{n}}(E_E) \geq \Psi_{\vec{n}}(\{p_i, p_j\})$. Therefore $\dfrac{\Psi_{\vec{n}}(E_A)}{\Psi_{\vec{n}}(E_E)} \leq 2$. □

6 Conclusions

The unboundedness of the ratio of areas came to us as a bit of surprise, since we believed otherwise, given Zarrabi-Chan's result on balls [11]. So the question is if this diminishes the practical utility of this algorithm for finding approximate mses. To find out we implemented the algorithm for the 2d-case. By clicking on the link **software** this implementation can be viewed at `http://cs.uwindsor.ca/~asishm`. Experiments with this implementation for randomly input point sets seem to suggest that the average behaviour of this algorithm may still be very good. We also experimented with points drwan from a uniform distribution; in these cases the approximation is very very good. It would be interesting to prop this up with some theoretical analysis.

It might be possible to improve the factor 2 for directional boundednss.

It is also possible to extend the results of this paper to d-dimensions. We have some preliminary results in this direction that we plan to report later and elsewhere.

References

[1] Bagchi, Chaudhary, Eppstein, and Goodrich. Deterministic sampling and range counting in geometric data streams. In *ACM Transactions on Algorithms (TALG)*, volume 3. 2007.

[2] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, volume 2 of *MPS/SIAM Series on Optimization*. SIAM, Philadelphia, 2001.

[3] Martin E. Dyer. A class of convex programs with applications to computational geometry. In *Symposium on Computational Geometry*, pages 9–15, 1992.

[4] Hershberger and Suri. Adaptive sampling for geometric problems over data streams. In *PODS: 23th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, 2004.

[5] Isenburg, Liu, Shewchuk, and Snoeyink. Illustrating the streaming construction of 2D delaunay triangulations (short). In *COMPGEOM: Annual ACM Symposium on Computational Geometry*, 2006.

[6] S. Muthukrishnan. Data streams: algorithms and applications. *Found. Trends Theor. Comput. Sci.*, 1(2):117–236, 2005.

[7] Mark J. Post. A minimum spanning ellipse algorithm. In *FOCS*, pages 115–122. IEEE, 1981.

[8] B. W. Silverman and D. M. Titterington. Minimum covering ellipses. *SIAM Journal on Scientific and Statistical Computing*, 1(4):401–409, December 1980.

[9] Eric W. Weisstein. Ellipse. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Ellipse.html.

[10] E. Welzl. Smallest enclosing disks (balls and ellipsoids). *Lecture Notes in Computer Science*, 555:359–370, 1991.

[11] Hamid Zarrabi-Zadeh and Timothy Chan. A simple streaming algorithm for minimum enclosing balls. In *Proceedings of the 18th Canadian Conference on Computational Geometry (CCCG'06)*, pages 139–142, 2006.

[12] Hamid Zarrabi-Zadeh and Asish Mukhopadhyay. Streaming 1-center with outliers in high dimensions. In *Proceedings of the 21st Canadian Conference on Computational Geometry (CCCG'09)*, pages 000–000, 2009.