Opposite-quadrant depth in the plane

Hervé Brönnimann[*]   János Pach[†]

Abstract

Given a set S of n points in arbitrary position in the plane, we define the *opposite-quadrant depth* of a point as the largest number d such that, for one of the two pairs (NW,SE) or (SW,NE) of opposite axis-aligned quadrants centered at p, each (closed) quadrant in the pair contains at least d points in S. We prove that there always is a point in S whose opposite-quadrant depth is at least $\lfloor n/8 \rfloor$. We also present a construction that cannot have a point or opposite quadrant depth bigger than $(n+2)/6$.

1 Introduction

Let S be a set of n points in arbitrary position in the plane. Given any point p, the horizontal and vertical lines through p define four quadrants $\mathrm{NW}(p)$, $\mathrm{NE}(p)$, and $\mathrm{SE}(p)$, and $\mathrm{SW}(p)$. We define the *opposite-quadrant depth* $\mathrm{opp}(p,S)$ of a point p with respect to S, or $\mathrm{opp}(p)$ when S is clearly understood, as the largest number d such that, for one of the two pairs $(\mathrm{NW},\mathrm{SE})$ or $(\mathrm{SW},\mathrm{NE})$, each (closed) quadrant in the pair contains at least d points in S. In notation:

$$\begin{aligned} \mathrm{opp}(p,S) &= \max\left(\min\left(|\mathrm{NW}(p)\cap S|, |\mathrm{SE}(p)\cap S|\right),\right.\\ &\qquad \left.\min\left(|\mathrm{SW}(p)\cap S|, |\mathrm{NE}(p)\cap S|\right)\right),\\ \mathrm{opp}(S) &= \max_{p\in\mathbb{R}^2} \mathrm{opp}(p,S). \end{aligned}$$

An *α-centerpoint* is a point p whose opposite-quadrant depth is at least $\lfloor \alpha n \rfloor$. For any given S, there is a (not necessarily unique) point of maximum opposite-quadrant depth which may be harder to compute than a centerpoint. If we do not require the centerpoint to be in S, then there is trivially a $\frac{1}{4}$-centerpoint, as the intersection of the vertical and horizontal halving lines will do. In fact, this is the best general bound that can be established, as can be seen for points uniformly distributed in a square, disk, or even circle. Note that there are sets with higher $\mathrm{opp}(S)$ (e.g. points on a diagonal line have depth $n/2$).

If the point p is restricted to be in S, however, it seems much harder to achieve opposite-quadrant depth $\mathrm{opp}(p,S)$ as high as $n/4$. Our main result is:

Theorem 1 *Let S be a set of n points in arbitrary position in the plane. There always exists a point in S of opposite-quadrant depth at least $\lfloor n/8 \rfloor$, and such a point can be computed in $O(n)$ time.*

The construction is inspired by the catline of Hubert and Rousseuw [2]. How restrictive is it to ask that the centerpoint belongs to S? Surprisingly, we can also show that the bound $n/4$ is not attainable in general.

Theorem 2 *There is a point set S such that $\max_{p\in S} \mathrm{opp}(p,S) = (n+2)/6$.*

The proof is omitted for lack of space. The best lower bound is somewhere $n/8$ and $(n+2)/6$, but we do not know or conjecture whether either bound is tight.

These notions are closely related to *hyperplane depth*, *multivariate regression depth*, and other measures of statistical depth [3]. Unlike opposite-quadrant depth, however, a centerpoint need not exist in the set itself. For instance, the hyperplane depth of any point set is well known to be at least $\left\lfloor \frac{n}{d+1} \right\rfloor$ in any dimension [1] yet if the set S is in convex position the hyperplane depth of any point in S is zero! In contrast, it is surprising that the notion of opposite-quadrant depth gives a centerpoint in the set, and even more so that the optimal fraction when restricted to the point set should differ from the unrestricted fraction $\frac{1}{4}$.

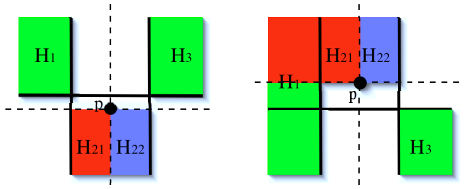
2 Proof of Theorem 1

First, notice that having points on the same vertical or horizontal line is not a problem since the quadrants are closed. Indeed, let us construct a perturbation by assigning to each point p_i a vector $\mathbf{v}_i = (i, i^2)$. Then there exists a $\varepsilon > 0$ small enough such that $S_\lambda =$

---

[*] CIS, Polytechnic University, Six metrotech, Brooklyn, NY 11201. `hbr@poly.edu`

[†] Courant Institute, Room 721, 251 Mercer Street, New York, NY 10012. `pach@cims.nyu.edu`



(a) Bitonic configuration (b) Monotone configuration

Figure 1: For the proof of Theorem 1. (a) The case of a bitonic configuration: both first and third heavy cells (labeled H_1 and H_3) have $\lfloor n/8 \rfloor$ points, and one of the two halves H_{21} and H_{22} of the middle heavy cell does as well. (b) The case of a monotone configuration: the third heavy cell (H_3) has $\lfloor n/8 \rfloor$ points, and so do either $\mathrm{NW}(p)$ (shaded in red), or $SW(p)$ and H_{22}.

$\{p_i + \lambda \mathbf{v}_i\}$ is in general position for any $\lambda \in (0, \varepsilon]$, and as a consequence the vertical and horizontal orders of the points in S_λ is the same over that interval. If the theorem holds for sets in general position, we can apply it to S_ε, and get an α-centerpoint $p_i + \lambda \mathbf{v}_i$, which is also an α-centerpoint for any $\lambda \in (0, \varepsilon]$. Bringing back points to their original position in S only adds points on the boundary of the two opposite quadrants of p_i, so p_i is also an α-centerpoint for S as well.

From now on, S will thus be assumed in general position. The proof proceeds by splitting the set S by two vertical lines such that in the left or right vertical strips, there are at least $\lfloor n/4 \rfloor$ points, and in the middle strip there are at least $\lfloor n/2 \rfloor$ points. Next, we find a median horizontal line, and this gives a 3×2 grid. We call a cell in that grid *heavy* if it contains at least half of the points in its column in the grid. Note that if all the cells on a row are heavy, then they must contain exactly half each, otherwise there wouldn't be enough points in the other row; in that case, the other cells are also heavy. Thus we can pick one heavy cell per column, such that all three are not on a single row. Hence, we may assume (up to horizontal or vertical reflexions) that the three heavy cells are in first, then second, then first row (we call this configuration *bitonic*), or in the first, then first, and then second row (the *monotone* configuration). The situation is depicted in Figure 1.

The case of the bitonic configuration is the easiest. We pick for p the highest point in the heavy middle cell (on the second row). The first and third heavy cells have at least $\lfloor n/8 \rfloor$ points of S, and the first is entirely contained in $\mathrm{NW}(p)$ while the third is in $\mathrm{NE}(p)$. The middle heavy cell contains at least $\lfloor n/4 \rfloor$ points, and is entirely contained in the halfplane $\mathrm{SW}(p) \cup \mathrm{SE}(p)$ by our choice of p. One of the two quadrants $\mathrm{SW}(p)$ or $\mathrm{SE}(p)$ must therefore contain $\lfloor n/8 \rfloor$ points. This concludes the proof for the bitonic configuration.

For the monotone configuration, we have to resort to a slightly different argument because we cannot assume any vertical relation between the points in the first and second heavy cells (they below to the same row). Nevertheless, we pick for p the lowest point in the heavy middle (on the first row). The first and third heavy cells have at least $\lfloor n/8 \rfloor$ points of S, and the halfplane $\mathrm{NW}(p) \cup \mathrm{SW}(p)$ contain the entire first column, thus at least $\lfloor n/4 \rfloor$ points. One of the two quadrants $\mathrm{NW}(p)$ or $\mathrm{SW}(p)$ must thus contain $\lfloor n/8 \rfloor$ points. Also, the middle heavy cell contains at least $\lfloor n/4 \rfloor$ points, and is entirely contained in the halfplane $\mathrm{NW}(p) \cup \mathrm{NE}(p)$ by our choice of p. Thus one of the two quadrants $\mathrm{NW}(p)$ or $\mathrm{NE}(p)$ must contain $\lfloor n/8 \rfloor$ points. If $\mathrm{NW}(p)$ does, then we are done since $\mathrm{SE}(p)$ contains the third heavy cell, at least $\lfloor n/8 \rfloor$ points. But if not, then by what precedes, both $\mathrm{NE}(p)$ and $\mathrm{SW}(p)$ must contain at least $\lfloor n/8 \rfloor$ points. This concludes the proof for the monotone configuration.

For the algorithm, it is clear that all the steps above can be performed in linear time, by using any linear-time selection algorithm.

Acknowledgements Thanks to Boris Aronov, Joshua Cooper, Raghavan Dhandapani, Jeff Erickson, John Iacono, and Radous Radoiucić for discussions and comments.

References

[1] H. Edelsbrunner. *Algorithms in Combinatorial Geometry.* Springer, Berlin, 1987.

[2] M. Hubert and P. Rousseeuw. The catline for deep regression. *J. Multivariate Analysis*, 66(2):270-296, 1998.

[3] R. Liu, J. Parelius, and K. Singh. Multivariate analysis by data depth: Descriptive statistics, graphics and inference. *Ann. Stat.* 27:783-858, 1999. Special Invited Paper with discussions.