A streaming algorithm for computing an approximate minimum spanning ellipse

Asish Mukhopadhyay* and Eugene Greene†
School of Computer Science
University of Windsor
Windsor, Ontario
Canada

Abstract

We propose a streaming algorithm to compute an approximate ellipse of a set of points. This extends the work of Zarrabi-Zadeh and Chan who proposed a very simple algorithm to compute a 3/2-approximate (in terms of radius) d-dimensional ball in the same model. We have been able to obtain only a very crude approximation ratio of $\frac{9k}{4\sqrt{1-e^2}}$ (in terms of area), where e is the eccentricity of the minimum spanning ellipse of the point set, and k is a constant defined in the text. The algorithm uses $O(1)$ space, and $O(1)$ time for each point.

1 Introduction

Of late geometric computation in the streaming model has received a lot of attention from researchers in the CG community [4, 1, 5].

In [10] Zarrabi-Zadeh and Chan proposed a simple algorithm for computing an approximate spanning ball of a set of n points $S = \{s_1, s_2, s_3, ..., s_n\}$ in the streaming model of computation. Though the algorithm is simple, they came up with a very clever and elegant analysis to show that the approximation ratio is 3/2 (in terms of radius). Inpired by the above paper, in this note we investigate the problem of extending their algorithm to compute an approximate *ellipse* in the same model.

The problem of computing a spanning ellipse of a planar set of points has been around for a very long time. Silverman and Titterington [7] proposed an $O(n^6)$ algorithm for this problem, which was improved to $O(n^4)$ by Post [6]. Much later, Dyer [3] proposed an $O(n)$ time deterministic algorithm for this (indeed in $O(n)$ time for any fixed dimension d). Welzl [9] proposed an $O(n)$ expected time algorithm.

In this note we show how to find an approximate minimum ellipse spanning S, with an approximation factor of $\frac{9k}{4\sqrt{1-e^2}}$ (in terms of area), where e is the eccentricity of the minimum spanning ellipse of S, and k is a constant defined later in the text. The algorithm uses $O(1)$ space and $O(n)$ time.

2 Streaming problem

At iteration $i + 1$ of the algorithm, we are given a point s_{i+1}, and also some previous ellipse E_i. We know that E_i is an approximate minimum enclosing ellipse for some unknown set S_i. We then want a new ellipse E_{i+1} that is an approximate minimum enclosing ellipse for $S_{i+1} = S_i \bigcup \{s_{i+1}\}$.

*E-mail: asishm@cs.uwindsor.ca
†E-mail: greene6@uwindsor.ca

This is a special case of the problem of computing a minimum ellipsoid that spans a given set of ellipsoids of full volume. Therefore it can be set up as a convex programming optimization problem [2].

There is a very simple streaming algorithm for balls [10] that picks the first input point as the center, and makes the ball just big enough to contain the farthest subsequent point. This has an approximation ratio of 2 (in terms of radius). 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 ellipse will have area close to that of the circle with the line segment as a radius. This note will suggest a better algorithm.

3 Minimum spanning balls

Zarrabi-Zadeh and Chan [10] give an algorithm for solving this problem for minimum enclosing balls. Each new ball in their algorithm is the smallest that contains the new point and also the previous ball. The algorithm produces a ball that has a radius no larger than 3/2 of the radius of the optimal ball.

We can apply the same idea to ellipses. Assume that E_i encloses all of S_i. At the $(i+1)$st iteration, we approximate the minimum ellipse by the smallest ellipse E_{i+1} enclosing both E_i, and s_{i+1}. Section 5 will describe how to accomplish this.

4 Ellipses

An ellipse E with center $p_0 = (x_0, y_0)$ is the set of points $p = (x, y)$ that satisfy

$$
\begin{aligned}
[p - p_0]^T A\, [p - p_0] &= 1 \\
\begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} &= 1
\end{aligned}
$$

This corresponds to an ellipse if and only if $\det(A) > 0$. Expanding this,

$$
\begin{aligned}
a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2 &= 1 \\
F(p) = a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2 - 1 &= 0
\end{aligned}
$$

The area of E will be $\frac{\pi}{\sqrt{\det(A)}} = \frac{\pi}{\sqrt{ac-b^2}}$. If E is aligned with the x and y axes, then $b = 0$.

The following symbols will be used to denote the geometry of an ellipse E (see Figure 1):

- E will have center $p_0 = (x_0, y_0)$;
- r_1 is half the length of the axis corresponding to ϕ, and r_2 is half the length of the other axis (r_1 can correspond to the minor or major axis);
- ϕ is the angle corresponding to the direction of r_1.

5 Our solution

When the algorithm receives its first point, the approximate minimum ellipse is just that point. We will only consider subsequent points that are not already in the current approximate ellipse. For the second point, the approximation becomes a segment. As long as subsequent points are on the supporting line of this segment, the approximate ellipse will be a segment. The next point not lying on the supporting line produces a non-degenerate ellipse defined by the new point and the endpoints of the segment. The approximate ellipse

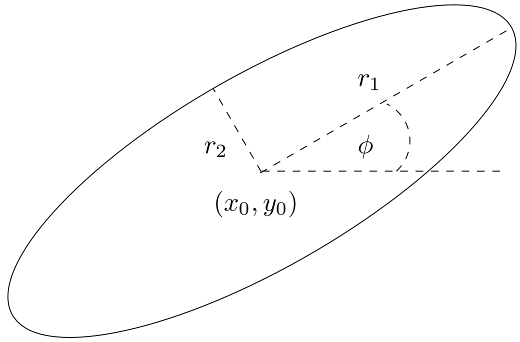

Figure 1: An ellipse

is thus equal to the exact ellipse up to this moment.

Given non-degenerate ellipse E_i and point s_{i+1}, we can derive a transformation T_i that maps E_i to the unit circle. Now define two necessary properties:

P1 For any ellipse E, the image $T_i(E)$ is an ellipse, and the inverse image $T_i^{-1}(E)$ is an ellipse;

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

Assuming T_i has these properties, we transform s_{i+1} into $s'_{i+1} = T_i(s_{i+1})$, and let E'_{i+1} be the smallest ellipse that contains the unit circle and s'_{i+1}. Then $E_{i+1} = T_i^{-1}(E'_{i+1})$ is the smallest ellipse that contains E_i and s_{i+1}. How do we find T_i?

5.1 Defining the transformation

Lemma 1. *Let* $T(x, y) = S \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. *Then* T *satisfies* $P1$ *and* $P2$.

Proof:
Let E be some ellipse

$$\begin{aligned} 1 &= [p - p_0]^T A [p - p_0] \\ &= \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \end{aligned}$$

For some p on E, let

$$\begin{aligned} p' &= T(p) \\ &= \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \alpha x \\ y \end{bmatrix} \end{aligned}$$

Then $p = S^{-1} p' = \begin{bmatrix} \frac{1}{\alpha} & 0 \\ 0 & 1 \end{bmatrix} p'$.
Let

$$\begin{aligned} p'_0 &= T(p_0) \\ &= \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\ &= \begin{bmatrix} \alpha x_0 \\ y_0 \end{bmatrix} \end{aligned}$$

Then $p_0 = S^{-1}p'_0 = \begin{bmatrix} \frac{1}{\alpha} & 0 \\ 0 & 1 \end{bmatrix} p'_0$. So

$$\begin{aligned} 1 &= \left[S^{-1}p' - S^{-1}p'_0\right]^T A \left[S^{-1}p' - S^{-1}p'_0\right] \\ &= \left[p' - p'_0\right]^T (S^{-1})^T A S^{-1} \left[p' - p'_0\right] \end{aligned}$$

Hence the transformed ellipse $T(E)$ can be represented by

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

Now

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

So $\det(S^{-1}AS^{-1}) = \frac{1}{\alpha^2}\det(A)$. Then $T(E)$ is an ellipse. (T^{-1} is also a x-scaling, so $T^{-1}(E)$ is also an ellipse.) In addition, the area of a transformed ellipse is just a constant multiple of that of the original ellipse. So T satisfies $P1$ and $P2$. □

We can make $T_i(x, y) = S^y(1/r_{2i})S^x(1/r_{1i})R(-\phi_i)[(x, y) - (x_{0i}, y_{0i})]^T$, where $S^x(r)$ and $S^y(r)$ are scaling matrices for scaling by r along the x and y axes respectively, and $R(\theta)$ is a rotation matrix for a counter-clockwise rotation about the origin by θ. We know that rotation and translation satisfy $P1$ and $P2$, and Lemma 1 shows that scaling along the x-axis satisfies $P1$ and $P2$. (Note that scaling along an arbitrary axis can be viewed as a rotation, a scaling along the x-axis, and another rotation.) So T_i satisfies $P1$ and $P2$.

We can further simplify the problem by adding another rotation to T_i, so that $s'_{i+1} = T_i(s_{i+1})$ is on the positive x-axis. Let $s'_{i+1} = (d, 0)$. So the problem of finding the smallest ellipse E_{i+1} containing an ellipse E_i and a point s_{i+1} is reduced to that of finding the smallest ellipse E'_{i+1} containing the unit circle and a point $s'_{i+1} = (d, 0)$ on the positive x-axis. Now how do we find E'_{i+1}?

5.2 Finding the ellipse

Lemma 2. *Let E and $\overline{E}$ be two 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:
We have $F(x, y) = a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2 - 1 = 0$ for E and, negating y and y_0, $\overline{F}(x, y) = a(x - x_0)^2 - 2b(x - x_0)(y - y_0) + c(y - y_0)^2 - 1 = 0$ for $\overline{E}$. The convex combination is $\lambda F(x, y) + (1-\lambda)\overline{F}(x, y) = a(x - x_0)^2 + 2b(2\lambda - 1)(x - x_0)(y - y_0) + c(y - y_0)^2 - 1 = 0$ Now $0 < \lambda < 1$, so $-1 < 2\lambda - 1 < 1$, and hence $ac - b^2(2\lambda - 1)^2 > ac - b^2 = ac - (-b)^2$. So the area of the convex combination is smaller than that of E (and $\overline{E}$). □

Lemma 3. *The smallest ellipse E'_{i+1} containing s'_{i+1} and the unit circle will have an axis on the x-axis.*

Proof:
Assume E'_{i+1} is not aligned with the x- and y-axes. Since the unit circle and s'_{i+1} are symmetric with respect to the x-axis, there must be another ellipse $\overline{E'_{i+1}}$ with the same area, that is just E'_{i+1} reflected in the x-axis. Now E'_{i+1} and $\overline{E'_{i+1}}$ have two points on the x-axis in common. The (strict) convex combination $\lambda F'_{i+1}(x, y) + (1-\lambda)\overline{F'_{i+1}}(x, y) = 0$ of E'_{i+1} and $\overline{E'_{i+1}}$ has these properties:

- If (x, y) is on the boundaries of both E'_{i+1} and $\overline{E'_{i+1}}$, then (x, y) is on the boundary of $\lambda E'_{i+1} + (1-\lambda)\overline{E'_{i+1}}$;
- If (x, y) is in the interiors of both E'_{i+1} and $\overline{E'_{i+1}}$, then (x, y) is in the interior of $\lambda E'_{i+1} + (1-\lambda)\overline{E'_{i+1}}$.

If we set $\lambda = \frac{1}{2}$, then the convex combination will be aligned with the x-axis, it will contain the unit circle and s'_{i+1}, and it will have a smaller area than that of E'_{i+1}. □

Now consider the circle C_{i+1} that is the smallest circle containing s_{i+1} and the unit circle. See Fig. 2.

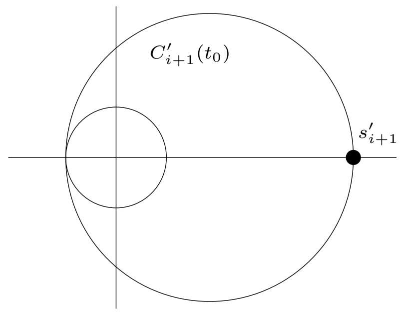

Figure 2: A circle containing and tangent to the unit circle, and passing through s'_{i+1}

We stretch C_{i+1} along the x- and y-axes, always maintaining three constraints regarding the stretched circle C'_{i+1} (which is an ellipse):

1. C'_{i+1} contains the unit circle;
2. C'_{i+1} is tangent to the unit circle at at least one point;
3. C'_{i+1} passes through s'_{i+1}.

See Figure 3. E'_{i+1} will have these properties as well. E'_{i+1} will be the C'_{i+1} with smallest area.

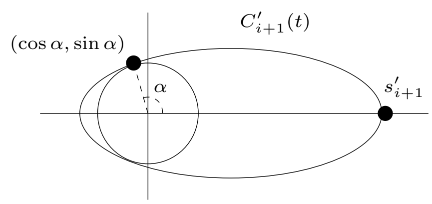

Figure 3: An ellipse containing and tangent to the unit circle, and passing through s'_{i+1}

Now we can shrink C_{i+1} 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. We can parameterize C'_{i+1} so that the ellipse $C'_{i+1}(t)$ varies in the previous manner as t varies. Define t_0 so that $C_{i+1} = C'_{i+1}(t_0)$. 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 $C'_{i+1}(t)$ remains constant, but the major axis increases in length. So $t < t_0 \Rightarrow area(C'_{i+1}(t)) > area(C'_{i+1}(t_0))$.

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

Then $\exists t' \in [t_0, t_2]$ such that $E'_{i+1} = C'_{i+1}(t')$. We should be able to write the area of $C'_{i+1}(t)$ in terms of just t, and then minimize that expression to find E'_{i+1}.

There is yet another restriction we can put on t. As we increase t from t_0, there is some value t_1 such that we can no longer shrink $C'_{i+1}(t)$ along the y-axis while fixing it through s'_{i+1} and $(-1, 0)$. See Fig. 5. If

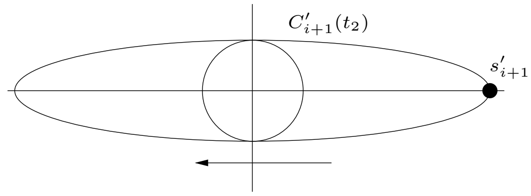

Figure 4: $C'_{i+1}(t)$ corresponding to $t = t_2$

$t \in [t_0, t_1)$, then $area(C'_{i+1}(t)) > area(C'_{i+1}(t_1))$. So we can focus on $t \in [t_1, t_2]$.

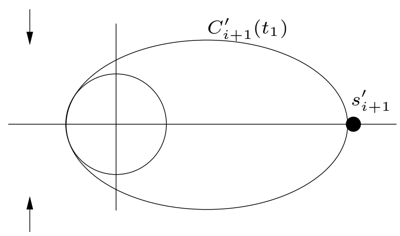

Figure 5: $C'_{i+1}(t)$ corresponding to $t = t_1$

The curvature of $C'_{i+1}(t)$ at $(r_1 \cos\theta, r_2 \sin\theta)$ is $\frac{r_1 r_2}{({r_2}^2 \cos^2\theta + {r_1}^2 \sin^2\theta)^{\frac{3}{2}}}$ [8]. 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 $C'_{i+1}(t_1)$ is then $\pi \left(\frac{d+1}{2}\right)^{\frac{3}{2}}$.

The equation of $C'_{i+1}(t)$ is $a(x - x_0)^2 + cy^2 = 1$. Since s'_{i+1} 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 $C'_{i+1}(t)$ 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

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

from the ellipse. Then it must be that $x_0 = 0$ and $c = 1$. This verifies that the area of $C'_{i+1}(t_2)$ 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 $C'_{i+1}(t)$ in terms of just $\beta = \cos\alpha$:

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

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 s'_{i+1} 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'_{i+1} is described by $a(x - x_0)^2 + cy^2 = 1$, where

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

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

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

6 Approximation ratio

To obtain an approximation ratio, we use the result of [10]. Suppose

$$ApproxEllipse \leq k \times ApproxDisk, \tag{8}$$

where k is a constant.

We have not been able to establish the value of k exactly, but extensive experimentations indicate that k lies between 5 and 6.

We know from [10] that

$$ApproxDisk \leq 9/4 MinDisk,$$

Now,

$$MinDisk \leq \frac{1}{\sqrt{1 - e^2}} MinEllipse,$$

where e is the eccentricity of the minimum spanning ellipse of the point set.

Thus, combining the above results we obtain the following approximation ratio:

$$ApproxEllipse \leq \frac{9k}{4\sqrt{1 - e^2}} MinEllipse, \tag{9}$$

If we set $k = 6$, we get

$$ApproxEllipse \leq \frac{27}{2\sqrt{1 - e^2}} MinEllipse, \tag{10}$$

Observation 1. *Let E be an ellipse centered at the origin, with major axis on the x-axis (r_1 and r_2 being the semi-major and semi-minor axis lengths respectively). Consider all points that are a fixed distance (greater than r_2) from the origin, and outside E. A point that maximizes the area of the smallest ellipse containing the point and E will lie on the y-axis.*

The reason for this is that scaling E to a circle will involve shrinking along the x-axis (or expanding along the y-axis). If the point is on the y-axis, then the distance between the circle and the transformed point is maximized, and hence the area of the smallest ellipse containing the circle and the transformed point is maximized.

We can use this observation to attempt to create a very bad approximate ellipse. Assuming that the exact minimum ellipse is the unit circle, how can we construct a corresponding sequence of points that will make the approximate ellipse as large as possible? One example of a bad point set is shown in Figure 6. The algorithm finds the left and right points first, then finds the points going from the origin to the top of the circle, and finally the points from the origin to the bottom of the circle. As we increase the number of points on the y-axis, the ratio of the areas of the approximate and exact ellipses gets bigger. See Figure 7.

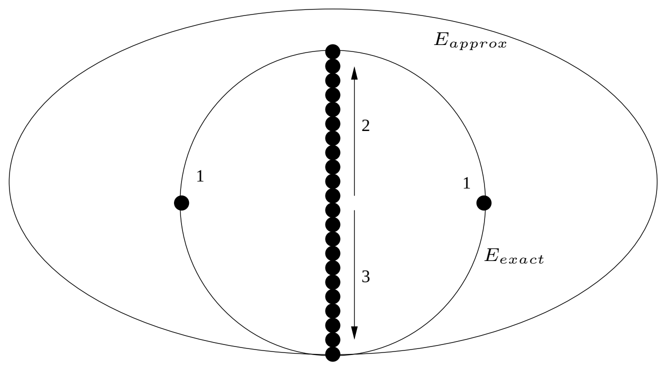

Figure 6: An undesirable sequence of points

# Points	Ratio
2×10^2	3.082697164225789
2×10^3	4.418103486895557
2×10^4	5.14471378862369
2×10^5	5.267371718841295
2×10^6	5.280568308868795
2×10^7	5.281897415706886
2×10^8	5.2820305804653405
2×10^9	5.28204328082953

Figure 7: Area ratios (approximate ellipse / exact ellipse)

Lemma 4. *Assume there is some (possibly infinite) sequence of points $< s_n >=< s_1, s_2, ... >$ that maximizes the area of $E_{approx}(< s_n >)$, subject to the constraint that the exact minimum enclosing ellipse $E_{exact}(S)$, of the set $S = \{s_1, s_2, ...\}$, is the unit circle.*

Let r_1 and r_2 be half of the axis lengths of $E_{approx}(< s_n >)$. Then $\frac{area(E_{approx}(<s_n>))}{area(E_{exact}(S))} = r_1 r_2$ is the approximation ratio (in terms of area) of the approximate ellipse algorithm.

Proof:
Assume that the above sequence of points $< s_n >$ exists.

Consider any bounded set $S' \subset \Re^2$, and any sequence $< s'_n >$ corresponding to S'. If the exact minimum ellipse $E_{exact}(S')$ is just a segment, then the approximate ellipse algorithm is exact. Otherwise, $E_{exact}(S')$ can be transformed by some T to the unit circle. Apply the same transformation to S' and $< s'_n >$, and call the results $T(S')$ and $T(< s'_n >)$.

Now $T(E_{approx}(< s'_n >)) = E_{approx}(T(< s'_n >))$ (and $T(E_{exact}(S')) = E_{exact}(T(S'))$), so

$$\begin{aligned}\frac{area(E_{approx}(< s'_n >))}{area(E_{exact}(S'))} &= \frac{area(T(E_{approx}(< s'_n >)))}{area(T(E_{exact}(S')))} \\ &= \frac{area(E_{approx}(T(< s'_n >)))}{area(T(E_{exact}(S')))}\end{aligned}$$

Obviously $area(E_{approx}(T(< s'_n >))) \leq area(E_{approx}(< s_n >))$, so

$$\begin{aligned}\frac{area(E_{approx}(< s'_n >))}{area(E_{exact}(S'))} &\leq \frac{area(E_{approx}(< s_n >))}{area(T(E_{exact}(S')))} \\ &= r_1 r_2\end{aligned}$$

□

6.1 Experimental results

We ran the algorithm on randomly generated (using Sun Java's Random class) point sets. There were ten different sets, each one with 50 points. Each set was randomly permuted 1000 times. For each permutation $< s_1, s_2, ..., s_n >$ of each set: For each pair of ellipses $E_{exact}(s_1, s_1, ..., s_i)$ and $E_{approx}(< s_1, s_2, ..., s_i >)$, we found the ratio of the respective areas, and stored it if it was the maximum for this permutation of this set. The frequencies of maximum ratios are shown in Figure 8.

7 Higher dimensions

We can always translate, rotate, and scale a non-degenerate D-dimensional ellipse E_i into the D-dimensional unit ball, and rotate so that the new point s_{i+1} is on the positive x_1-axis. Now the new ellipse E'_{i+1} will have to be symmetric with respect to the x_1-axis. The plane containing the x_1-axis and some other axis x_j will look the same, no matter what other axis we use. Therefore it must be that E'_{i+1} has the representation

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

$$\left[\begin{array}{cccc} x_1 - x_0 & x_2 & ... & x_D \end{array} \right] \left[\begin{array}{cccc} a & 0 & ... & 0 \\ 0 & c & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & ... & 0 & c \end{array} \right] \left[\begin{array}{c} x_1 - x_0 \\ x_2 \\ \vdots \\ x_D \end{array} \right] = 1,$$

where x_0, a, and c are the same as those in equations (5), (6), and (7). The approximate ellipse will be exact at least up to and including the $(D+1)$th point.

Let the current ellipse be $[p - p_{0i}]^T A_i [p - p_{0i}] = 1$. Assuming we have the right half $H_i = S_i R_i$ of the current matrix $A_i = H_i^T H_i$,

1. $s''_{i+1} \leftarrow H_i [s_{i+1} - p_{0i}]$
2. $R_{i+1} \leftarrow$ the matrix that will rotate s''_{i+1} onto the positive x_1-axis
3. $d \leftarrow distance(origin, s''_{i+1})$; x_0, a, c are as in equations (5), (6), (7)

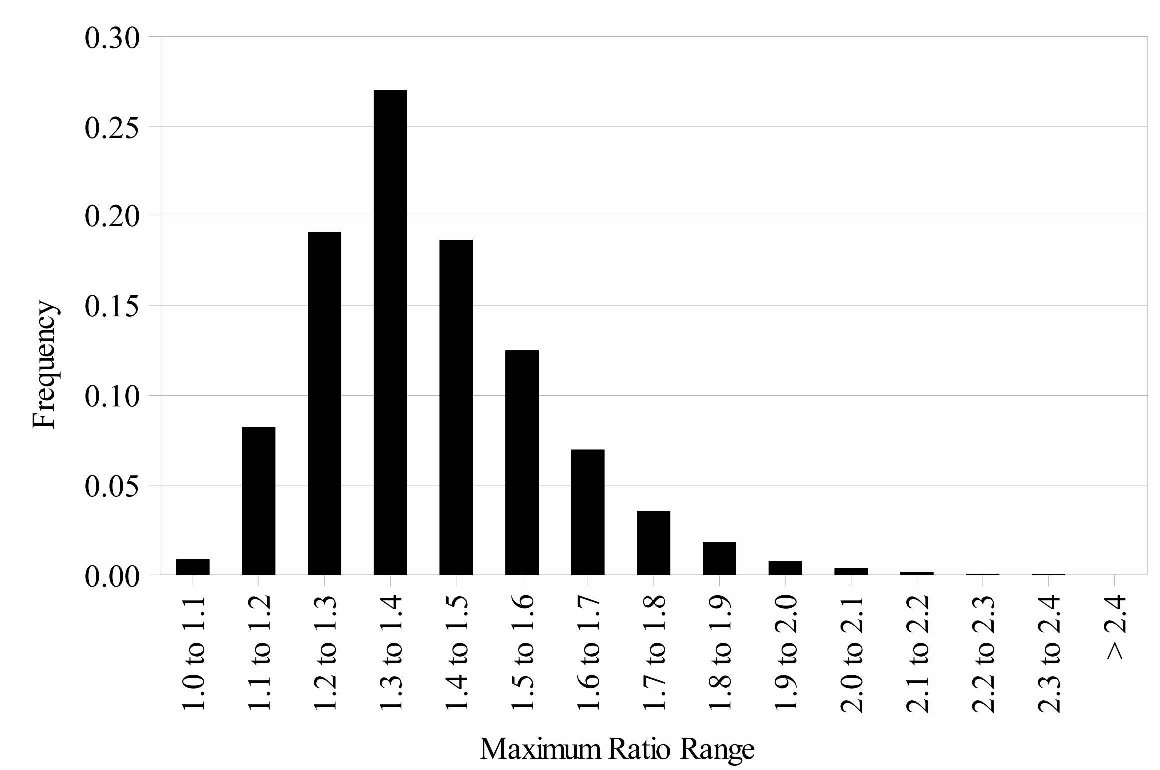

Figure 8: Maximum ratio frequencies for 1000 permutations of ten point sets

4. $S_{i+1} \leftarrow \begin{bmatrix} \sqrt{a} & 0 & ... & 0 \\ 0 & \sqrt{c} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & ... & 0 & \sqrt{c} \end{bmatrix}$; ${p_o}'_{i+1} \leftarrow (x_0, 0, ..., 0)$

5. Solve for p_{0i+1} in $R_{i+1}H_i[p_{0i+1} - p_{0i}] = {p_0}'_{i+1}$

6. $H_{i+1} \leftarrow S_{i+1}R_{i+1}H_i$

Then E_{i+1} has equation $[p - p_{0i+1}]^T H_{i+1}^T H_{i+1}[p - p_{0i+1}] = 1$. If steps 2 and 5 can be done in $O(D^2)$ space and $O(D^\omega)$ time (where $O(D^\omega)$ is the cost of multiplying $D \times D$ matrices), then only this much space and time are required for each new point.

8 Conclusions

The main open question is to obtain an approximation ratio that does not involve the eccentricity of the minimum spanning ellipse.

We have implemented this algorithm. This impementation can be viewed at `http://cs.uwindsor.ca/~asishm` by clicking on the link software.

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

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

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

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

[10] 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.