Small Weak Epsilon Nets

Boris Aronov* Franz Aurenhammer† Ferran Hurtado‡ Stefan Langerman§
David Rappaport¶ Shakhar Smorodinsky‖ Carlos Seara**

1 Introduction

Let P be a set of n points in $\mathbb{R}^2$. A point q (not necessarily in P) is called a *centerpoint* of P if each closed halfplane containing q contains at least $\frac{n}{3}$ points of P. Note that, equivalently, any convex set that contains more than $\frac{2}{3}n$ points of P also has to contain q. It is a well known fact that a centerpoint always exists and the constant $\frac{2}{3}$ is the best possible (see, e.g., [6] for more details). Can we improve this constant by using, say, two points, or some other constant number of points? What happens when we replace convex sets by, say, axis-parallel rectangles? Let us start by generalizing the notion of a centerpoint.

Definition 1 ([5]) *Let P be an n-point set in $\mathbb{R}^2$. Consider a family $\mathcal{S}$ of sets in $\mathbb{R}^2$. A set $Q \subset \mathbb{R}^2$ is called a* weak ϵ-net *for P (with respect to $\mathcal{S}$) if for any $S \in \mathcal{S}$ with $|S \cap P| > \epsilon n$ we have $S \cap Q \neq \emptyset$. Further, if $Q \subseteq P$, then Q is called a (strong) ϵ-net for P with respect to $\mathcal{S}$.*

Any centerpoint of P constitutes a 1-point weak $\frac{2}{3}$-net of P with respect to the class of convex sets. The concepts of ϵ-net and weak ϵ-net were first defined by Haussler and Welzl [5] and quickly found many applications in range searching, approximation algorithms and geometric optimization. When the VC-dimension of the range space $(P, \mathcal{S})$ is $d < \infty$, it is known that there exists an ϵ-net (and so a weak ϵ-net) of size $O((d/\epsilon)\log(d/\epsilon))$ [5]. When $\mathcal{C}$ is the family of all convex sets, the VC-dimension of the range space $(P, \mathcal{C})$ is infinite and the previous result does not apply. Nevertheless, it is known that for any ϵ and for any set P of points in the plane, there exists a weak ϵ-net for P with respect to $\mathcal{C}$ with size at most $O(\frac{1}{\epsilon^2})$, see [1]. The best known lower bound is the trivial $\Omega(\frac{1}{\epsilon})$ bound, which holds already when all points are on a line. See, e.g., [6] for more details on weak ϵ-nets.

In this paper, we will mainly consider weak ϵ-nets of small constant size. Let $0 < \epsilon_i^{\mathcal{S}} < 1$ denote the minimum real number such that for any finite point set P there exists a set Q of i points that is an weak $\epsilon_i^{\mathcal{S}}$-net for P with respect to $\mathcal{S}$. We provide upper and lower bounds for $\epsilon_i^{\mathcal{S}}$ for small sizes i, when $\mathcal{S}$ is the family of all convex sets or the family of all axis-parallel rectangles.

2 Convex sets

Let $\mathcal{C}$ denote the family of all convex sets in the plane. In this section, we derive various bounds on the quantity $\epsilon_i^{\mathcal{C}}$, for $i \geq 2$. We start by proving a lower bound on $\epsilon_2^{\mathcal{C}}$ and $\epsilon_3^{\mathcal{C}}$.

Theorem 2 $\epsilon_2^{\mathcal{C}} \geq \frac{5}{9}$ *and* $\epsilon_3^{\mathcal{C}} \geq \frac{5}{12}$.

Proof. For any n, we construct a set P of n points such that, for any pair of points q, r in the plane, there is a convex set K that avoids q, r and contains at least $\frac{5}{9}n$ points of P. See Figure 1. The

*Department of Computer and Information Science, Polytechnic University, Brooklyn, NY 11201, USA; http://cis.poly.edu/~aronov/; research supported by grant SAB2003-0097, Min. Educ. y Ciencia, España. Also partially supported by NSF ITR Grant CCR-00-81964 and by a grant from US-Israel Binational Science Foundation.

†Institute for Theoretical Computer Science, Graz University of Technology, Graz, Austria; auren@igi.tugraz.at.

‡Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, ferran.hurtado@upc.edu, partially supported by projects MCYT-FEDER-BFM2003-00368, Gen-Cat-2001SGR00224 and Acción Integrada Austria-España MCYT HU2002-0010.

§Chercheur qualifié du FNRS, Département d'Informatique, Université Libre de Bruxelles, stefan.langerman@ulb.ac.be.

¶School of Computing, Queen's University, Kingston, ON K7L 3N6, Canada; daver@cs.queensu.ca.

‖Courant Institute for Mathematical Sciences, New York University, 251 Mercer st, New York, NY, 10012, USA; http://www.cims.nyu.edu/~shakhar/; shakhar@cims.nyu.edu. Work on this paper was supported by the NSF Mathematical Sciences Postdoctoral Fellowship award 0402492.

**Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, carlos.seara@upc.edu, partially supported by projects MCYT-FEDER-BFM2003-00368, Gen-Cat-2001SGR00224 and Acción Integrada Austria-España MCYT HU2002-0010.

set P comes in three groups, each group consisting of three subsets arranged in a triangular shape. Each of the nine subsets, call them $1, 2, \dots, 9$, lies in some disk of diameter δ (for δ sufficiently small) and contains $\frac{n}{9}$ points. For any choice of q and r, let L

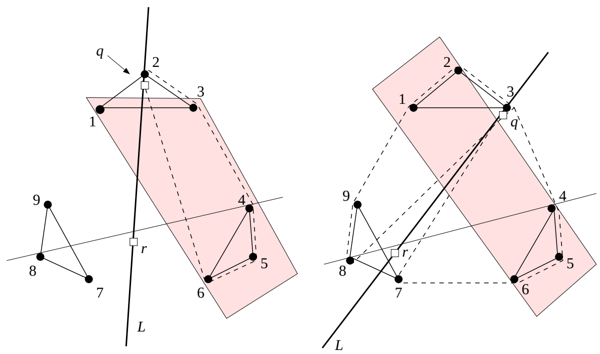

Figure 1: Lower bound construction for $\epsilon_2^{\mathcal{C}}$

be the line through q and r. By construction of P, line L can intersect the convex hull of at most two of the subsets $1, \dots, 9$. We may assume that L has at least three out of the nine subsets fully contained on each side. We may also assume that L intersects the convex hull of at least one subset. For, otherwise, at least $\frac{5}{9}n$ points of P lie in a fixed open halfplane bounded by L. Write $CH(i, j, \dots)$ for the convex hull of the point set $i \cup j \cup \dots$. Without loss of generality, assume that L intersects $CH(1, 2, 3)$. Then, in order to stab $CH(4, 5, 6, 7, 8)$, one of the points in question, say r, has to lie on or below the upper tangent of $CH(4)$ and $CH(8)$. For placing the point q, two cases remain.

Case (1) Line L intersects $CH(2)$; see Figure 1, left side. Exploiting symmetries, it is no loss of generality to assume that L be not closer to 6 than to 7. As we must have $q \in L \cap CH(2, 3, 4, 5, 6)$, q will be arbitrarily close to $CH(2)$ if the disk we assumed to contain the set 2 becomes arbitrarily small. So, for sufficiently small disk diameter δ, $K = CH(1, 3, 4, 5, 6)$ will avoid both q and r.

Case (2) Line L intersects $CH(3)$ (or, symmetrically, $CH(1)$); see Figure 1, right side. If L is not closer to 8 than to 7, then we need $q \in L \cap CH(1, 2, 3, 8, 9)$. Otherwise, we need $q \in L \cap CH(3, 4, 5, 6, 7)$. In both cases, q becomes arbitrarily close to $CH(3)$ if δ is chosen to be sufficiently small. But now $K = CH(1, 2, 4, 5, 6)$ avoids both q and r.

In conclusion, for any two given points, we can find a convex set K that avoids both points and satisfies $|K \cap P| \geq \frac{5}{9}n$. The lower bound for $\epsilon_3^{\mathcal{C}}$ follows from circularly placing 4 'triangular shapes' instead of the 3 and using similar arguments. Details appear in the full version. $\square$

We now turn to upper bounds. For arbitrary sets of n points in the plane, we want to construct weak ϵ-nets of given size i and with 'deficiency' ϵ as small as possible. The tools we use are ham-sandwich cuts [6], as well as weak ϵ-nets of size at most $i-1$ that we will have already shown to exist (starting with $i = 1$, the well-known centerpoint).

The following terminology is used. Let P be any n-point set in the plane. Let l be a vertical line that splits the set P into two subsets of, say, r red points and b blue points. Denote by h a line that simultaneously halves the red subset and the blue subset. This so-called ham-sandwich line h is well known to exist. Finally, define $q_0 = l \cap h$. Refer to Figure 2 where the constructions described below are shown schematically. The point q_0 is drawn in square shape, and the red side of h is shaded.

Theorem 3 *$\epsilon_2^{\mathcal{C}} \leq \frac{5}{8}$, $\epsilon_3^{\mathcal{C}} \leq \frac{7}{12}$, $\epsilon_4^{\mathcal{C}} \leq \frac{4}{7}$, and $\epsilon_5^{\mathcal{C}} \leq \frac{1}{2}$.*

Proof. Let us first prove $\epsilon_2^{\mathcal{C}} \leq \frac{5}{8}$. Choose the vertical line l such that $r = \frac{n}{4}$ (and thus $b = \frac{3n}{4}$). Let q_1 be a centerpoint for the blue subset of P. We claim that the set $\{q_0, q_1\}$ is a weak $\frac{5}{8}$-net for P.

Let K be any convex set with $q_0, q_1 \notin K$. As $q_0 \notin K$, the set K avoids at least one of the four quadrants defined by the line l and the halving line h. (By convexity, K would contain q_0, otherwise.) If this quadrant is blue then K avoids at least $\frac{3}{8}n$ (blue) points. If this quadrant is red then K avoids at least $\frac{1}{8}n$ red points. In addition, as $q_1 \notin K$, and q_1 is a centerpoint for the blue points, K also avoids at least $\frac{1}{3} \cdot \frac{3n}{4} = \frac{1}{4}n$ blue points. Altogether, K avoids at least $\frac{3}{8}n$ points again. So, in both cases, K cannot contain more that $\frac{5}{8}n$ points of P.

Next, we show $\epsilon_3^{\mathcal{C}} \leq \frac{7}{12}$. To this end, choose line l such that $r = \frac{n}{2}$. Then each of the quadrants defined by the lines l and h contains $\frac{n}{4}$ points of P. Take q_1 as a centerpoint for the red points, and take q_2 as a centerpoint for the blue points. Put $Q = \{q_0, q_1, q_2\}$. Then Q is a weak $\frac{7}{12}$-net for P by the arguments below.

Let K be any convex set that avoids Q. Since K does not contain q_0, it must avoid some quadrant. Assume, without loss of generality, that this quadrant is blue. Then K can contain at most $\frac{n}{4}$ blue points. Since K avoids q_1, and q_1 is a centerpoint for the red points, K contains at most $\frac{2}{3} \cdot \frac{n}{2}$ red points. In total, at most $(\frac{1}{3}+\frac{1}{4})\cdot n = \frac{7}{12}n$ points of P can lie in K.

To see that $\epsilon_4^{\mathcal{C}} \le \frac{4}{7}$ holds, we proceed as below. Choose $r = \frac{n}{7}$, which gives $b = \frac{6n}{7}$. From above, we know that a weak $\frac{7}{12}$-net of size three exists for the blue points. Let Q_b be such a net. We consider the set $Q = Q_b \cup \{q_0\}$ that will be a weak $\frac{4}{7}$-net for P.

If a convex set K avoids Q then it avoids one quadrant defined by l and h. If it is a blue quadrant then K contains at most $\frac{4}{7}n$ points of P. If this quadrant is red then K contain at most $(\frac{1}{14}+\frac{7}{12}\cdot\frac{6}{7})\cdot n = \frac{4}{7}n$ points of P as well. This follows from the fact that K contains at most one red quadrant and at most a fraction of $\frac{7}{12}$ of the blue points.

Finally, let us argue that $\epsilon_5^{\mathcal{C}} \le \frac{1}{2}$. As done for the net of size three, we choose l and h such that each quadrant contains $\frac{n}{4}$ points of P. For the corresponding four subsets P_j of P, let q_j be a centerpoint, for $1 \le j \le 4$. Then $Q = \{q_0, \ldots, q_4\}$ represents a weak $\frac{1}{2}$-net for P.

Each convex set K that avoids Q totally avoids one subset, say P_1. In addition, K avoids a fraction of $\frac{1}{3}$ of the points in each of P_2, P_3, and P_4, because the centerpoints of these subsets are in Q. Thus at least $n\cdot(\frac{1}{4}+\frac{1}{3}\cdot\frac{3}{4}) = \frac{n}{2}$ points are avoided by K. $\square$

Note that there exist other possibilities of combining ham-sandwich cuts with weak ϵ-nets. For example, when constructing a weak ϵ-net of size 3, we could use a weak $\frac{5}{8}$-net of size 2 rather than two centerpoints. Then the best vertical cut, $r = \frac{1}{5}$, evaluates to $\epsilon = \frac{3}{5}$, which is slightly worse than $\epsilon = \frac{7}{12}$ obtained in the proof above. For weak ϵ-nets of size 4 there also is another construction, namely, using one centerpoint and a size-2 net. We obtain the same bound, however. For size 5, no other construction competes with $\epsilon = \frac{1}{2}$.

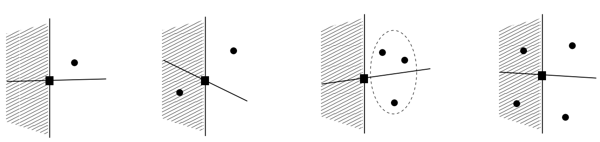

Figure 2: Small weak ϵ-nets for convex sets

In general, it is preferable to use nets that are as small as possible. To obtain an upper bound on $\epsilon_i^{\mathcal{C}}$ for arbitrary net size i, we may apply the construction for $\epsilon_5^{\mathcal{C}}$ recursively. This evaluates to

$$\epsilon_i^{\mathcal{C}} \le \frac{2}{3}\cdot\left(\frac{3}{4}\right)^k \text{ for } i = \frac{1}{3}\cdot(4^{k+1}-1), \quad k \ge 0.$$

A rough calculation shows that a weak ϵ-net of size $O(\frac{1}{\epsilon^5})$ with respect to $\mathcal{C}$ is obtained. Unfortunately (but not surprisingly) this by far falls short of the best known bound $O(\frac{1}{\epsilon^2})$ in [1]; see also [2]. Still, for small nets, our constructions are superior. For example, to achieve $\epsilon = \frac{1}{2}$ a net of size eight (rather than five) is needed in [1, 2].

3 Axis-parallel rectangles

This section presents bounds on $\epsilon_i^{\mathcal{R}}$, where $\mathcal{R}$ denotes the family of all axis-parallel rectangles in the plane.

Theorem 4 $\epsilon_1^{\mathcal{R}} = \frac{1}{2}$, $\epsilon_2^{\mathcal{R}} = \frac{2}{5}$, $\epsilon_3^{\mathcal{R}} = \frac{1}{3}$, $\epsilon_4^{\mathcal{R}} \le \frac{5}{16}$, $\epsilon_5^{\mathcal{R}} \le \frac{1}{4}$, $\epsilon_8^{\mathcal{R}} \le \frac{1}{5}$, $\epsilon_{10}^{\mathcal{R}} \le \frac{1}{6}$, $\epsilon_{12}^{\mathcal{R}} \le \frac{1}{7}$, $\epsilon_{6x+4}^{\mathcal{R}} \le \epsilon_x^{\mathcal{R}}/3$.

Proof. We only present here the proof of the second equality. The other proofs appear in the full version. We start by showing that $\epsilon_2^{\mathcal{R}} \le \frac{2}{5}$. Let l_1 be a vertical line with exactly $\frac{2}{5}n$ points of P to its left and let l_2 be a vertical line with exactly $\frac{2}{5}n$ points of P to its right. Similarly consider a line μ_1 (resp., μ_2) with exactly $\frac{2}{5}n$ points of P below it (resp., above it). Let q_1, q_2, q_3, q_4 be the vertices of the rectangle formed by the intersection points of these lines. See Figure 3 for an illustration. We will show that at least one of the sets $Q_1 = \{q_1, q_3\}$ or the set $Q_2 = \{q_2, q_4\}$ is a weak $\frac{2}{5}$-net for P. Assume to the contrary that neither of these sets is a weak $\frac{2}{5}$-net for P. Observe that every axis parallel rectangle that contains at least $\frac{2}{5}n$ points of P must contain one of the vertices $\{q_1, q_2, q_3, q_4\}$. Assume that there exists a rectangle avoiding the set Q_1 and containing more than $\frac{2}{5}n$ points of P. Such a rectangle must lie either above of q_1 and to the left of q_3 or to the right of q_1 and below q_3. Assume without loss of generality that the later situation occurs. Symmetrically, there must exists a rectangle below q_2 and to the left of q_4 that contains at least $\frac{2}{5}n$ points of P. Let A, B, C, D, E, F be the number of points in each of the six rectangles induces by the arrangement of the lines l_1, l_2, μ_1, μ_2 that lie below μ_2. See Figure 3. We have:

$$A + B + C = \frac{n}{5} \tag{1}$$

$$D + E + F = \frac{2n}{5} \tag{2}$$

$$B + C + E + F > \frac{2n}{5} \tag{3}$$

$$A + B + D + E > \frac{2n}{5} \tag{4}$$

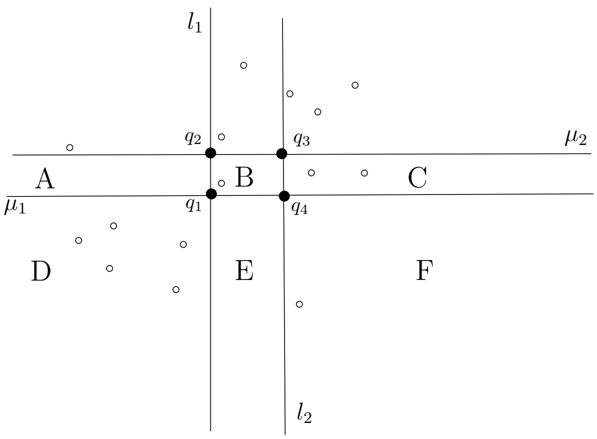

Figure 3: The two vertical lines l_1, l_2 and two horizontal lines μ_1, μ_2 intersect in four points q_1, q_2, q_3, q_4.

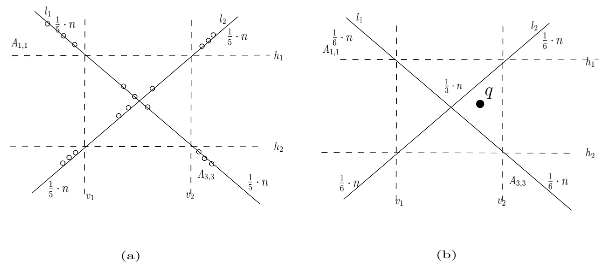

Figure 4: (a) A lower bound construction showing that $\epsilon_2^{\mathcal{R}} \geq \frac{2}{5}$. (b) A similar lower bound construction showing $\epsilon_3^{\mathcal{R}} \geq \frac{1}{3}$.

Notice that by summing the last two inequalities and subtracting the first two equalities we get $B + E > \frac{n}{5}$, a contradiction. Hence, $\epsilon_2^{\mathcal{R}} \leq \frac{2}{5}$.

To show that $\epsilon_2^{\mathcal{R}} \geq \frac{2}{5}$, we place a set P of n points on two diagonal lines l_1, l_2 as depicted in Figure 4(a). Let h_1 be a horizontal line with $\frac{2}{5}n$ points above it. Similarly h_2 has $\frac{2}{5}n$ points below it, v_1 is vertical and has $\frac{2}{5}n$ points to its left and v_2 is vertical and has $\frac{2}{5}n$ points to its right. The four lines h_1, h_2, v_1, v_2 partition the plane into 9 axis parallel rectangles. Denote those rectangles by $A_{i,j}$ for $i, j = 1, \ldots, 3$ where $A_{i,j}$ is the rectangle defined by the i'th row and the j'th column. If a pair of points q_1, q_2 is a weak $\frac{2}{5}$-net for P with respect to axis parallel rectangles, then it is easily seen that either $A_{1,3} \cup A_{3,1}$ contains $\{q_1, q_2\}$ or $A_{1,1} \cup A_{3,3}$ contains $\{q_1, q_2\}$. Assume without loss of generality the former case. Then $A_{1,1} \cup A_{1,2} \cup A_{2,1} \cup A_{2,2}$ is an axis-parallel rectangle containing at least $\frac{2}{5}n$ points of P and avoiding $\{q_1, q_2\}$, a contradiction. The other case is treated similarly. $\square$

4 Remarks

It is interesting to note that some results on weak ϵ-nets follow rather directly from classical results. We illustrate this fact for $\mathcal{D}$, the collection of all discs in the plane.

Theorem 5 $\epsilon_4^{\mathcal{D}} \leq \frac{1}{2}$.

Proof. Let P be a set of n points in the plane. We need to show that there exists a set Q of four points such that every disc d for which $|d \cap P| > \frac{n}{2}$ must intersect Q. Consider the collection $\mathcal{D}' \subset \mathcal{D}$ of all discs d that contains more than $\frac{n}{2}$ points of P. Obviously every pair of discs of $\mathcal{D}'$ must have a non-empty intersection. By the result of [3], there exists a set Q of four points that stab all discs in $\mathcal{D}'$. This completes the proof. $\square$

In [4] it was proved that for any finite collection of pairwise intersecting unit discs, there exists three points that stab those discs. Thus, using the same analysis as in the proof above we have that $\epsilon_3^{\mathcal{U}} \leq \frac{1}{2}$, where $\mathcal{U}$ is the collection of all unit discs in the plane.

References

[1] N. Alon, I. Bárány, Z. Füredi, and D. Kleitman. Point selections and weak ϵ-nets for convex hulls. *Combin., Probab., Comput.,* 1(3):189–200, 1992.

[2] B. Chazelle. The Discrepancy Method. Cambridge University Press, 2000.

[3] L. Danzer. Zur Lösung des Gallaischen Problems über Kreisscheiben in der euklidischen Ebene. *Studia Sci. Math. Hungar.*, 21:111–134, 1986.

[4] H. Hadwiger, H. Debrunner and V. Klee. *Combinatorial Geometry in the Plane.* Holt, Reinhart and Winston, New York, 1964.

[5] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.,* 2:127–151, 1987.

[6] J. Matoušek. *Lectures in Discrete Geometry* Springer-Verlag, 2002.