On the Bichromatic k-Set Problem

Timothy M. Chan*

September 2, 2009

Abstract

We study a bichromatic version of the well-known *k-set problem*: given two sets R and B of points of total size n and an integer k, how many subsets of the form $(R \cap h) \cup (B \setminus h)$ can have size exactly k over all halfspaces h? In the dual, the problem is asymptotically equivalent to determining the worst-case combinatorial complexity of the *k-level* in an arrangement of n *halfspaces*.

Disproving an earlier conjecture by Linhart (1993), we present the first nontrivial upper bound for all $k \ll n$ in two dimensions: $O(nk^{1/3} + n^{5/6-\varepsilon}k^{2/3+2\varepsilon} + k^2)$ for any fixed $\varepsilon > 0$. In three dimensions, we obtain the bound $O(nk^{3/2} + n^{0.5034}k^{2.4932} + k^3)$. Incidentally, this also implies a new upper bound for the original k-set problem in four dimensions: $O(n^2k^{3/2} + n^{1.5034}k^{2.4932} + nk^3)$, which improves the best previous result for all $k \ll n^{0.923}$. Extensions to other cases, such as arrangements of disks, are also discussed.

1 Introduction

Two dimensions. In the standard *k-set* problem, we consider a set P of n points in the plane and count the number of distinct subsets of the form $P \cap h$ that have size exactly k over all halfplanes h. The goal is to bound the maximum value of this number asymptotically as a function of n and k. This is one of the all-time favorite open problems in combinatorial geometry. The case $k = n/2$ in particular is known as the *halving lines* problem. An asymptotically equivalent dual formulation is known as the *k-level* problem: for an arrangement of n lines, bound the number of vertices (or faces) that have level k, where the level of a point (or face) refers to the number of lines below it. The problem is also equivalent to bounding the number of combinatorial changes of the k-th smallest element in a set of n linearly moving points in one dimension. These different points of view, the close relationship to numerous fundamental algorithmic problems in computational geometry, as well as the simplicity of the original problem statement itself, explain the problem's immense popularity [16, 24, 28].

In the early 70s, Lovász [22] and Erdős *et al.* [19] established the first upper bound of $O(n\sqrt{k})$ and lower bound of $\Omega(n \log k)$. At FOCS'89, Pach, Steiger, and Szemerédi [29] obtained the first improved upper bound of $O(n\sqrt{k}/\log^* k)$. At FOCS'97, Dey [14] obtained a substantially better $O(nk^{1/3})$ upper bound, which has remained the current record. This upper bound is still far from

*School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada (tmchan@uwaterloo.ca).
Work supported by NSERC. A preliminary version of this paper has appeared in *Proc. 19th ACM-SIAM Sympos. Discrete Algorithms*, pages 561–570, 2008.

the current best lower bound of $n2^{\Omega(\sqrt{\log k})}$, due to Tóth [36]. Erdős *et al.* [19] conjectured that the true answer is $o(n^{1+\varepsilon})$ for any fixed constant $\varepsilon > 0$.

In this paper, we will not make progress on this longstanding problem (regrettably). Instead we study the following a *bichromatic* variant which we feel is just as interesting and has been unfairly neglected:

> Consider a set R of red points and a set B of blue points in the plane with $|R| + |B| = n$, and count the number of subsets of the form $(R \cap h) \cup (B \setminus h)$ that have size exactly k over all halfplanes h. Bound the maximum value of this number asymptotically as a function of n and k.

Asymptotically, this is equivalent to bounding the number of lines that pass through two points and "misclassify" exactly k points, where a line ℓ misclassifies a red (resp. blue) point p iff ℓ is above (resp. below) p. In the dual, the problem has an even more natural formulation:

> For an arrangement of n (closed) halfplanes in $\mathbb{R}^2$, bound the number of vertices that have level k, where the level of a point q now refers to the number of halfplanes that q lies outside.

The original k-level/k-set problem corresponds to the special case where all the halfplanes are lower halfplanes (i.e., contain $(0, -\infty)$), since the k-level in an arrangement of lower halfplanes is the same as the k-level in the arrangement of the lines bounding the halfplanes. In our problem, however, we can have both upper halfplanes (i.e., those containing $(0, \infty)$) and lower halfplanes, and the halfplanes do not necessarily have a common intersection. The problem has obvious connections with computational problems related to linear programming with violations (e.g., enumerating all basic solutions that violate k constraints) and geometric optimization with outliers (see Section 5 for more on the algorithmic significance).

As a function of n alone it is not possible to obtain a subquadratic bound for this variant, unlike the original k-set problem, as the example in Figure 1 illustrates. This fact perhaps explains why despite its naturalness, our problem has not been as intensively studied. However, the related bichromatic versions of the $(\le k)$*-set*/$(\le k)$*-level* problem (counting the total number of vertices of levels between 0 and k in an arrangement of n halfplanes) has been considered by several researchers. Sharir [30] observed that Clarkson and Shor's technique [13] can be applied to yield asymptotically tight bounds for the $(\le k)$-level for general classes of arrangements, including arrangements of closed Jordan curves in the plane. In the case of halfplanes, the bound is $\Theta(nk)$, just as in the case of lines. For $(\le k)$-levels of halfspaces in $\mathbb{R}^d$, the bound is $\Theta(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil})$ (as in the case of hyperplanes); the constant factors have recently been improved by Wagner [37]. These upper bounds for $(\le k)$-levels obviously translate to upper bounds for k-levels, but these authors did not investigate the corresponding k-set/k-level problem (perhaps believing no nontrivial bounds are possible).

A paper by Linhart [21] actually did address our problem of k-levels for halfplanes in $\mathbb{R}^2$. However, he was interested in deriving $O(nk)$-type upper bounds with smaller constant factors. In fact, the paper explicitly contained the conjecture that $O(nk)$ is asymptotically tight for all $0 \le k \le n/2$.

In Section 2, we disprove Linhart's conjecture by proving the following interesting upper bound for the 2-d bichromatic k-set problem or k-levels of halfplanes:

$$O(nk^{1/3} + n^{5/6-\varepsilon}k^{2/3+2\varepsilon} + k^2).$$

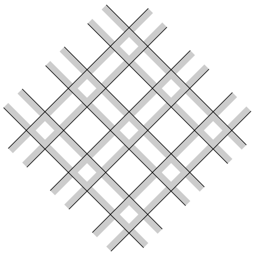


Figure 1: An arrangement of halfplanes where a level may have a quadratic number of vertices.

(Throughout the paper, ε denotes an arbitrarily small positive constant.) For all $k = o(n)$, this bound is strictly better than $O(nk)$. For $k \ll \sqrt{n}$, the first term dominates and our bound is the best possible with current knowledge, in the sense that it matches Dey's upper bound for the original k-set/k-level problem. On the other hand, for $k \gg n^{5/8+\varepsilon}$, the last term dominates and our bound is the best possible as well, because $\Omega(k^2)$ is an obvious lower bound by a construction similar to Figure 1.

Three dimensions. In 3-d, the k-set problem has received just as much (if not more) attention. Bárány, Füredi, and Lovász [5] gave the first proof of a nontrivial upper bound $O(n^{3-1/343}) = O(n^{2.9971})$. Shortly after, Aronov *et al.* [4] improved it to $O(n^{8/3} \operatorname{polylog} n)$, where the polylogarithmic factor was then further refined by Eppstein [18]. Dey and Edelsbrunner [15] later completely removed the polylogarithmic factor. Their result had remained the record holder for several years, until Sharir, Smorodinsky, and Tardos [33] obtained the best bound in terms of n known to date: $O(n^{5/2})$. By applying a technique of Agarwal *et al.* [1], this result can be converted to a k-sensitive upper bound of $O(nk^{3/2})$. The current best lower bound is $nk2^{\Omega(\sqrt{\log k})}$, which is lifted from Tóth's 2-d lower bound [36].

The bichromatic version of the k-set problem can be defined in the same way. The dual problem now corresponds to k-levels in arrangements of (closed) halfspaces. We obtain the first nontrivial upper bound in 3-d:

$$O(nk^{3/2} + n^{0.5034}k^{2.4932} + k^3).$$

For all $k = o(n)$, this bound is strictly better than Clarkson and Shor's ($\leq k$)-level bound $O(nk^2)$. For $k \ll \sqrt{n}$, the first term dominates and our bound matches the current best upper bound for the k-set/k-level problem by Sharir, Smorodinsky, and Tardos. On the other hand, for $k \gg n^{0.994}$, the last term dominates and our bound is optimal, because $\Omega(k^3)$ is an obvious lower bound.

Four dimensions. The higher-dimensional analog of the original k-set problem appears to be even more challenging. A slightly nontrivial bound of $O(n^{d-\alpha_d})$, with a rather small $\alpha_d = 1/(4d-3)^d$, can be obtained by combining Bárány *et al.*'s proof with Živaljević and Vrećica's multicolored Tverberg theorem [38] (whose only known proof currently requires tools from algebraic topology). Agarwal *et al.*'s technique [1] can turn this into a bound depending on k: $O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - \alpha_d})$. Just last year, Matoušek *et al.* [25] obtained an $O(n^{4-2/45})$ upper bound for the 4-d case, which becomes $O(n^2 k^{2-2/45})$.

One advantage in studying the k-level problem for arbitrary halfspaces is that we can directly reduce the problem in $\mathbb{R}^d$ into n instances of the problem in $\mathbb{R}^{d-1}$, because within each of n bounding hyperplanes, we get an subarrangement of halfspaces in $d-1$ dimensions. (This reduction would not work for k-levels of hyperplanes, i.e., the original k-set problem.) Automatically, our 3-d result implies the following bound for k-levels in arrangements of halfspaces (and thus bichromatic k-sets) in 4-d:

$$O(n^2 k^{3/2} + n^{1.5034} k^{2.4932} + nk^3).$$

Remarkably, this improves the best previous bound ($O(n^2 k^{2-2/45})$) for the original 4-d k-set problem by Matoušek *et al.* for quite a wide range of values of k, namely, when $k \ll n^{0.923}$. Furthermore, for $k \ll \sqrt{n}$, the first term dominates and our bound is particularly strong. Although we get no improvement for the most popular case $k = \Theta(n)$, any new result for the original k-set problem is significant, in the author's opinion.

Further extensions. A series of papers have addressed the k-level problem for arrangements of x-monotone curves and curve segments in two dimensions [35, 6, 8]. The author [8] has obtained subquadratic bounds (of about $O(nk^{1-1/2s})$) for any family of such curves and curve segments that pairwise intersect at most a constant s number of times; when specialized to the original case of lines ($s = 1$), this incidentally leads to a new proof of Erdős *et al.*'s old bound $O(n\sqrt{k})$. Improvements are known for certain curve families. For instance, we have $O(nk^{1/2} \log k)$ for pseudo-parabolas ($s = 2$) [8, 10]. These improvements are obtained by further exploiting results for a related problem of *cutting* curves into pseudo-segments [35, 6, 2, 23]. In a recent paper, the author [9] obtained the first nontrivial results for the k-level in certain arrangements of nonlinear, xy-monotone surfaces in three dimensions; the bound is $O(n^{2+\varepsilon} k^{1.9966})$ for pseudo-spherical patches where each triple intersects at most 2 times.

In this paper, we also derive new results for nonlinear objects. For any arrangement of n circular disks in 2-d (where the boundary of each pair intersects at most $s = 2$ times), we show that the number of vertices at level k is

$$O(nk^{1/2+\varepsilon} + k^2),$$

where the level of a point is the number of disks not containing the point. Note that levels in arrangements of circles have been considered as far back as in Tamaki and Tokuyama's paper [35], but there the level of a point is defined as the number of semicircular arcs below the point. Our definition is arguably more natural. Note that for $k \ll n^{2/3}$, the first term dominates and our bound almost matches the current upper bound for pseudo-parabolas. For $k \gg n^{2/3+\varepsilon}$, the second term dominates and our bound is optimal.

For the k-level in an arrangement of n spherical balls in 3-d, our bound is $O(n^{2+\varepsilon} k^{1.9966} + k^3)$. Again, for small k, this matches the known result for xy-monotone spherical patches, and for large k, this is optimal.

Techniques. As the first terms of our bounds coincide with known bounds for the original k-set/k-level problem, the reader might guess that our result for k-levels of halfspaces in 2-d and 3-d are obtained by modifying the proofs by Dey [14] and Sharir, Smorodinsky, and Tardos [33]. Indeed, this is the case, although new ideas are required in the modifications (especially in the 3-d case). The second and third terms come from adapting an entirely different known proof, namely, the inequalities of the author [8, 9] originally used to address levels in arrangements of curves and surfaces. In addition, another known technique by Agarwal *et al.* [1] for obtaining k-sensitive results plays a critical role.

Although it might appear that we are simply borrowing from existing techniques, determining the best way to apply these techniques, and realizing that they can actually be combined, are far from obvious, and represent the main contributions of this paper. We hope that the paper would add to better understanding into the relative strengths of the different known approaches to the original k-set problem.

2 Two Dimensions

2.1 Dey's approach

We begin by exploring Dey's approach [14] to solve the bichromatic k-set problem in 2-d. Dey actually gave two presentations of the proof: a dual version (from the conference paper) and a primal version (from the journal paper). We follow the dual presentation.

Let H be the given set of n halfplanes in $\mathbb{R}^2$. For the purposes of subsequent improvements, it helps to study a slightly generalized problem of bounding the combined complexity of $O(i)$ consecutive levels (where originally Dey obtained the bound $O(nk^{1/3}i^{2/3})$). Let Δ be a region (polygon) of constant complexity. Let T be the set of all vertices in the arrangement at levels between $k-i$ and $k+i$ inside Δ. Let $t = |T|$.

Dey's approach proceeds in two steps: 1. we bound t in terms of an intermediate quantity X, and 2. we bound X in terms of n. Here we define X to be the number of *vertex-vertex tangents*, which are pairs of the form (v_1, v_2) where $v_1 = \partial h_1 \cap \partial h_2 \in T$ and $v_2 = \partial h_3 \cap \partial h_4 \in T$ $(h_1, \ldots, h_4 \in H)$ such that $h_1 \cap h_2$ lies entirely on one side of the line through $v_1 v_2$, and $h_3 \cap h_4$ also lies entirely on one side of this line. Call (v_1, v_2) an *external* (resp. *internal*) vertex-vertex tangent if furthermore $h_1 \cap h_2$ and $h_3 \cap h_4$ lie on the same (resp. different) side of the line.

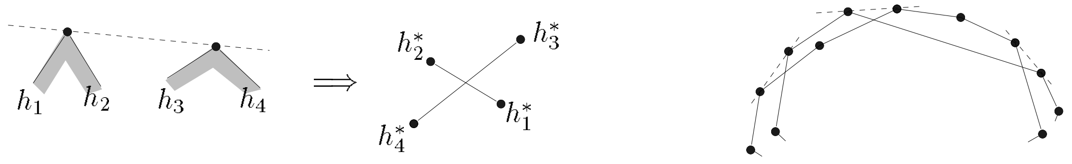
For step 1, Dey observed that an upper bound on t directly follows from the well-known Crossing Lemma [3, 20, 24, 28] for geometric graphs, which can be stated[1] as follows: any graph drawn in the plane with n nodes, m arcs, and X crossings must satisfy $m = O(n + n^{2/3}X^{1/3})$. To see the connection, we map each halfplane h with bounding line $y = ax + b$ to its dual point $h^* = (a, -b)$, and draw the resulting n points as nodes in our graph. The locus of tangent lines to a vertex $v = h_1 \cap h_2$ dualizes to either the line segment $\overline{h_1^* h_2^*}$ (see Figure 2(a)), or its complement inside the line $\overleftrightarrow{h_1^* h_2^*}$, which consists of two rays (for the original problem with only lower halfplanes, the latter case does not occur). If we draw such $O(t)$ line segments/rays for all vertices $v \in T$, a vertex-vertex tangent dualizes precisely to a crossing in the drawing. We can add two extra nodes at $(-\infty, 0)$ and $(\infty, 0)$ and bend the rays to terminate at these two points. The Crossing Lemma immediately implies:

Lemma 2.1 $t = O(n + n^{2/3}X^{1/3})$.

For step 2, Dey used a clever argument involving the "convex/concave chains" decomposition (first introduced in [1]) to prove that $X = O(n^2)$ in the original case of lower halfplanes with $i = 1$. This immediately leads to his $O(n^{4/3})$ result.

To solve the more general problem for arbitrary halfplanes, we need to change the way the chains are constructed; in particular, the number of chains required is now larger. We introduce a parameter M, defined as the number of local x-minima in T (i.e., vertices of the form $v = \partial h_1 \cap \partial h_2 \in T$ where $h_1 \cap h_2$ is entirely to the right of the vertical line through v).

[1] The Crossing Lemma is more commonly stated in the following equivalent form: if $m \geq cn$ for a sufficiently large constant c, then $X = \Omega(m^3/n^2)$.



(a) (b)

Figure 2: (a) An (external) vertex-vertex tangent and its dual. (b) Proof of Lemma 2.2.

Our new step 2 is described below:

Lemma 2.2 $X = O(n^2 i^2 + M^2)$.

Proof: For $j \in [k-i, k+i]$, let S_j be the set of all local x-minima at level j inside Δ, together with all points at levels at most j on the boundary $\partial\Delta$. For each $j \in [k-i, k+i]$ and $v \in S_j$, we create at most two x-monotone chains starting at v, initially along one of the lines incident to v. As we sweep from left to right, whenever we hit a vertex at level j, we make the chain turn at this vertex to follow this vertex's other incident line. As soon as we hit a local x-maximum at level j or touch $\partial\Delta$, we stop the chain.

Note that the level of a point on the chain is always kept at most j. The chain is either concave or convex: if the initial line bounds a lower (resp. upper) halfplane, the entire chain goes through boundaries of only lower (resp. upper) halfplanes. Each vertex at level j is a vertex of one of these chains, since we can reverse the sweep until we return to a point of S_j. Conversely, a vertex can be the vertex of at most two chains. A point can lie on at most $O(i)$ chains ($O(1)$ chains per choice of j). In particular, the number of chains that touch $\partial\Delta$ is $O(ni)$. The total number of chains is thus $C = O(M + ni)$.

The key observation is that the number of common (external or internal) tangents between two convex/concave chains is at most the number of intersection points between the two chains plus $O(1)$ (see Figure 2(b)). Thus, summing up over all $O(C^2)$ pairs of chains, we get that the total number of common tangents is $O(n^2 i^2 + C^2)$, since each of the $O(n^2)$ intersection points is counted at most $O(i^2)$ times. □

Combining the two preceding lemmas, we get:

Lemma 2.3 *For an arrangement of n halfplanes in $\mathbb{R}^2$, a region Δ of constant complexity, and $1 \le i \le k \le n$, the number of vertices at levels between $k-i$ and $k+i$ inside Δ is $f_i(n, M) = O(n^{4/3} i^{2/3} + n^{2/3} M^{2/3})$, where M is the number of local x-minima among such vertices.*

It is known that the number of local x-minima at levels between $k-i$ and $k+i$ (in fact, the number of local x-minima at all levels at most $O(k)$) is $M = O(k^2)$ [27]. Thus, Lemma 2.3 implies an $O(n^{4/3} + n^{2/3} k^{4/3})$ bound for the k-level of halfplanes. It is not difficult to reduce the first term to $O(nk^{1/3})$ by simple modifications of the above proof. However, we present an even better way to obtain k-sensitive results, which can improve the second term as well. The technique involves a geometric divide-and-conquer that builds upon a technique of Agarwal *et al.* [1] previously designed to obtain k-sensitive results for the original k-level problem. Below, the importance of having the analysis depend on the parameter M is revealed.

Theorem 2.4 *For an arrangement of n halfplanes in $\mathbb{R}^2$ and $1 \le i \le k \le n$, the number of vertices at levels between $k-i$ and $k+i$ is $O(nk^{1/3}i^{2/3} + n^{1/3}k^{5/3})$.*

Proof: As we have mentioned, the number of local x-minima at levels between $k-i$ and $k+i$ is $O(k^2)$. It is also known [13, 30] that the $(\le 3k)$-level of the lower halfplanes and the $(\le 3k)$-level of the upper halfplanes have $O(nk)$ complexity. Let $k' \in (2k, 3k]$ be an index that minimizes the sum of the complexities of the k'-level L^- of the lower halfplanes and the k'-level L^+ of the upper halfplanes. It follows that L^- and L^+ have $O(nk/k) = O(n)$ number of vertices. Divide the plane into $O(n/k)$ vertical slabs by drawing a vertical line at every k-th vertex of L^- and L^+ and drawing a vertical line at every $\lceil k^2/(n/k)\rceil$-th local x-minimum. Then each slab contains $O(k)$ vertices of L^- and L^+ and $O(k^3/n)$ local x-minima.

For each slab Δ, let H_Δ be the subset of all halfplanes that define vertices of L^- and L^+ inside the slab Δ, together with the k' lowest lower halfplanes and the k' highest upper halfplanes at the left wall of Δ. Then $|H_\Delta| = O(k)$. Observe that the j-level of H inside Δ coincides with the j-level of H_Δ inside Δ for every $j \le k+i < k'$. Applying Lemma 2.3 to H_Δ and summing over all slabs Δ yield the upper bound

$$O\left(\frac{n}{k}\cdot f_i(O(k), O(k^3/n))\right) \;=\; O\left(\frac{n}{k}\cdot\left[k^{4/3}i^{2/3} + k^{2/3}\left(\frac{k^3}{n}\right)^{2/3}\right]\right).$$

□

Thus, the complexity of the k-level is $O(nk^{1/3} + n^{1/3}k^{5/3})$, which matches Dey's bound for small $k \ll n^{1/2}$. The bound is worse for large k, but is always better or as good as the $(\le k)$-level bound $O(nk)$.

2.2 An inequality-based approach

We next explore a different approach proposed by the author [8]. For the original k-level problem, this approach led to a bound $(O(n\sqrt{k}))$ which is worse than Dey's. However, the idea turns out to have new implications for our bichromatic problem.

The approach is based on a simple inequality relating the combined size of levels $k-i, \ldots, k+i$ to the sizes of levels $k-i$ and $k+i$. To solve our bichromatic problem, we need to generalize this inequality from arrangements of lines to arrangements of halfplanes. Fortunately, another paper by the author [9] provides precisely such a generalization (the objective of that paper was actually to study k-levels in 3-d).

Lemma 2.5 *For an arrangement of n halfplanes in $\mathbb{R}^2$, let t_i^{bi} be the number of vertices at levels between $k-i$ and $k+i$ that are local x-extrema (i.e., local x-minima and x-maxima), and let t_i^{mo} be the number of vertices at levels between $k-i$ and $k+i$ that are not local x-extrema. Let $\Delta t_i^{\mathrm{mo}} = t_{i+1}^{\mathrm{mo}} - t_i^{\mathrm{mo}}$ and $\Delta t_i^{\mathrm{bi}} = t_{i+1}^{\mathrm{bi}} - t_i^{\mathrm{bi}}$. Then*

$$t_i^{\mathrm{mo}} \;\preceq\; c_1 i(\Delta t_i^{\mathrm{mo}} + \Delta t_i^{\mathrm{bi}}) + c_2 t_i^{\mathrm{bi}} + O(n_i i),$$

where n_i denotes the number of vertices at levels between $k-i$ and $k+i$ at $x = \pm\infty$, and $f \preceq g$ means $f \le (1 + O(1/i))g$.

In [9], lines bounding upper/lower halfplanes are colored red/blue, and a local x-extremum is referred to as a "bichromatic" vertex as it is defined by a pair of lines of different colors; the remaining vertices are thus "monochromatic". The paper [9] provided two proofs of the above lemma. A simpler proof gives the constants $c_1 = c + 2$ and $c_2 = 2 + 2/c$ for any fixed $c > 0$. A more complicated proof improves c_2 to a value strictly smaller than 2 at the expense of increasing c_1; for example, $c_1 \approx 15.83$ and $c_2 \approx 1.8343$ (from the latest improved version of the paper [9]). In our application here, the simpler proof suffices, since c_2 is less important than c_1.

Theorem 2.6 *For an arrangement of n halfplanes in $\mathbb{R}^2$, the number of vertices at level k is $O(nk^{1/2+\varepsilon} + k^2)$.*

Proof: Note that $t_i^{\mathrm{bi}}, \Delta t_i^{\mathrm{bi}} = O(k^2)$ by the known bound on local minima. Let $t_i = t_i^{\mathrm{mo}} + t_i^{\mathrm{bi}}$ and $\Delta t_i = \Delta t_i^{\mathrm{mo}} + \Delta t_i^{\mathrm{bi}}$. By Lemma 2.5 with $c_1 = 2 + c$ and $c_2 = 2 + 2/c$ for an arbitrarily small constant $c > 0$,

$$t_i \;\preceq\; (2+\varepsilon)i\Delta t_i + O(ni + k^2), \qquad \text{i.e.,} \qquad t_i \;\preceq\; \frac{(2+\varepsilon)i}{(2+\varepsilon)i+1} t_{i+1} + O\left(n + \frac{k^2}{i}\right).$$

By expanding the recurrence for j iterations and noting in general that $\prod_{\ell=1}^{i} \frac{a\ell}{a\ell+1} = O(1/i^{1/a})$ for any $a > 0$,

$$t_1 \;=\; O\left(\frac{t_j}{j^{1/2-\varepsilon}} + \sum_{i=1}^{j}\left(n + \frac{k^2}{i}\right)\frac{1}{i^{1/2-\varepsilon}}\right) \;=\; O\left(\frac{t_j}{j^{1/2-\varepsilon}} + nj^{1/2+\varepsilon} + k^2\right) \tag{1}$$

(after readjusting ε). We choose $j = k$ and use the base case $t_k = O(nk)$ [13, 30]. $\square$

Although the first term is worse than that of Theorem 2.4, interestingly we get an optimal $O(k^2)$ bound for large $k \gg n^{2/3+\varepsilon}$. Dey's concave-chains approach does not seem capable of getting such an $O(k^2)$ bound.

2.3 Combining the two approaches

The preceding two subsections illustrate well the merits of both proof techniques: neither one is superior to the other for all values of k. We can actually combine the two to get a better result for intermediate values of k:

Theorem 2.7 *For an arrangement of n halfplanes in $\mathbb{R}^2$, the number of vertices at level k is $O(nk^{1/3} + n^{5/6-\varepsilon}k^{2/3+2\varepsilon} + k^2)$.*

Proof: We follow the proof of Theorem 2.6 but use a different base case, namely the bound provided by Theorem 2.4. For any $1 \le j \le k$, (1) becomes

$$t_1 \;=\; O\left(\frac{nk^{1/3}j^{2/3} + n^{1/3}k^{5/3}}{j^{1/2-\varepsilon}} + nj^{1/2+\varepsilon} + k^2\right).$$

For $k > \sqrt{n}$, setting $j = k^2/n$ yields $t_1 = O(n^{5/6-\varepsilon}k^{2/3+2\varepsilon} + n^{1/2-\varepsilon}k^{1+2\varepsilon} + k^2)$, where the second term is absorbed by the first. For $k \le \sqrt{n}$, the bound in Theorem 2.4 already gives $O(nk^{1/3})$. $\square$

Remark: By using a nonconstant value for ε, we can lower (and even eliminate) the extra $(k^2/n)^\varepsilon$ factor in the second term, but the third term would increase slightly.

3 Three Dimensions

3.1 The Sharir–Smorodinsky–Tardos approach

We now tackle the bichromatic k-set problem in 3-d. In this subsection, we adapt Sharir, Smorodinksy, and Tardos's approach [33]. Their proof was originally described in primal space but we find it more convenient to work in the dual.

Let H be the given set of n halfspaces in $\mathbb{R}^3$. As in the previous section, we consider a slightly more general problem, of bounding the size t of the set T of all vertices at level between $k-i$ and $k+i$ inside a given region (polyhedron) Δ of constant complexity.

Sharir *et al.*'s approach proceeds in two steps, like Dey's: 1. we bound t in terms of an intermediate quantity X, and 2. we bound X in terms of n. Here in the 3-d (dual) setting, we[2] define X to be the number of *line-vertex tangents*, which are pairs of the form (ℓ, v) where $\ell = \partial h_1 \cap \partial h_2$ and $v = \partial h_3 \cap \partial h_4 \cap \partial h_5 \in T$ $(h_1, \ldots, h_5 \in H)$ such that $h_3 \cap h_4 \cap h_5$ lies entirely on one side of the plane through ℓ and v. If (ℓ, v) forms a line-vertex tangent, we say that v is a *local ℓ-extremum.*

The 3-d k-level problem is more challenging than the 2-d problem, because of the lack of a suitable generalization of convex/concave chain decomposition to convex/concave surfaces. Sharir *et al.*'s idea, however, is to work within each of n 2-d subarrangements separately; inside each plane, we can build the convex/concave chains as before, and by summing, we may be able to deduce consequences in 3-d. Although the proof uses convex/concave chains, it is quite different from Dey's proof, in that step 1 is the innovative part where chains are used, while step 2 simply follows from a known fact (Lovász's Lemma).

We begin with step 1. Here, we can essentially borrow Sharir *et al.*'s proof without significant changes, except that we recast the proof completely in dual space and incorporate the parameter i:

Lemma 3.1 $t = O(n^2 i + \sqrt{nX})$.

Proof: Fix a halfspace $h \in H$. Consider the subarrangement $\mathcal{A}_h$ of the arrangement of H restricted within the plane ∂h; this is a 2-d arrangement of halfplanes. Follow the construction in the proof of Lemma 2.2 to obtain a collection of convex and concave chains inside the region $\Delta \cap \partial h$ within this 2-d subarrangement $\mathcal{A}_h$. Recall that a point can lie on at most $O(i)$ chains. Let t_h be the number of vertices of T within $\mathcal{A}_h$. Let C_h be the number of chains, which is asymptotically at least the number of local x-extrema in T with respect to the 2-d subarrangement $\mathcal{A}_h$.

Let X_h be the number of external vertex-vertex tangents in T within $\mathcal{A}_h$. We claim that

$$X_h = \Omega(C_h - O(ni))^2 - O(t_h ni).$$

To see this, without loss of generality assume that there are more concave chains than convex chains. The number of pairs of concave chains is $\Omega(C_h^2)$, or $\Omega(C_h - O(ni))^2$ if we exclude the $O(ni)$ chains that touch $\partial\Delta \cap \partial h$. If two concave chains have disjoint x-spans and do not touch $\partial\Delta \cap \partial h$, then they must have at least one common external tangent (see Figure 3(a)). On the other hand, if two chains have overlapping x-spans, then one chain γ must intersect the vertical line through some vertex v of the other chain (see Figure 3(b)); there are at most $O(ni)$ such chains γ for each of the $O(t_h)$ choices of v.

Next, we sum the above inequality over all n halfspaces h. Note that $\sum_h t_h = 3t$ and $\sum_h C_h = \Omega(t)$: each vertex $v = \partial h_1 \cap \partial h_2 \cap \partial h_3$ lies in three 2-d subarrangements $\mathcal{A}_{h_1}$, $\mathcal{A}_{h_2}$, and $\mathcal{A}_{h_3}$; at least

[2]Technically, our definition of X here is a slight variation of Sharir *et al.*'s.

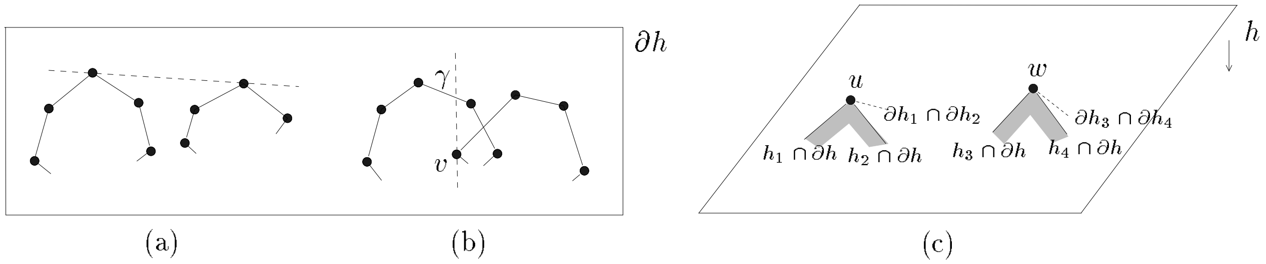


Figure 3: Proof of Lemma 3.1.

one of three faces of $h_1 \cap h_2 \cap h_3$, say the one on ∂h_1, lies entirely on one side of the vertical plane through v, so that v is a local x-extremum with respect to $\mathcal{A}_{h_1}$.

Moreover, $\sum_h X_h = O(X)$: if $(u = \partial h_1 \cap \partial h_2 \cap \partial h,\ w = \partial h_3 \cap \partial h_4 \cap \partial h)$ is an external vertex-vertex tangent in the 2-d arrangement $\mathcal{A}_h$ (see Figure 3(c)), then we can rotate a plane π in 3-d, starting at ∂h, around the line $\overleftrightarrow{uw}$, while remaining tangent to $h_1 \cap h_2 \cap h$ and $h_3 \cap h_4 \cap h$; if π hits the edge $\partial h_1 \cap \partial h_2 \cap h$ first before $\partial h_3 \cap \partial h_4 \cap h$, then $(\partial h_1 \cap \partial h_2, w)$ is a line-vertex tangent, otherwise $(\partial h_3 \cap \partial h_4, u)$ is.

Thus,

$$X = \frac{\Omega(t - O(n^2 i))^2}{n} - O(tni)$$

(since $\sum_{i=1}^n a_i^2 \geq \left(\sum_{i=1}^n a_i\right)^2 / n$), and the conclusion follows (the tni term is dominated by t^2/n if $t = \omega(n^2 i)$). $\square$

For step 2, in the original case of lower halfspaces with $i = 1$, we can apply a 3-d version of Lovász's Lemma [5, 1], which states[3] that any fixed line ℓ can define at most $O(n^2)$ local extremal vertices at level k. By summing over all $O(n^2)$ lines ℓ, see that the total number X of ℓ-local extrema is at most $O(n^4)$, which immediately leads to Sharir *et al.*'s $O(n^{5/2})$ upper bound. In fact, the bound on the number of ℓ-local extrema can be reduced to $O(k^2)$ [1, 12, 26], so X can be reduced to $O(n^2k^2)$.

Unfortunately, for the more general case of arbitrary halfspaces, the number of ℓ-local extrema is $O(k^3)$ [27], so we only get $X = O(n^2k^3)$. Plugging into Lemma 3.1 gives $O(n^2 + n^{3/2}k^{3/2})$ for $i = 1$. Although the first term can be refined to become sensitive to k, the second term always exceeds the known $(\leq k)$-level bound $O(nk^2)$. So, the above proof alone is not sufficient to yield any new result for our problem.

Our idea is to bound X in terms of another quantity M, like in Section 2.1, which enables us to use geometric divide-and-conquer later. We define M to be the number of local ℓ_0-extrema in T, where ℓ_0 is a fixed line (the choice of ℓ_0 turns out to be inconsequential).

Our step 2 uses a new argument. Interestingly, we use concave/convex chains for this part as well:

Lemma 3.2 $X = O(n^4 i + n^2 M)$.

Proof: Define $\mathcal{A}_h$, C_h, and chains in the same manner as in the proof of Lemma 3.1.

[3]Lovász's Lemma is more commonly stated in the dual form: in 3-d, any fixed line can intersect at most $O(n^2)$ number of k-*facets* (triangles formed by 3 data points whose incident plane has exactly k points underneath).

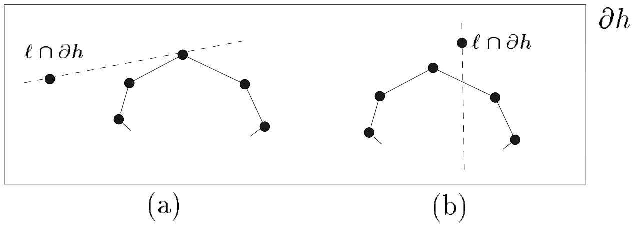


Figure 4: Proof of Lemma 3.2.

Fix a line ℓ. Let X_ℓ be the number of local ℓ-extrema in T. Let $X_{\ell h}$ be the number of local ℓ-extrema in T with respect to the 2-d subarrangement $\mathcal{A}_h$, i.e., vertices of the form $v = \partial h_1 \cap \partial h_2 \cap \partial h$ such that $h_1 \cap h_2 \cap \partial h$ lies entirely on one side of the line through $\ell \cap \partial h$ and v within ∂h. In such a case, call v an *ℓ-maximum* (resp. *ℓ-minimum*) if $h_1 \cap h_2 \cap \partial h$ is below (resp. above) the line within ∂h.

We claim that

$$X_{\ell h} \;=\; C_h \pm O(ni).$$

To see this, observe that each concave (resp. convex) chain has at most two local ℓ-maxima (resp. ℓ-minima). If the chain is completely to the left or to the right of $\ell \cap \partial h$ and does not touch $\partial\Delta \cap \partial h$, then the chain has precisely one local ℓ-maximum (resp. ℓ-minimum) (see Figure 4(a)). On the other hand, there are at most $O(ni)$ chains that intersect the vertical line through $\ell \cap \partial h$ or touch $\partial\Delta \cap \partial h$ (see Figure 4(b)).

Next, note the identity $\sum_h X_{\ell h} = 3X_\ell + (t - X_\ell)$: for each $v = \partial h_1 \cap \partial h_2 \cap \partial h_3$, if v is a local ℓ-extremum in 3-d, then it is a local ℓ-extremum with respect to three 2-d subarrangements; on the other hand, if v is not a local ℓ-extremum, then precisely one face of $h_1 \cap h_2 \cap h_3$ lies entirely on one side of the plane through ℓ and v, so that v is a local ℓ-extremum with respect to precisely one of the three 2-d subarrangements.

Summing the above equation over all n halfspaces h gives $\sum_h C_h \pm O(n^2 i) = 3X_\ell + (t - X_\ell)$. Since the left-hand side is independent of ℓ, we see that X_ℓ must be identical for all ℓ up to an $O(n^2 i)$ error term. In particular, $X_\ell = M \pm O(n^2 i)$.

Thus,

$$X \;=\; \sum_\ell X_\ell \;=\; O(n^2 \cdot (M + n^2 i)).$$

□

Combining the two preceding lemmas, we get:

Lemma 3.3 *For an arrangement of n halfspaces in $\mathbb{R}^3$, a region Δ of constant complexity, and $1 \le i \le k \le n$, the number of vertices at levels between $k - i$ and $k + i$ inside Δ is $f_i(n, M) = O(n^{5/2} i^{1/2} + n^{3/2} M^{1/2})$, where M is the number of local ℓ_0-extrema among such vertices.*

We are finally ready to apply divide-and-conquer. The simple strategy from the proof of Theorem 2.4 would not work in 3-d, and we need a random sampling strategy which is based on (but is slightly more involved than) another technique by Agarwal *et al.* [1].

Theorem 3.4 *For an arrangement of n halfspaces in $\mathbb{R}^3$ and $1 \le i \le k \le n$, the number t of vertices at levels between $k - i$ and $k + i$ is $O(nk^{3/2} i^{1/2} + n^{1/2} k^{5/2})$.*

Proof: The total number of local ℓ_0-extrema at levels between $k-i$ and $k+i$ is $O(k^3)$ by known bounds [27]. Let $\widetilde{H}$ be a set containing the given halfspaces, each duplicated $\mu := \lceil k^3/n \rceil$ times, together with the $O(k^3)$ local ℓ_0-extrema. Thus $\widetilde{H}$ contains $O(\mu n)$ halfspaces and $O(\mu n)$ points.

Take a random sample R of $\widetilde{H}$, where each element (halfspace or point) is selected with probability $1/(3\mu k)$. Thus, $E[|R|] = O(n/k)$. Let $\mathcal{P}(R)$ denote the intersection of the halfspaces in R. Consider the canonical triangulation [11] of $\mathcal{P}(R)$. For each cell in the triangulation, divide the cell into subcells by adding vertical walls at every point of R in the cell. Let $\mathcal{T}(R)$ be the resulting triangulation. The number of subcells is still $O(|R|)$. For each subcell Δ of a cell Δ', let $\widetilde{H}_\Delta$ contain all halfspaces of $\widetilde{H}$ whose bounding planes intersect Δ', together with all points of $\widetilde{H}$ that lie inside Δ. According to Clarkson and Shor's analysis [13], we have $E\left[\sum_{\Delta\in\mathcal{T}(R)} |\widetilde{H}_\Delta|^a\right] = O((n/k)\cdot(\mu k)^a)$ for any $a \geq 0$.

Let Z be the number of vertices at levels between $k-i$ and $k+i$ that lie inside $\mathcal{P}(R)$. For each subcell Δ, observe that any $2i$ consecutive levels of H inside Δ coincide with some $2i$ consecutive levels of the halfspaces in $\widetilde{H}_\Delta$ inside Δ. The number of distinct halfspaces in $\widetilde{H}_\Delta$ is at most $|\widetilde{H}_\Delta|/\mu$, and the number of local ℓ_0-extrema inside Δ is at most $|\widetilde{H}_\Delta|$. Thus, applying Lemma 3.3 to the halfspaces in $\widetilde{H}_\Delta$ and summing over all subcells Δ yield

$$\begin{aligned} E[Z] &\leq E\left[\sum_{\Delta\in\mathcal{T}(R)} f_i(|\widetilde{H}_\Delta|/\mu, |\widetilde{H}_\Delta|)\right] = E\left[\sum_{\Delta\in\mathcal{T}(R)} O\left((|\widetilde{H}_\Delta|/\mu)^{5/2} i^{1/2} + (|\widetilde{H}_\Delta|/\mu)^{3/2}|\widetilde{H}_\Delta|^{1/2}\right)\right] \\ &= O\left(\frac{n}{k}\cdot\left[k^{5/2}i^{1/2} + k^{3/2}\left(\frac{k^4}{n}\right)^{1/2}\right]\right). \end{aligned}$$

On the other hand, for a fixed vertex v at level at most $2k$, the probability that v lies outside $\mathcal{P}(R)$ is at most $2\mu k \cdot 1/(3\mu k) = 2/3$. Thus, $E[Z] = \Omega(t)$. The conclusion follows. □

3.2 The inequality-based approach

We now give a different solution that uses the approach proposed by the author [8, 9]. For the original k-level problem, this approach led to a bound ($O(n^{2.9966})$) much worse than Sharir, Smorodinksy, and Tardos'. However, for the bichromatic problem, we actually get better (in fact optimal $O(k^3)$) bounds for large k.

Theorem 3.5 *For an arrangement of n halfspaces in $\mathbb{R}^3$, the number of vertices at level k is $O(nk^{1.9966} + k^3)$.*

Proof: We apply Lemma 2.5 to each 2-d subarrangement $\mathcal{A}_h$ of halfplanes and sum over all halfspaces h. We use the more sophisicated version of the lemma, with $c_1 \approx 15.83$ and $c_2 \approx 1.8343$.

For each vertex $v = \partial h_1 \cap \partial h_2 \cap \partial h_3$, if v is a local x-extremum in 3-d, then it is a local x-extremum with respect to three 2-d subarrangements; there are $O(k^3)$ local x-extrema in 3-d by known results [27]. Otherwise, precisely one face of $h_1 \cap h_2 \cap h_3$ lies entirely on one side of the vertical plane through v, so that v is a local ℓ-extremum with respect to one of the three 2-d subarrangements, and is not a local ℓ-extremum with respect to two of the three. At $x = \pm\infty$, the number of vertices at levels between $k-i$ and $k+i$ can be trivially bounded by $O(nk)$ (the 2-d ($\leq O(k)$)-level bound).

Letting t_i be the number of vertices at levels between $k-i$ and $k+i$ in the 3-d arrangement, we thus get:

$$2t_i \preceq 3c_1 i\Delta t_i + c_2 t_i + O(nki + k^3)$$

$$\implies t_i \preceq c_0 i\Delta t_i + O(nki + k^3), \qquad \text{i.e.,} \qquad t_i \preceq \frac{c_0 i}{c_0 i + 1} t_{i+1} + O\left(nk + \frac{k^3}{i}\right),$$

where $c_0 = 3c_1/(2 - c_2) < 287$. Expanding the recurrence for j iterations gives

$$t_1 = O\left(\frac{t_j}{j^{1/c_0}} + \sum_{i=1}^{j}\left(nk + \frac{k^3}{i}\right)\frac{1}{i^{1/c_0}}\right) = O\left(\frac{t_j}{j^{1/c_0}} + nkj^{1-1/c_0} + k^3\right). \tag{2}$$

We choose $j = k$ and use the base case $t_k = O(nk^2)$ [13]. □

3.3 Combining the two approaches

Again we can get a better result for intermediate values of k by combining the two approaches:

Theorem 3.6 *For an arrangement of n halfspaces in $\mathbb{R}^3$, the number of vertices at level k is $O(nk^{3/2} + n^{0.5034}k^{2.4932} + k^3)$.*

Proof: We follow the proof of Theorem 3.5 but use a different base case, namely the bound provided by Theorem 3.4. For any $1 \le j \le k$, (2) implies

$$t_1 = O\left(\frac{nk^{3/2}j^{1/2} + n^{1/2}k^{5/2}}{j^{1/c_0}} + nkj^{1-1/c_0} + k^3\right),$$

with $c_0 < 287$. For $k > \sqrt{n}$, setting $j = k^2/n$ yields $t_1 = O(n^{1/2+1/c_0}k^{5/2-2/c_0} + n^{1/c_0}k^{3-2/c_0} + k^3)$, where the second term is absorbed by the first. For $k \le \sqrt{n}$, the bound in Theorem 3.4 already gives $O(nk^{3/2})$. □

4 Extensions

In this section, we discuss some extensions of our results to arrangements of nonlinear objects.

One of the most natural cases is 2-d disks. Although Dey's approach is still viable after applying known techniques for *cutting* pseudo-parabolas into (extendible) pseudo-segments [6, 23, 35] and using a larger number of chains, the inequality-based approach turns out to give a strictly better result in this case.

The paper [9] provided a generalization of Lemma 2.5 for curves. The version specialized for disks is stated below. The overhead term $O(n^{3/2}i^{1/2}\log(n/i))$ comes from the best known results on cutting pseudo-parabolas by Marcus and Tardos [23]. The constants c_1 and c_2 are the same as before.

Lemma 4.1 *For an arrangement of n (pseudo-)disks in $\mathbb{R}^2$, let t_i^{bi} be the number of vertices at levels between $k - i$ and $k + i$ that are local x-extrema, and let t_i^{mo} be the number of vertices at levels between $k - i$ and $k + i$ that are not local x-extrema. Let $\Delta t^{\mathrm{mo}} = t_{i+1}^{\mathrm{mo}} - t_i^{\mathrm{mo}}$ and $\Delta t^{\mathrm{bi}} = t_{i+1}^{\mathrm{bi}} - t_i^{\mathrm{bi}}$. Then*

$$t_i^{\mathrm{mo}} \preceq c_1 i(\Delta t_i^{\mathrm{mo}} + \Delta t_i^{\mathrm{bi}}) + c_2 t_i^{\mathrm{bi}} + O(n^{3/2}i^{1/2}\log(n/i)).$$

We use the same approach as in Section 2.2, in conjunction with Agarwal *et al.*'s technique:

Theorem 4.2 *For an arrangement of n disks (or convex pseudo-disks) in $\mathbb{R}^2$, the number of vertices at level k is $O(nk^{1/2+\varepsilon} + k^2)$.*

Proof: Let M be the number of local x-extrema between levels 0 and $2k$. We proceed as in the proof of Theorem 2.6. By Lemma 2.5 with $c_1 = 2 + c$ and $c_2 = 2 + 2/c$ for an arbitrarily small constant $c > 0$, we see that the number t_i of vertices at levels between $k - i$ and $k + i$ satisfies the recurrence

$$t_i \preceq (2+\varepsilon)i\Delta t_i + O(n^{3/2+\varepsilon'} i^{1/2-\varepsilon'} + M),$$

which solves to

$$t_1 = O\left(\frac{t_j}{j^{1/2-\varepsilon}} + \sum_{i=1}^{j}\left(\frac{n^{3/2+\varepsilon'}}{i^{1/2+\varepsilon'}} + \frac{M}{i}\right)\frac{1}{i^{1/2-\varepsilon}}\right) = O\left(\frac{t_j}{j^{1/2-\varepsilon}} + n^{3/2+\varepsilon'} j^{\varepsilon-\varepsilon'} + M\right)$$

(after readjusting ε and for a sufficiently small $\varepsilon' < \varepsilon$). Choosing $j = n$ and using the base case $t_n = O(n^2)$, we get the upper bound $f(n, M) := O(n^{3/2+\varepsilon} + M)$. Note that the bound still holds if we restrict ourselves to vertices inside a region Δ of constant complexity.

To make the first term k-sensitive, we next proceed as in the proof of Theorem 2.4. The divide-and-conquer strategy gives $O(n/k)$ slabs Δ, where within each slab Δ, only $O(k)$ disks are relevant. Let M_Δ be the number of local x-minima of levels between 0 and $2k$ inside Δ. We then get the upper bound

$$\sum_{\Delta} f(O(k), M_\Delta) = O\left(\left(\frac{n}{k}\right) k^{3/2+\varepsilon} + \sum_{\Delta} M_\Delta\right).$$

To conclude, observe that $\sum_\Delta M_\Delta$, the total number of local x-minima of levels between 0 and $2k$, is still $O(k^2)$, by the same proof for halfplanes/halfspaces [27]. (The main property used in this proof is that for any subset of objects, the number of local x-minima of the intersection of the objects is at most 1—this is true by convexity of disks.) □

For 3-d balls, Sharir, Smorodinsky, and Tardos' approach does not seem easily generalizable, but we can adapt the inequality-based approach of Section 3.2, again in conjunction with Agarwal *et al.*'s technique [1]:

Theorem 4.3 *For an arrangement of n balls in $\mathbb{R}^3$, the number of vertices at level k is $O(nk^{1.9966} + k^3)$.*

Proof: Within the upper/lower boundary of each disk, we get a 2-d subarrangement of pseudo-disks. We apply Lemma 2.5 to each 2-d subarrangement, and sum up. We need the more sophisicated version of the lemma, with $c_1 \approx 15.83$ and $c_2 \approx 1.8343$.

For each vertex v, we still have the following properties: If v is a local x-extremum in 3-d, then it is a local x-extremum in with respect to three 2-d subarrangements. Otherwise, v is a local ℓ-extremum with respect to one of the three 2-d subarrangements, and is not a local ℓ-extremum with respect to two of the three.

Letting M be the number of local x-extrema at levels between 0 and $2k$ in 3-d, t_i be the number of vertices at levels between $k - i$ and $k + i$ in the 3-d arrangement, we thus get:

$$2t_i \preceq 3c_1 i\Delta t_i + c_2 t_i + O(n^{5/2+\varepsilon} i^{1/2-\varepsilon} + k^3) \implies t_i \preceq c_0 i\Delta t_i + O(n^{5/2+\varepsilon} i^{1/2-\varepsilon} + M)$$

for $c_0 = 3c_1/(2-c_2) < 287$. This recurrence solves to

$$t_1 \;=\; O\left(\frac{t_j}{j^{1/c_0}} + \sum_{i=1}^{j}\left(\frac{n^{5/2+\varepsilon}}{i^{1/2+\varepsilon}} + \frac{M}{i}\right)\frac{1}{i^{1/c_0}}\right) \;=\; O\left(\frac{t_j}{j^{1/c_0}} + n^{5/2+\varepsilon} i^{1/2-1/c_0+\varepsilon} + M\right).$$

Choosing $j = n$ and using the base case $t_n = O(n^3)$, we get the upper bound $f(n, M) := O(n^{3-1/c_0} + M)$. Again the bound still holds if we restrict ourselves to vertices inside a region Δ.

To make the first term k-sensitive, we next proceed as in the proof of Theorem 3.4, with a slightly simplified divide-and-conquer strategy. We take a random sample R where each ball is chosen with probability $1/(3k)$. A decomposition of the intersection of the balls in R has complexity $O(|R|^{2+\varepsilon})$ [31]. In each cell Δ, let H_Δ be the subset of all balls whose boundary intersect Δ. According to Clarkson and Shor's analysis [13], we have $E\left[\sum_\Delta |H_\Delta|^a\right] = O((n/k)^{2+\varepsilon} k^a)$ for any $a \geq 0$. Let M_Δ be the number of local x-extrema at levels between 0 and $2k$ inside Δ. We then get the upper bound

$$\sum_\Delta f(|H_\Delta|, M_\Delta) \;=\; O\left(\left(\frac{n}{k}\right)^{2+\varepsilon} k^{3-1/c_0} \;+\; \sum_\Delta M_\Delta\right).$$

To conclude, observe that $\sum_\Delta M_\Delta$, the total number of local x-minima at levels between 0 and $2k$, is still $O(k^3)$, again by a known proof [27] (due to convexity of balls). $\square$

Remark: The same approach works for arrangements of other objects in 2-d and 3-d, if the 2-d subarrangements admit subquadratic cutting bounds and the objects are convex.

5 Further Remarks

As mentioned in the introduction, one of the notable consequences of our work is a new bound for the original k-set problem in 4-d for $k \ll n^{0.923}$. Although our results say nothing about the most important case $k = n/2$, they do show a phenomenon not observed before—that it may be potentially easier to get new results for k-sets for *small* k when the dimension is at least 4. To put things into perspective, recall that as Agarwal *et al.* [1] showed, an upper bound for the case $k = \Theta(n)$ is sufficient to yield an upper bound for all k; specifically, an upper bound of $f(n)$ for the former implies an upper bound of $O((n/k)^{\lfloor d/2 \rfloor} f(k))$ for the latter. For $d \leq 3$, the converse is easily seen to be true: a lower bound $f(n)$ for the former translates to a lower bound of $\Omega((n/k)f(k))$ for all k, by simply making n/k separate affine copies of a bad point set of size k. However, such an argument for the converse apparently fails for $d \geq 4$.

For small k, our k-set bound in 4-d is $O(n^2 k^{3/2})$, and without new breakthroughs on the current 3-d bound ($O(nk^{3/2})$), this is basically as good as one can wish for. But is it really? A lower bound of $f(n, k)$ in $d-1$ dimensions easily lifts to a lower bound of $\Omega(k \cdot f(n, k))$ in d dimensions, but not necessarily $\Omega(n \cdot f(n, k))$. At the moment, it is conceivable that one could prove, say, an $O(n^2 k^{1+\varepsilon} + k^{O(1)})$ upper bound for the 4-d k-set problem without changing current knowledge on the 2-d and 3-d k-set problems. In particular, naively lifting Tóth's 2-d lower bound [36] does not give a meaningful lower bound for k-sets when k is small and $d \geq 4$.

Actually, it might be possible to prove new bounds for k-sets in higher fixed dimensions for small k by using the inequality-based approach, as the following theorem indicates. (As mentioned in the introduction, the current k-set bound for higher dimensions is very weak.) The catch, however, is that the current best value of c_2 from [9] is around 1.69, which makes the constant κ_d defined below negative and renders the theorem useless for any dimension $d \geq 4$. (If c_2 can be made arbitrarily close to 1, we would have a different story.)

Theorem 5.1 *For an arrangement of n halfspaces in $\mathbb{R}^d$ and $1 \leq i \leq k \leq n$, the number of vertices at levels between $k-i$ and $k+i$ is $O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - 1/\kappa_d} i^{1/\kappa_d} + n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1})$, where $\kappa_d := \frac{(2\lceil d/2 \rceil - 1)c_1}{\lceil d/2 \rceil - (\lceil d/2 \rceil - 1)c_2}$.*

Proof: We apply Lemma 2.5 to each 2-d subarrangement and sum up.

For each vertex $v = \partial h_1 \cap \cdots \cap \partial h_d$ at levels between $k-i$ and $k+i$, consider the d edges of $h_1 \cap \cdots \cap h_d$. Suppose exactly j of these edges point leftward. Note that v is local x-maximum within a flat f^- of dimension j and a local x-minimum within a flat f^+ of dimension $d-j$. If $j \geq \lceil d/2 \rceil + 1$, we charge v to f^-; there are $O(n^{d-j})$ such flats and each contains at most $O(k^j)$ local x-extrema by known results [27]. If $j \geq \lfloor d/2 \rfloor - 1$, we charge v to f^+; there are $O(n^j)$ such flats and each contains at most $O(k^{d-j})$ local x-extrema. The number of vertices counted thus far is $O(n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1})$. In the remaining cases, we have $j = \lfloor d/2 \rfloor$ or $j = \lceil d/2 \rceil$, and here v is not a local x-extremum with respect to precisely $\lfloor d/2 \rfloor \lceil d/2 \rceil$ of the $\binom{d}{2}$ 2-d subarrangements.

By induction on the dimension, we may assume that at $x = \pm\infty$, the number of vertices at levels between $k-i$ and $k+i$ is bounded by $O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - 1/\kappa_{d-1}} i^{1/\kappa_{d-1}} + n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1})$.

Letting t_i be the number of vertices at levels between $k-i$ and $k+i$ in the d-dimensional arrangement, we thus get:

$$\begin{aligned} \left\lfloor \frac{d}{2} \right\rfloor \left\lceil \frac{d}{2} \right\rceil t_i &\preceq \binom{d}{2} c_1 i \Delta t_i + \left(\binom{d}{2} - \left\lfloor \frac{d}{2} \right\rfloor \left\lceil \frac{d}{2} \right\rceil \right) c_2 t_i + O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - \frac{1}{\kappa_{d-1}}} i^{\frac{1}{\kappa_{d-1}}} + n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1}) \\ \Longrightarrow \quad t_i &\preceq \kappa_d i \Delta t_i + O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - \frac{1}{\kappa_{d-1}}} i^{\frac{1}{\kappa_{d-1}}} + n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1}). \end{aligned}$$

since $\frac{\binom{d}{2} c_1}{\lfloor d/2 \rfloor \lceil d/2 \rceil - (\binom{d}{2} - \lfloor d/2 \rfloor \lceil d/2 \rceil) c_2}$ simplifies to κ_d. With the base case $t_k = O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil})$ [13], the recurrence solves to $t_i = O(n^{\lfloor d/2 \rfloor} k^{\lceil d/2 \rceil - 1/\kappa_d} i^{1/\kappa_d} + n^{\lfloor d/2 \rfloor - 1} k^{\lceil d/2 \rceil + 1})$. □

On the bichromatic k-set problem in 4-d, can Matoušek *et al.*'s proof [25] be adapted to give any new result? Also, can one get $O(k^4)$ for large k? On the bichromatic k-set problem in 2-d, can the recent improved technique in [10] be adapted?

In Section 3.1, we have given a dual reinterpretation of Sharir, Smorodinsky, and Tardos' proof for the 3-d k-sets or k-levels of planes. It has been pointed out that the proof might be difficult to adapt for k-levels of pseudo-planes [9], but with this reinterpretation, there seems more of a chance now. Perhaps previous 2-d ideas [34, 32], of using "diamond"-related configurations to replace concepts like tangents, might help. If so, we might have another avenue to deal with k-levels for nonlinear surface families (via techniques for cutting curves into pseudo-segments).

Finally, we briefly mention possible algorithmic applications of our combinatorial results. In robust variants of geometric optimization problems, it is of interest to find solutions that violate at most k of n given constraints for a given parameter k. Our results bound the combinatorial complexity of the region $\mathcal{R}$ of all such solutions (i.e., the number of vertices/edges along $\partial\mathcal{R}$). If the constraints are halfspaces and the objective function is linear (this is the problem of *linear programming with violations*), we can find an optimal solution faster than constructing the entire region $\mathcal{R}$, as it turns out. For example, for 2-d halfplanes, the author has given an $O((n+k^2)\log k)$-time randomized algorithm [7]. However, if the application requires a description of the entire region $\mathcal{R}$, or if we are dealing with a nonconvex objective function, then our results imply new computational bounds. For example, in the 2-d halfplane case, $\mathcal{R}$ has size $O(nk^{1/3} + n^{5/6-\varepsilon}k^{2/3+2\varepsilon} + k^2)$, and we can construct $\mathcal{R}$ in time within a polylogarithmic factor of this bound, by first finding all local x-minima by the algorithm in [7], and then using repeated ray shooting queries to generate all vertices of $\mathcal{R}$ as in a standard k-level algorithm [17]. We get similar applications for our results in 3-d.

Acknowledgement. Many thanks to the SODA reviewers for their comments.

References

[1] P. K. Agarwal, B. Aronov, T. M. Chan, and M. Sharir. On levels in arrangements of lines, segments, planes, and triangles. *Discrete Comput. Geom.*, 19:315–331, 1998.

[2] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *J. ACM*, 51:139–186, 2004.

[3] M. Ajtai, V. Chavátal, M. Newborn, and E. Szemerédi. Crossing-free subgraphs. *Annals of Discrete Math.*, 12:9–12, 1982.

[4] B. Aronov, B. Chazelle, H. Edelsbrunner, L. J. Guibas, M. Sharir, and R. Wenger. Points and triangles in the plane and halving planes in space. *Discrete Comput. Geom.*, 6:435–442, 1991.

[5] I. Bárány, Z. Füredi, and L. Lovász. On the number of halving planes. *Combinatorica*, 10:175–183, 1990.

[6] T. M. Chan. On levels in arrangements of curves. *Discrete Comput. Geom.*, 29:375–393, 2003.

[7] T. M. Chan. Low-dimensional linear programming with violations. *SIAM J. Comput.*, 34:879–893, 2005.

[8] T. M. Chan. On levels in arrangements of curves, II: a simple inequality and its consequences. *Discrete Comput. Geom.*, 34:11–24, 2005.

[9] T. M. Chan. On levels in arrangements of surfaces in three dimensions. In *Proc. 16th ACM-SIAM Sympos. Discrete Algorithms*, pages 232–240, 2005. Full version submitted to *Discrete Comput. Geom.*

[10] T. M. Chan. On levels in arrangements of curves, III: further improvements. In *Proc. 24th ACM Sympos. Comput. Geom.*, pages 85–93, 2008.

[11] K. L. Clarkson. A randomized algorithm for closest-point queries. *SIAM J. Comput.*, 17:830–847, 1988.

[12] K. L. Clarkson. A bound on local minima of arrangements that implies the Upper Bound Theorem. *Discrete Comput. Geom.*, 10:427–233, 1993.

[13] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.

[14] T. K. Dey. Improved bounds on planar k-sets and related problems. *Discrete Comput. Geom.*, 19:373–382, 1998.

[15] T. K. Dey and H. Edelsbrunner. Counting triangle crossings and halving planes. *Discrete Comput. Geom.*, 12:281–289, 1994.

[16] H. Edelsbrunner. *Algorithms in Combinatorial Geometry.* Springer-Verlag, Berlin, 1987.

[17] H. Edelsbrunner and E. Welzl. Constructing belts in two-dimensional arrangements with applications. *SIAM J. Comput.*, 15:271–284, 1986.

[18] D. Eppstein. Improved bounds for intersecting triangles and halving planes. *J. Combin. Theory Ser. A*, 62:176–182, 1993.

[19] P. Erdős, L. Lovász, A. Simmons, and E. Straus. Dissection graphs of planar point sets. In *A Survey of Combinatorial Theory* (J. N. Srivastava, ed.), North-Holland, Amsterdam, Netherlands, pages 139–154, 1973.

[20] T. Leighton. New lower bound techniques for VLSI. *Math. Systems Theory*, 17:47–70, 1984.

[21] J. Linhart. Arrangement of oriented hyperplanes. *Discrete Comput. Geom.*, 10:435–446, 1993.

[22] L. Lovász. On the number of halving lines. *Ann. Univ. Sci. Budapest, Eötvös, Sec. Math.*, 14:107–108, 1971.

[23] A. Marcus and G. Tardos. Intersection reverse sequences and geometric applications. *J. Combin. Theory Ser. A*, 113:675–691, 2006.

[24] J. Matoušek. *Lectures on Discrete Geometry.* Springer-Verlag, 2002.

[25] J. Matoušek, M. Sharir, S. Smorodinsky, and U. Wagner. On k-sets in four dimensions. *Discrete Comput. Geom.*, 35:177–191, 2006.

[26] K. Mulmuley. Dehn-Sommerville relations, Upper Bound Theorem, and levels in arrangements. In *Proc. 9th ACM Sympos. Comput. Geom.*, pages 240–246, 1993.

[27] K. Mulmuley. Output sensitive and dynamic constructions of higher order Voronoi diagrams and levels in arrangements. *J. Comput. Sys. Sci.*, 47:437–458, 1993.

[28] J. Pach and P. K. Agarwal. *Combinatorial Geometry.* Wiley-Interscience, New York, 1995.

[29] J. Pach, W. Steiger, and E. Szemerédi. An upper bound on the number of planar k-sets. *Discrete Comput. Geom.*, 7:109–123, 1992.

[30] M. Sharir. On k-sets in arrangements of curves and surfaces. *Discrete Comput. Geom.*, 6:593–613, 1991.

[31] M. Sharir and P. K. Agarwal. *Davenport-Schinzel Sequences and Their Geometric Applications.* Cambridge University Press, New York, 1995.

[32] M. Sharir and S. Smorodinksy. Extremal configurations and levels in pseudoline arrangements. In *Proc. 8th Workshop Algorithms Data Struct.*, Lect. Notes Comput. Sci., vol. 2748, Springer-Verlag, pages 127–139, 2003.

[33] M. Sharir, S. Smorodinsky, and G. Tardos. An improved bound for k-sets in three dimensions. *Discrete Comput. Geom.*, 26:195–204, 2001.

[34] H. Tamaki and T. Tokuyama. A characterization of planar graphs by pseudo-line arrangements. In *Proc. 8th Int. Sympos. Algorithms and Computation*, Lect. Notes Comput. Sci., vol. 1350, Springer-Verlag, pages 123–132, 1997.

[35] H. Tamaki and T. Tokuyama. How to cut pseudoparabolas into segments. *Discrete Comput. Geom.*, 19:265–290, 1998.

[36] G. Tóth. Point sets with many k-sets. *Discrete Comput. Geom.*, 26:187–194, 2001.

[37] U. Wagner. On a geometric generalization of the Upper Bound Theorem. In *Proc. 47th IEEE Sympos. Found. Comput. Sci.*, pages 635–645, 2006.

[38] R. Živaljević and S. Vrećica. The colored Tverberg's problem and complexes of injective functions. *J. Combin. Theory Ser. A*, 61:309–318, 1992.