Conflict-Free Coloring of Points with Respect to Rectangles and Approximation Algorithms for Discrete Independent Set

Timothy M. Chan[*]

March 26, 2012

Abstract

In the conflict-free coloring problem, for a given range space, we want to bound the minimum value $F(n)$ such that every set P of n points can be colored with $F(n)$ colors with the property that every nonempty range contains a unique color. We prove a new upper bound $O(n^{0.368})$ with respect to orthogonal ranges in two dimensions (i.e., axis-parallel rectangles), which is the first improvement over the previous bound $O(n^{0.382})$ by Ajwani, Elbassioni, Govindarajan, and Ray [SPAA'07]. This result leads to an $O(n^{1-0.632/2^{d-2}})$ upper bound with respect to orthogonal ranges (boxes) in dimension d, and also an $O(n^{1-0.632/(2^{d-3}-0.368)})$ upper bound with respect to dominance ranges (orthants) in dimension $d \geq 4$.

We also observe that combinatorial results on conflict-free coloring can be applied to the analysis of approximation algorithms for *discrete* versions of *geometric independent set* problems. Here, given a set P of (weighted) points and a set S of ranges, we want to select the largest(-weight) subset $Q \subset P$ with the property that every range of S contains at most one point of Q. We obtain, for example, a randomized $O(n^{0.368})$-approximation algorithm for this problem with respect to orthogonal ranges in the plane.

1   Introduction

The main topic of this paper concerns three inter-related problems in combinatorial geometry:

Definition 1.1 *(CF-coloring of elements w.r.t. ranges)* Let S be a range space (or set system) [22, 23]. A *conflict-free (or CF-)coloring* of a set P of elements is a coloring χ of P such that for every range s in S with $P \cap s \neq \emptyset$, the multiset $\{\chi(p) \mid p \in P \cap s\}$ contains a unique color. Let CF-COLOR(n, S) be the minimum number f such that every set of n elements in S has a CF-coloring with f colors.

Definition 1.2 *(DG$_k$-independent set of elements w.r.t. ranges)* Given a set P of elements in a range space S and a value k, define the *"order-k" Delaunay graph* DG$_k(P, S)$ to be the graph where the vertices are the elements of P, and pq is an edge iff there exists a range s in S such that $p, q \in s$ and $|P \cap s| \leq k$. A DG$_k$-*independent set* of P refers to an independent set of the graph DG$_k(P, S)$. Let DG-INDEP$_k(n, S)$ be the maximum number g such that every set of n elements in S has a DG$_k$-independent set of size g.

---

[*]Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, `tmchan@uwaterloo.ca`

Definition 1.3 *(DG$_k$-coloring of elements w.r.t. ranges)* A DG$_k$-*coloring* of a set P of elements refers to a vertex coloring of the graph DG$_k(P, \mathcal{S})$. Let DG-COLOR$_k(n, \mathcal{S})$ be the minimum number f such that every set of n elements has a DG$_k$-coloring with f colors.

Of the above three, CF-coloring has garnered the most fame (hence, we use the term in the title); see [25] for a recent survey. Frequency assignment in radio networks is often mentioned as one possible application; another application will be described in Section 1.3.

The standard Delaunay graph corresponds to the $k = 2$ case; for simplicity, we drop the subscript in the above definitions when $k = 2$, e.g., DG-INDEP$(n, \mathcal{S})$ = DG-INDEP$_2(n, \mathcal{S})$. In the case of orthogonal ranges, i.e., boxes,[1] the concept is further simplified because for every two points p, q, there is a unique minimal enclosing box, which we denote by $B(p, q)$. We get similar definitions for CF-coloring/DG$_k$-independent set/DG$_k$-coloring of *ranges w.r.t. elements* when we dualize the range space.

As explained below, the problems of bounding CF-COLOR$(n, \mathcal{S})$, DG-INDEP$(n, \mathcal{S})$, and DG-COLOR$(n, \mathcal{S})$ are known to be equivalent up to logarithmic factors:

Proposition 1.4 *Let α, c be constants. Assume the following monotonicity property (which holds for all range spaces considered in this paper): for every set P of elements and every range s in $\mathcal{S}$ with $|P \cap s| > 2$, there exists a range s' in $\mathcal{S}$ with $P \cap s' \subset P \cap s$ and $|P \cap s'| = 2$. Then[2]*

(i) DG-INDEP$_k(n, \mathcal{S}) = \widetilde{\Omega}(n^\alpha / k^c)$ *iff* DG-COLOR$_k(n, \mathcal{S}) = \widetilde{O}(n^{1-\alpha} k^c)$;

(ii) DG-INDEP$(n, \mathcal{S}) = \widetilde{\Omega}(n^\alpha)$ *iff* CF-COLOR$(n, \mathcal{S}) = \widetilde{O}(n^{1-\alpha})$.

Proposition 1.4(i) is easy to see, due to connection between independent set and coloring in graphs (since our graph family DG$_k$ is hereditary). Proposition 1.4(ii) follows from a simple algorithm which appeared in Even et al.'s original paper [16] on CF-coloring and was re-explained in Har-Peled and Smorodinsky's paper [19]. For completeness, a short proof is included in Appendix A.

1.1 Main Result on Points w.r.t. Rectangles in Two Dimensions

Many results on CF-coloring are known in the literature [25]. For example, there is an $O(\log n)$ upper bound in the dual for disks w.r.t. points in two dimensions (or more generally, for planar objects with linear union complexity w.r.t. points), and polylogarithmic upper bounds for two-dimensional rectangles and three-dimensional boxes w.r.t. points [16, 19]. Perhaps the most perplexing open problem in the area—the simplest case for which no polylogarithmic upper bound is currently known and yet a sublinear bound is possible—is CF-coloring of *points w.r.t. rectangles* in the plane. Let $\mathcal{B}^d$ be the range space where the elements are all points in $\mathbb{R}^d$ and the ranges are all boxes in $\mathbb{R}^d$. We provide a brief history of the problem of bounding CF-COLOR$(n, \mathcal{B}^2)$:

There are at least two different proofs that CF-COLOR$(n, \mathcal{B}^2) = O(\sqrt{n})$, i.e., DG-INDEP$(n, \mathcal{B}^2) = \Omega(\sqrt{n})$. One proof [19] uses the Erdős–Szekeres theorem to find a chain of $\Omega(\sqrt{n})$ points which is monotone in x and y simultaneously; then an independent set can be obtained by taking, say, all the even-indexed points in the chain. Another proof is divided into two cases: every graph with maximum degree at most $\sqrt{n}$ has independence number $\Omega(\sqrt{n})$; on the other hand, if some vertex in the Delaunay graph has degree exceeding $\sqrt{n}$, then its neighborhood would contain a monotone chain

---

[1] All rectangles and boxes are axis-parallel in this paper.

[2] The $\widetilde{O}$ or $\widetilde{\Omega}$ notation hides $\log^{O(1)} n$ and $\log^{O(1)} m$ factors in this paper.

of size $\Omega(\sqrt{n})$, from which an independent set of size $\Omega(\sqrt{n})$ can again be obtained. This second proof can be refined to yield a slight improvement CF-COLOR$(n, \mathcal{B}^2) = O(\sqrt{n/\log n})$ by using better bounds on the independence number of graphs with sparse neighborhoods [2], as several researchers have independently observed [19, 24].

The square-root barrier was finally broken in a more substantial way by Ajwani et al. [1], by building on ideas from a previous paper of Elbassioni and Mustafa [13]. Ajwani et al. showed that CF-COLOR$(n, \mathcal{B}^2) = O(n^{2-\phi+\varepsilon}) = O(n^{0.382})$, where ϕ is the golden ratio $(1 + \sqrt{5})/2$ and ε is an arbitrarily small constant. Consequently, DG-INDEP$(n, \mathcal{B}^2) = \Omega(n^{\phi-1-\varepsilon}) = \Omega(n^{0.618})$. In the opposite direction, Chen et al. [10] proved that DG-INDEP$(n, \mathcal{B}^2) = O(n \log^2 \log n / \log n)$.

In Section 2, we refine Ajwani et al.'s proof approach [1] to obtain the following result:

Theorem 1.5 DG-INDEP$(n, \mathcal{B}^2) = \widetilde{\Omega}(n^\alpha)$ *for some* $\alpha > 0.632$.

This implies that CF-COLOR$(n, \mathcal{B}^2) = \widetilde{O}(n^{1-\alpha}) = O(n^{0.368})$. Although the improvement is not huge, it represents the first progress in several years, and deepens the mystery surrounding the true value of the exponent. Our value of 0.632 or 0.368 originates from an interesting recurrence relation, and is a result of a computer-assisted search (alas, not a nice number like the golden ratio).

1.2 Implications for Larger k and Higher Dimensions

We next examine consequences of our result, extending Theorem 1.5 to larger $k > 2$ as well as to boxes in higher dimensions. We also consider the case of dominance ranges, i.e., orthants. Let $\mathcal{O}^d$ be the range space where the elements are points in $\mathbb{R}^d$ and the ranges are orthants of the form $(-\infty, x_1] \times \cdots \times (-\infty, x_d]$. In Section 3, we prove the following:

Theorem 1.6 *Suppose that* DG-INDEP$(n, \mathcal{B}^2) = \widetilde{\Omega}(n^\alpha)$. *Then*

(i) DG-INDEP$_k(n, \mathcal{B}^2) = \widetilde{\Omega}([n/k]^\alpha)$, *and more generally,* DG-INDEP$_k(n, \mathcal{B}^d) = \widetilde{\Omega}([n/k]^{\alpha/2^{d-2}})$ *for any fixed* $d \geq 2$;

(ii) DG-INDEP$_k(n, \mathcal{O}^d) = \widetilde{\Omega}([n/k]^{\alpha/(2^{d-3}-1+\alpha)})$ *for any fixed* $d \geq 4$.

In the two-dimensional case, the extension to larger k in Theorem 1.6(i) is not supposed to be trivial and exploits ideas specific to the range space $\mathcal{B}^2$. In higher dimenisons, Har-Peled and Smorodinksy [19] previously stated a weaker bound CF-COLOR$(n, \mathcal{B}^d) = O(n^{1-1/2^{d-1}})$, i.e., DG-INDEP$(n, \mathcal{B}^d) = \Omega(n^{1/2^{d-1}})$. The proof of Theorem 1.6(ii) requires further ideas, and in particular, the consideration of an intermediate case of points w.r.t. 3-dimensional 4-sided boxes. The bounds deteriorate quickly as the dimension increases, but the bounds for small d are still interesting, e.g., we have DG-INDEP$(n, \mathcal{O}^4) = \widetilde{\Omega}(n^{\alpha/(1+\alpha)}) = \Omega(n^{0.387})$.

The dual $(\mathcal{B}^d)^*$ of the range space $\mathcal{B}^d$ may be viewed as a special case of $\mathcal{O}^{2d}$ by mapping a box $[a_1, b_1] \times \cdots \times [a_d, b_d]$ to the point $(a_1, -b_1, \ldots, a_d, -b_d) \in \mathbb{R}^{2d}$. Thus, as a corollary of Theorem 1.6(ii), DG-INDEP$_k(n, (\mathcal{B}^d)^*) = \widetilde{\Omega}([n/k]^{\alpha/(2^{2d-3}-1+\alpha)})$.

1.3 Application to Discrete Independent Set

We also point out an application to approximation algorithms for the following optimization problem, which is actually the original motivation of this paper:

3

Problem 1.7 *(Discrete independent set of elements w.r.t. ranges)* Given a set P of n (weighted) elements and a set S of m ranges in a range space $\mathcal{S}$, find the largest(-weight) subset $Q \subseteq P$ such that each range $s \in S$ contains at most one element of Q.

We obtain a similar optimization problem of discrete independent set of *ranges w.r.t. elements* when we dualize the range space: namely, given a set P of m elements and a set S of n (weighted) ranges in a range space $\mathcal{S}$, find the largest(-weight) subset $Q \subseteq S$ such that each element $p \in P$ is contained in at most one range of Q. When P is taken as the entire universe, this dual problem is equivalent to finding maximum independent sets in intersection graphs—this "continuous" version of geometric independent set has been studied extensively (e.g., see [6, 7, 9, 18]). The unweighted discrete independent set problem can be viewed as the LP-dual of geometric *set cover* and geometric *hitting set*, both of which have also received considerable attention (e.g., see [5, 8, 12]).

In contrast, the discrete version of geometric independent set has not been studied as often. We are aware of only two papers: Chan and Har-Peled's work [9] has a short section showing that their $O(1)$-approximation algorithm for independent set of pseudo-disks can be extended to discrete independent set of pseudo-disks w.r.t. points. Ene, Har-Peled, and Raichel's work [14] contains multiple results on a problem that generalizes discrete independent set (namely, replacing the "at most one element/range" with "at most the capacity of the element/range" in the constraints). For example, they gave polylogarithmic-approximation algorithms for discrete independent set of two-dimensional rectangles and three-dimensional boxes w.r.t. points.

We observe the following connection:

Theorem 1.8 *Consider the following linear program (LP) for Problem 1.7:*

$$
\begin{aligned}
\text{maximize} \quad & \sum_{p \in P} w_p x_p \\
\text{subject to} \quad & \sum_{p \in s} x_p \leq 1 \quad \forall s \in S \\
& 0 \leq x_p \leq 1 \quad \forall p \in P,
\end{aligned}
$$

where $w_p > 0$ denotes the weight of p. Let opt *be the optimal value of the LP and* OPT *be the optimal value of the original problem. Suppose* DG-INDEP$_k(n, \mathcal{S}) = \widetilde{\Omega}(n^\alpha / k^{O(1)})$. *Then* OPT $\geq \widetilde{\Omega}(\text{opt}^\alpha) \geq \widetilde{\Omega}(\text{OPT}^\alpha)$ *in the unweighted case, and* OPT $\geq \widetilde{\Omega}(\text{opt}/n^{1-\alpha}) \geq \widetilde{\Omega}(\text{OPT}/n^{1-\alpha})$ *in the weighted case.*

The proof, given in Section 4, uses random sampling as in [9] and is straightforward (especially since we are not too concerned with optimizing logarithmic factors here, unlike in [9]). For this LP, there are unweighted instances where OPT $\leq$ DG-INDEP$(n, \mathcal{S})$ but opt $\geq n/2$: just let S consist of all (combinatorially different) ranges in $\mathcal{S}$ that contain exactly two points in the set P that realizes DG-INDEP$(n, \mathcal{S})$, and set all x_p's to $1/2$. Thus, the above theorem is in some sense tight (ignoring logarithmic factors).

From the bound on DG-INDEP$_k(n, \mathcal{B}^2)$ implied by Theorems 1.5 and 1.6(i), we immediately obtain an $O(n^{0.368})$-approximation algorithm for weighted discrete independent set of points w.r.t. rectangles in two dimensions, or an $\widetilde{O}(\text{OPT}^{0.368})$-approximation algorithm in the unweighted case; to the author's knowledge, no previous approximation algorithms were given for this particular problem. From the bound on DG-INDEP$_k(n, (\mathcal{B}^d)^*)$ implied by Theorems 1.5 and 1.6(ii), we obtain the first sublinear approximation algorithm for discrete independent set of boxes w.r.t. points in dimensions $d \geq 4$; Ene et al. [14] previously stated the $d = 4$ case as an open problem. Since the proofs of all the theorems are constructive (using randomization in some places), it can be verified that the resulting approximation algorithms run in Las Vegas polynomial time.

We remark that our linking of LP-based approximation algorithms to combinatorial geometry problems continues an existing trend. For example, bounding the integrality gap of the standard LP for unweighted geometric set cover and hitting set problems is equivalent (up to constant factors) to the ε-net problem [17, 21]. For (continuous or discrete) independent set of planar objects w.r.t. points, Chan and Har-Peled [9] have shown previously that an upper bound on the integrality gap can be obtained from the *union complexity* (but unfortunately boxes may have high union complexity even in two dimensions).

2 Points w.r.t. Rectangles in Two Dimensions (Proof of Theorem 1.5)

In this section, we prove that $\text{DG-COLOR}(n, \mathcal{B}^2) = \widetilde{O}(n^{1-\alpha})$ for some $\alpha > 0.632$, thereby showing that $\text{DG-INDEP}(n, \mathcal{B}^2) = \widetilde{\Omega}(n^\alpha)$. Our proof heavily builds on Ajwani et al.'s grid-based approach [1], but our description will be kept self-contained. We will solve a generalization of the problem, introducing two additional parameters v and h:

Problem 2.1 Given a set P of $\leq n$ points in the plane and a set V of $< v$ vertical lines and a set H of $< h$ horizontal lines, define a *"generalized Delaunay graph"* $\text{DG}(P, V, H, \mathcal{B}^2)$ where the vertices are the points in P, and pq is an edge iff $B(p,q) \cap P = \{p, q\}$, and p *and* q *are separated by both a vertical line in* V *and a horizontal line in* H. We want to find a coloring of the graph $\text{DG}(P, V, H, \mathcal{B}^2)$, which in this section will be referred simply as a *coloring of P relative to V and H*.

Let $F(n, v, h)$ be the minimum number of colors needed to solve this problem. It will be more convenient to work with the function $G(n, v, h) = n/F(n, v, h)$.

Initially, we can set V to consist of $n-1$ vertical lines separating all the points and H to consist of $n-1$ horizontal lines separating all the points; then a coloring of P relative to V and H would be a DG-coloring of P. Thus, our goal is to bound $G(n, n, n)$.

For "base cases" of our recursive algorithm, we can start with a weak bound $F(n, v, h) = O(\sqrt{n})$ by one of the known methods [19]. Another upper bound is $F(n, v, h) \leq v$, which follows from the trivial method of assigning the same color to all points between two consecutive vertical lines of V. Rewriting in terms of the function G, we have

$$G(n, v, h) \geq \Omega(\sqrt{n}) \tag{1}$$
$$G(n, v, h) \geq n/v. \tag{2}$$

Symmetry gives the following additional relationship:

$$G(n, v, h) = G(n, h, v). \tag{3}$$

It turns out a further generalization involving a fourth parameter r_0 will be useful for our recursive algorithm:

Definition 2.2 We say that (x_1, y_1) *(++)-dominates* (x_2, y_2) if $x_1 > x_2$ and $y_1 > y_2$. Similarly, (x_1, y_1) *(+−)-dominates* (x_2, y_2) if $x_1 > x_2$ and $y_1 < y_2$. For any sign sequence σ, we say that a point p is *σ-maximal* in a set P if $p \in P$ and p is σ-dominated by no other point in P.

Given a grid $\mathcal{G}_0$ formed by a set of vertical and horizontal lines, we say that a point set P *respects* $\mathcal{G}_0$ if for every grid cell τ in $\mathcal{G}_0$, every point in $P \cap \tau$ is (++)-maximal or (+−)-maximal in $P \cap \tau$.
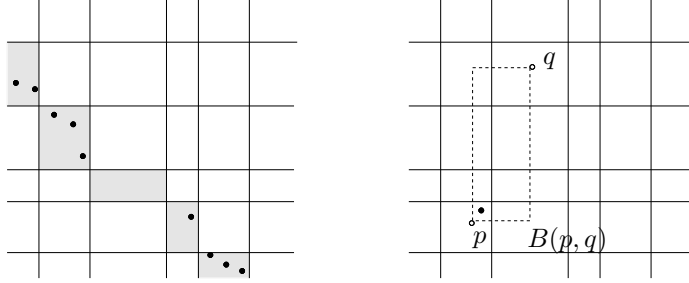
Figure 1: (Left) The $(++)$-maximal points in the cells of one diagonal of a grid $\mathcal{G}_0$. (Right) Two unmarked points p and q lying in different columns and different rows of a grid $\mathcal{G}$.

Let $F_{r_0}(n, v, h)$ be the minimum number of colors needed to solve Problem 2.1, under the knowledge that the point set P respects an $r_0 \times r_0$ grid $\mathcal{G}_0$. Let $G_{r_0}(n, v, h) = n/F_{r_0}(n, v, h)$. (Of course, every point set P respects a sufficiently fine, $n \times n$ grid $\mathcal{G}_0$, so $G(n, v, h) = G_n(n, v, h)$.)

For base case, we have the upper bound $F_{r_0}(n, v, h) = O(r_0)$: This follows by observing that the $(++)$-maxima (resp. $(+-)$-maxima) in each grid cell in $\mathcal{G}_0$ form a monotone chain, i.e., a downward (resp. upward) "staircase". Since the grid cells of $\mathcal{G}_0$ can be partitioned into $O(r_0)$ northwest-to-southeast or southwest-to-northeast diagonals (see Figure 1(left) for one such diagonal), the whole point set P can be decomposed into $O(r_0)$ monotone chains. In each chain, we can easily color the points alternately with two colors. The claim follows. In terms of the function G,

$$G_{r_0}(n, v, h) \;\geq\; \Omega(n/r_0). \tag{4}$$

2.1   The Main Algorithm

We now present our main recursive algorithm to solve Problem 2.1. Let $r \in [1, \min\{v, h\}]$ be a parameter. The algorithm consists of the five simple steps:

1. Form a grid $\mathcal{G}$ such that each column contains $\leq n/r$ points of P and $< v/r$ vertical lines of V, and each row contains $\leq n/r$ points of P and $< h/r$ horizontal lines of H. At most $O(r)$ vertical and horizontal grid lines suffice.

2. For each grid cell τ of $\mathcal{G}$, mark the $(++)$-maxima and $(+-)$-maxima of $P \cap \tau$. Find a coloring of all the marked points in P. Let C_0 be the palette used.

3. For each column ξ of $\mathcal{G}$, recursively find a coloring of the unmarked points in $P \cap \xi$ *relative to the vertical lines of V inside ξ and the horizontal grid lines of $\mathcal{G}$*. Use a common palette C in all the columns. Make sure that each coloring is "distributed fairly"—see details below.

4. For each row η of $\mathcal{G}$ and each $i \in C$, recursively find a coloring of the unmarked points in $P \cap \eta$ that have received color i in Step 3, *relative to all the vertical lines of V and the horizontal lines of H inside η*. Use a common palette C' in all the columns.

5. An unmarked point that has received color i in Step 3 and color j in Step 4 is given a final color (i, j) in the palette $C \times C'$ (assumed to be disjoint from C_0).

6

Correctness. We now verify that the result is a valid coloring of $\mathrm{DG}(P, V, H, \mathcal{B}^2)$. Consider two points $p, q \in P$ such that $B(p, q) \cap P = \{p, q\}$ and p and q are separated by a vertical line of V and a horizontal line of H. If both p and q are marked, they must have different colors by Step 2. If one is marked and the other is unmarked, they obviously have different colors. From now on, we may assume that p and q are both unmarked. Three cases may arise:

- p and q lie in different columns and different rows of $\mathcal{G}$. If p is to the left of and below q as in Figure 1(right) (all other cases are similar), then $B(p, q)$ must contain a $(++)$-maximal point inside the grid cell containing p; this point must be different from p and q since p is unmarked—a contradiction.

- p and q lie in the same column but in different rows of $\mathcal{G}$. Then p and q are separated by a horizontal grid line of $\mathcal{G}$. They must receive different colors in Step 3, and thus different final colors.

- p and q lie in the same row of $\mathcal{G}$. If they receive the same color in Step 3, they will receive different colors in Step 4. Thus, they must have different final colors.

Details. Step 2 can be done using a palette C_0 of $F_{cr}(n, v, h)$ colors for some constant c, since the marked points respect the $cr \times cr$ grid $\mathcal{G}$.

Step 3 can be done using a palette C of $f = F(n/r, v/r, cr)$ colors. We ensure that the coloring of the unmarked points in $P \cap \xi$ for each column ξ is *distributed fairly*, i.e., satisfies the following properties: (a) the maximum size of the color classes is $O(|P \cap \xi|/f)$, and (b) each point has probability $O(1/f)$ of receiving any fixed color.

To ensure (a), when a color class contains more than $2|P \cap \xi|/f$ points, we can just subdivide it into color subclasses of size $\Theta(|P \cap \xi|/f)$. The number of colors remains $O(f)$. To ensure (b), we can just pick a random permutation π_ξ of the color palette and rename color i as color $\pi_\xi(i)$.

Let P_i denote the subset of all unmarked points that have received color i after Step 3. Consider a row η and a color i. We analyze the random variable $|P_i \cap \eta|$. Since each point has probability $O(1/f)$ of receiving color i by (b), $|P_i \cap \eta|$ has expected value $\mu = O(n/(rf))$. Furthermore, $|P_i \cap \eta| = \sum_\xi |P_i \cap \xi \cap \eta|$, where the sum is over all columns ξ. The terms in the sum are independent random variables, since the permutations π_ξ's are independent. Because $|P_i \cap \xi| \leq O(|P \cap \xi|/f) = O(n/(rf))$ by (a), each term in the sum is upper-bounded by $M = O(n/(rf))$. By the Chernoff–Hoeffding inequality [20] for sums of bounded random variables (after rescaling the variables and the mean by a factor of $1/M$),

$$\Pr[|P_i \cap \eta| > (1 + \delta)\mu] \leq \left[\frac{e^\delta}{(1 + \delta)^{1+\delta}} \right]^{\mu/M}$$

for any $\delta > 0$ (possibly large). Choosing $\delta = \Theta((1 + M/\mu)\log n)$, we have $|P_i \cap \eta| \leq O((n/(rf))\log n)$ with high probability (e.g., probability $1 - o(1/n^2)$). So, by repeating for an expected $O(1)$ number of trials, we can find a choice of random permutations that ensure $|P_i \cap \eta| \leq \widetilde{O}(n/(rf))$ for all rows η and all colors i at the same time.

Step 4 can therefore be done using a palette C' of $F(c'n/(rf), v, h/r)$ colors for some $c' = \widetilde{O}(1)$.

7

The recurrence. The final palette $C_0 \cup (C \times C')$ has size $F_{cr}(n, v, h) + O(f) F(c'n/(rf), v, h/r)$, where $f = F(n/r, v/r, cr)$. We conclude that for any $r \in [1, \min\{v, h\}]$,

$$F(n, v, h) \; \leq \; F_{cr}(n, v, h) \; + \; O(F(n/r, v/r, cr)) \, F\left(\frac{c'n/r}{F(n/r, v/r, cr)}, v, h/r\right).$$

When expressed in terms of the function $G(n, v, h) = n/F(n, v, h)$, this peculiar recurrence is simplified somewhat. Make the following definition: $x \oplus y = z$ iff $1/x + 1/y = 1/z$. Then for any $r \in [1, \min\{v, h\}]$,

$$G(n, v, h) \; \geq \; G_{cr}(n, v, h) \; \oplus \; \widetilde{\Omega}(r) \, G\left(c' G\left(\frac{n}{r}, \frac{v}{r}, cr\right), v, \frac{h}{r}\right). \tag{5}$$

Note that one can think of the operator $\oplus$ as a shorthand for min, since they have the same behavior asymptotically when the number of terms is small: more precisely, $(1/k) \min\{x_1, \ldots, x_k\} \leq x_1 \oplus \cdots \oplus x_k \leq \min\{x_1, \ldots, x_k\}$ for all $x_1, \ldots, x_k > 0$.

2.2 First Analysis of the Recurrence

We now analyze the recurrence (5) (in conjunction with (1)–(4)). First, we ignore the effects of the parameters v and h. Assume that a bound $G(n, v, h) = \widetilde{\Omega}(n^\alpha)$ is known. Then (5) together with (4) implies that

$$G(n, v, h) \; \geq \; \widetilde{\Omega}\left(n/r \; \oplus \; r[n/r]^{\alpha^2}\right).$$

By setting $r = n^{(1-\alpha^2)/(2-\alpha^2)}$ to balance the two terms, we obtain a new bound $G(n, v, h) = \widetilde{\Omega}(n^{1/(2-\alpha^2)})$.

We can apply this process a constant number of times, starting with the base case (1). It is straightforward to verify that the sequence $\alpha_0 = 1/2$, $\alpha_{i+1} = 1/(2 - \alpha_i^2)$ converges to a root of the equation $\alpha = 1/(2 - \alpha^2)$, namely, $\phi - 1 = (\sqrt{5} - 1)/2 > 0.618$. For any arbitrarily small constant $\varepsilon > 0$, we thus get $G(n, n, n) = \widetilde{\Omega}(n^{\phi-1-\varepsilon}) = \Omega(n^{0.618})$ and $F(n, n, n) = O(n^{0.382})$, recovering the previous result by Ajwani et al. [1].

Although our explanation might be somewhat cleaner, our algorithm is identical to Ajwani et al.'s when V and H are absent. Our main innovation is in considering colorings *relative to the horizontal grid lines* during Step 3. Because Ajwani et al. did not do this, the case when unmarked points p and q lie in both the same column and the same row is "handled" twice—our idea eliminates this redundancy.

It may not be immediately obvious that taking into account v and h can indeed lead to an improvement, however. One needs more than one level of recursion to see an effect. The following is a shortest chain of applications of (5) that yield a new result. In each instance, we use (4) for the first term $G_r(n, r, h)$ and the bound $G(n, v, h) = \Omega(n^{0.618})$ for the outer G in the second term of the right-hand side of (5):

$$G(n, n, n) \; \geq \; \widetilde{\Omega}\left(n^{0.622} \; \oplus \; n^{0.378}[G(n^{0.622}, n^{0.622}, cn^{0.378})]^{0.618}\right)$$

$$G(n^{0.622}, n^{0.622}, cn^{0.378}) \; = \; G(n^{0.622}, cn^{0.378}, n^{0.622})$$

$$\geq \; \widetilde{\Omega}\left(n^{0.395} \; \oplus \; n^{0.227}[G(n^{0.395}, cn^{0.151}, cn^{0.227})]^{0.618}\right)$$

$$G(n^{0.395}, cn^{0.151}, cn^{0.227}) \; \geq \; \widetilde{\Omega}\left(n^{0.272} \; \oplus \; n^{0.123}[G(n^{0.272}, cn^{0.028}, cn^{0.123})]^{0.618}\right)$$

$$G(n^{0.272}, cn^{0.028}, cn^{0.123}) \; \geq \; \Omega(n^{0.272-0.028}) \; = \; \Omega(n^{0.244}) \; \text{by (2)}.$$

8

Substituting in reverse, we conclude that $G(n, n, n) = \widetilde{\Omega}(n^{0.622})$, i.e., $F(n, n, n) = \widetilde{\Omega}(n^{0.378})$.

The recurrence appears hard to solve exactly, but clearly further improvements should be possible with more iterations. Later in Section 2.4, we will describe a way to make the setting of r and the verification process more systematic. However, we will not explore the best bound achievable by (5) at this point, since a different version of the algorithm and recurrence turns out to be more effective...

2.3 Variants of the Algorithm

For example, a similar algorithm gives the following alternative recurrence for any $r \in [1, \min\{v, h\}]$:

$$G(n, v, h) \ \geq \ G_{cr}(n, v, h) \ \oplus \ \widetilde{\Omega}(r) \, G\left(c'G\left(\frac{n}{r}, \frac{v}{r}, h\right), cr, \frac{h}{r}\right). \tag{6}$$

This is obtained by modifying the algorithm so that the colorings in Step 3 are now *relative to the vertical lines of V inside ξ and all the horizontal lines of H* and the colorings in Step 4 are now *relative to the vertical grid line of $\mathcal{G}$ and the horizontal lines of H inside η*. In the correctness proof, Step 3 takes care of the case when p and q lie in the same column, and Step 4 takes care of the case when p and q lie in the same row but in different columns.

Better still, we can prove the following for any $r \in [1, \min\{r_0, v, h\}]$:

$$G_{r_0}(n, v, h) \ \geq \ G_{cr}(n, v, h) \ \oplus \ \widetilde{\Omega}(r \cdot (r/r_0)) \, G\left(c'G\left(\frac{n}{r}, \frac{v}{r}, cr\right), cr, \frac{h}{r}\right). \tag{7}$$

Here, we need several modifications to the algorithm, knowing that the given point set P respects an $r_0 \times r_0$ grid $\mathcal{G}_0$: First, in step 1, by adding $O(r)$ vertical and horizontal grid lines to $\mathcal{G}$, we can ensure that each column of $\mathcal{G}$ contains $< r_0/r$ vertical grid lines of $\mathcal{G}_0$, and each row of $\mathcal{G}$ contains $< r_0/r$ horizontal grid lines of $\mathcal{G}_0$. As before, the colorings in Step 3 are *relative to the vertical lines of V inside ξ and the horizontal grid lines of $\mathcal{G}$*, but the colorings in step 4 are now *relative to the vertical grid line of $\mathcal{G}$ and the horizontal lines of H inside η*. We perform an extra step: inside each grid cell τ of $\mathcal{G}$, we compute a coloring of the unmarked points in $P \cap \tau$ using a common palette C''; since $P \cap \tau$ respects an $\lfloor r_0/r \rfloor \times \lfloor r_0/r \rfloor$ grid, C'' requires $O(r_0/r)$ colors by (4) (which accounts for the extra r/r_0 factor in (7)). In Step 5, an unmarked point that has received color i in Step 3, color j in Step 4, color ℓ in the extra step is given a final color (i, j, ℓ) in the palette $C \times C' \times C''$.

In the correctness proof, Step 3 takes care of the case when p and q lie in the same column but in different rows of $\mathcal{G}$, and Step 4 takes care of the case when p and q lie in the same row but in different columns of $\mathcal{G}$. The extra step takes of the case when p and q lie in both the same row and the same column of $\mathcal{G}$. This establishes (7).

Using (7) to bound $G(n, v, h)$ requires some preparation, since we need to keep the ratio r_0/r small in each application. We first start with the following: for $r^* = \min\{v, h\}$,

$$G(n, v, h) \ \geq \ G_{cr^*}(n, v, h) \ \oplus \ \widetilde{\Omega}(r^*) \, G\left(c'G\left(\frac{n}{r^*}, \frac{v}{r^*}, cr^*\right), cr^*, \frac{h}{r^*}\right).$$

If $h \leq v$, then $r^* = h$ and the above inequality follows from (5) since by (2), $G(x, v, 1) = x = G(x, cr^*, 1)$. If $v \leq h$, then $r^* = v$ and the above inequality follows from (6) since by (2), $G(x, 1, h) = x = G(x, 1, cr^*)$.

Next, by applying (7) with $r_0 = cr^*/2^i$ and $r = r^*/2^{i+1}$ repeatedly for logarithmically many iterations and terminating with the base case (4) with $r = r_*$, we obtain

$$G(n, v, h) \geq \Omega(n/r_*) \oplus \min_{r_* \leq r \leq \min\{h,v\}} \widetilde{\Omega}(r) \, G\left(c'G\left(\frac{n}{r}, \frac{v}{r}, cr\right), cr, \frac{h}{r}\right) \tag{8}$$

for any $r_* \in [1, \min\{v, h\}]$. (Remember that $\oplus$ behaves like min, since the number of terms is $O(\log n) = \widetilde{O}(1)$.)

We will see that the min in the above formula tends to be attained by the term for $r = r_*$, and if so, (8) indeed subsumes both (5) and (6) simultaneously in some sense. From now on, we work with (8) in place of (5)–(7).

2.4   Better Analysis of the Alternative Recurrence

We now obtain bounds on $G(n, v, h)$ using the recurrence (8) together with (1)–(3). Specifically, we show how to derive a new bound from two known bounds.

Assume that $G(n, v, h) = \widetilde{\Omega}(n^{\alpha_1} v^{\beta_1} h^{\gamma_1})$ and $G(n, v, h) = \widetilde{\Omega}(n^{\alpha_2} v^{\beta_2} h^{\gamma_2})$, with $\alpha_1, \alpha_2 \in [0, 1]$ and $\beta_1, \beta_2, \gamma_1, \gamma_2 \leq 0$. Then by (8),

$$G(n, v, h) \geq \widetilde{\Omega}\left(n/r_* \oplus \min_{r_* \leq r \leq \min\{v,h\}} r\left[(n/r)^{\alpha_1}(v/r)^{\beta_1} r^{\gamma_1}\right]^{\alpha_2} r^{\beta_2}[h/r]^{\gamma_2}\right)$$

$$\geq \widetilde{\Omega}\left(n/r_* \oplus \min_{r_* \leq r \leq \min\{v,h\}} n^{\alpha_1\alpha_2} v^{\beta_1\alpha_2} h^{\gamma_2} r^{1-(\alpha_1+\beta_1-\gamma_1)\alpha_2+\beta_2-\gamma_2}\right).$$

Assuming $1 - (\alpha_1 + \beta_1 - \gamma_1)\alpha_2 + \beta_2 - \gamma_2 \geq 0$, we see that the minimum above is attained when $r = r_*$. By setting $r_* = [n^{\alpha_1\alpha_2-1} v^{\beta_1\alpha_2} h^{\gamma_2}]^{-1/(2-(\alpha_1+\beta_1-\gamma_1)\alpha_2+\beta_2-\gamma_2)}$ to balance the two terms,[3] we conclude that $G(n, v, h) \geq \widetilde{\Omega}(n^\alpha v^\beta h^\gamma)$, where

$$(\alpha, \beta, \gamma) = \mathbf{g}((\alpha_1, \beta_1, \gamma_1), (\alpha_2, \beta_2, \gamma_2)) = (1, 0, 0) + \frac{1}{2-(\alpha_1+\beta_1-\gamma_1)\alpha_2+\beta_2-\gamma_2}(\alpha_1\alpha_2 - 1, \, \beta_1\alpha_2, \, \gamma_2). \tag{9}$$

If $1 - (\alpha_1 + \beta_1 - \gamma_1)\alpha_2 + \beta_2 - \gamma_2 < 0$, we make $\mathbf{g}$ undefined.

We can apply the above process in conjunction with (3) a constant number of times, in any order, starting with the base case (2) and the known bound $G(n, v, h) = \Omega(n^{0.618})$. In other words, letting

$$\mathbf{s}(\alpha, \beta, \gamma) = (\alpha, \gamma, \beta), \quad \mathbf{u} = (1, -1, 0), \quad \mathbf{v}_0 = (0.618, 0, 0), \tag{10}$$

we have $G(n, v, h) = \widetilde{\Omega}(n^\alpha v^\beta h^\gamma)$ for any vector (α, β, γ) that can be generated by a constant number of applications of the mappings $\mathbf{g}$ and $\mathbf{s}$ to the initial vectors $\mathbf{u}$ and $\mathbf{v}_0$.

Numerical calculations.   For example, the specific chain of inequality used in Section 2.2 is analogous to taking the expression

$$\mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{u}, \mathbf{v}_0), \mathbf{v}_0)), \mathbf{v}_0),$$

which indeed evaluates to a vector exceeding $(0.622, 0, 0)$. We can immediately do better by iterating:

$$\mathbf{v}_{i+1} = \mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{u}, \mathbf{v}_i), \mathbf{v}_i)), \mathbf{v}_i)$$

---

[3]It is easy to check that $r_* \geq 1$. The bound remains true for $r_* > \min\{v, h\}$, since (2) gives $G(n, v, h) \geq \Omega(n/\min\{v, h\}) \geq \Omega(n/r_*)$.

gives a vector $\geq (0.624, 0, 0)$ for a sufficiently large constant i. (Of course, at this stage, the evaluations are best done on a computer than a hand calculator.) Modifying the expression to

$$\mathbf{v}_{i+1} = \mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{u}, \mathbf{s}(\mathbf{g}(\mathbf{u}, \mathbf{s}(\mathbf{u}))))), \mathbf{v}_i)), \mathbf{v}_i)$$

gives $\geq (0.627, 0, 0)$. Expanding and further tweaking the expression, we find that letting $\mathbf{w}_0 = \mathbf{s}(\mathbf{u})$, $\mathbf{w}_{i+1} = \mathbf{g}(\mathbf{u}, \mathbf{w}_i)$, and

$$\mathbf{v}_{i+1} = \mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{w}_5, \mathbf{s}(\mathbf{g}(\mathbf{w}_3, \mathbf{s}(\mathbf{g}(\mathbf{w}_6, \mathbf{s}(\mathbf{g}(\mathbf{w}_2, \mathbf{w}_0))))))))), \mathbf{v}_i)), \mathbf{v}_i)$$

gives $\geq (0.629, 0, 0)$. With more ad hoc experimentation, we discover that

$$\mathbf{v}_{i+1} = \mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{g}(\mathbf{w}_7, \mathbf{g}(\mathbf{w}_7, \mathbf{w}_2)), \mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{w}_2, \mathbf{w}_0)), \mathbf{s}(\mathbf{g}(\mathbf{w}_6, \mathbf{s}(\mathbf{g}(\mathbf{w}_2, \mathbf{w}_0))))))))), \mathbf{v}_i)), \mathbf{v}_i)$$

gives $\geq (0.630, 0, 0)$. Continuing still further and letting $\mathbf{z}_0 = \mathbf{s}(\mathbf{w}_1)$, $\mathbf{z}_{i+1} = \mathbf{g}(\mathbf{u}, \mathbf{z}_i)$, and

$$\begin{aligned}\mathbf{v}_{i+1} \;=\; & \mathbf{g}(\mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{g}(\mathbf{g}(\mathbf{g}(\mathbf{u}, \mathbf{g}(\mathbf{u}, \mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{w}_2, \mathbf{s}(\mathbf{z}_1)))))), \mathbf{g}(\mathbf{g}(\mathbf{u}, \mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{w}_2, \mathbf{s}(\mathbf{z}_1)))), \mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{z}_2, \\ & \mathbf{s}(\mathbf{g}(\mathbf{w}_2, \mathbf{z}_0))), \mathbf{s}(\mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{w}_2, \mathbf{s}(\mathbf{z}_1))))))))), \mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{g}(\mathbf{z}_3, \mathbf{g}(\mathbf{z}_2, \mathbf{s}(\mathbf{g}(\mathbf{u}, \mathbf{s}(\mathbf{z}_1))))), \mathbf{g}(\mathbf{g}(\mathbf{z}_2, \\ & \mathbf{s}(\mathbf{g}(\mathbf{z}_2, \mathbf{s}(\mathbf{z}_1)))), \mathbf{s}(\mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{z}_2, \mathbf{s}(\mathbf{z}_1)))))), \mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{w}_4, \mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{z}_2, \mathbf{s}(\mathbf{z}_1)))), \mathbf{s}(\mathbf{g}(\mathbf{g}(\mathbf{z}_2, \\ & \mathbf{s}(\mathbf{g}(\mathbf{w}_2, \mathbf{s}(\mathbf{z}_1)))), \mathbf{s}(\mathbf{g}(\mathbf{w}_3, \mathbf{g}(\mathbf{z}_2, \mathbf{s}(\mathbf{z}_1))))))))))))), \mathbf{v}_i)), \mathbf{v}_i)\end{aligned}$$

give $\geq (0.631, 0, 0)$.

This last expression is not generated by hand (as the reader might guess), but automatically by a C program which employs a simple greedy expansion strategy. Roughly speaking, in each round, the program evaluates the arguments (n, v, h) at all nodes of the current expression tree; then at each leaf, it decides whether to apply the $\mathbf{s}$ map first before applying $\mathbf{g}$, based on whether $v > h$, to produce the next level of the tree.

The same program finds an expression that gives a vector exceeding $(0.632, 0, 0)$, although the expression is just a bit more complicated to write down—the curious reader can see the whole thing in Appendix B. (We have an even longer expression giving $(0.6322, 0, 0)$, but decide not to waste more paper.) We conclude that $G(n, n, n) = \widetilde{\Omega}(n^\alpha)$ and $F(n, n, n) = \widetilde{O}(n^{1-\alpha})$ for some $\alpha > 0.632$. This proves Theorem 1.5.

Remarks. Although there is no guarantee the exponent obtained by our greedy heuristics should converge to the optimal value for our recurrence, new ideas are probably required to get a more substantial improvement.

In fact, our recurrence cannot yield a bound for α greater than $2/3 \approx 0.667$. For example, the function $G(n, v, h) = \max\{n^{2/3}, n/\sqrt{v}, n/\sqrt{h}\}$ (which has $G(n, n, n) = n^{2/3}$) obeys (8) and (1)–(3), as one can verify by hand with a tedious case analysis. One could try to find worse examples by using another computer search, experimenting with more complicated functions of the form $G(n, v, h) = \max_{(\alpha, \beta, \gamma) \in S} n^\alpha v^\beta h^\gamma$ for some finite sets S.

3 Implications

In this section, we consider the case of boxes in higher dimensions. Given $d \leq j \leq 2d$, let $\mathcal{B}^{d,j}$ be the range space where the elements are all points in $\mathbb{R}^d$ and the ranges are all j-sided boxes of the form $[x_1, x'_1] \times [x_{j-d}, x'_{j-d}] \times (-\infty, x_{j-d+1}] \times \cdots \times (-\infty, x_d]$. For example, $\mathcal{B}^{d,2d} = \mathcal{B}^d$ and $\mathcal{B}^{d,d} = \mathcal{O}^d$.

3.1 Points w.r.t. 3-Dimensional 4-Sided Boxes

For the case of points w.r.t. 3-dimensional orthants in $\mathcal{O}^3$ (3-sided boxes), which is equivalent to orthants in $\mathcal{O}^3$ w.r.t. points by self-duality, good bounds can be obtained because the union complexity is low, which generally implies that the order-k Delaunay graph is sparse. Aloupis et al. [4] observed this fact specifically for the case of pseudo-disks w.r.t. points; they used a sampling-based proof. Chan and Har-Peled [9] gave a different sampling-based proof in the style of "shallow cuttings." In Appendix A, we include for completeness an adaptation of the latter proof to the 3-dimensional orthant case.

Proposition 3.1

(i) *The graph* $\mathrm{DG}_k(P, \mathcal{O}^3)$ *has* $O(nk)$ *edges.*

(ii) $\mathrm{DG}\text{-}\mathrm{INDEP}_k(P, \mathcal{O}^3) = \Omega(n/k)$.

We are thus led to the next nontrivial case of 3-dimensional 4-sided boxes in $\mathcal{B}^{3,4}$; this case turns out to be useful later in Section 3.3.

First the concept of dominance and maximality in Definition 2.2 can be extended to three dimensions in the obvious way: for example, (x_1, y_1, z_1) $(+--)$-dominates (x_2, y_2, z_2) if $x_1 > x_2$ and $y_1 < y_2$ and $z_1 < z_2$. The σ-*orthant* of a point q is the region of all points that σ-dominate q.

We reduce the case of 3-dimensional 4-sided boxes to the 2-dimensional rectangular case in a series of three stages:

Lemma 3.2 *Suppose that* $\mathrm{DG}\text{-}\mathrm{INDEP}(n, \mathcal{B}^2) = \widetilde{\Omega}(n^\alpha)$.

(i) *Let P be an n-point set in $\mathbb{R}^3$ such that every point in P is both $(---)$-maximal and $(+--)$-maximal. Then we can find an independent set of $\mathrm{DG}(P, \mathcal{B}^{3,4})$ of size $\widetilde{\Omega}(n^\alpha)$.*

(ii) *Let P be an n-point set in $\mathbb{R}^3$ such that every point in P is $(---)$-maximal. Then we can find an independent set of $\mathrm{DG}(P, \mathcal{B}^{3,4})$ of size $\widetilde{\Omega}(n^\alpha)$.*

(iii) $\mathrm{DG}\text{-}\mathrm{INDEP}(n, \mathcal{B}^{3,4}) = \widetilde{\Omega}(n^\alpha)$.

Proof:

(i) Consider the yz-projection $\widehat{P}$ of P. All points in $\widehat{P}$ are $(--)$-maximal, i.e., they form a descending staircase. Thus, for any value z, the range $\mathbb{R} \times (-\infty, z]$ contains the same points of $\widehat{P}$ as the range $[y, \infty) \times \mathbb{R}$ for some y. Consequently, the 3-dimensional 4-sided box $[x, x'] \times (-\infty, y'] \times (-\infty, z]$ contains the same points of P as the range $[x, x'] \times [y, y'] \times \mathbb{R}$. We can find a DG-independent set of P w.r.t. $\mathcal{B}^{3,4}$ by finding a DG-independent set of the xy-projection of P w.r.t. $\mathcal{B}^2$.

(ii) We describe a simple divide-and-conquer algorithm to color $\mathrm{DG}(P, \mathcal{B}^{3,4})$ with $\widetilde{O}(n^{1-\alpha})$ colors. Use a vertical line to divide P into a left point set P_L and a right point set P_R, each of size $n/2$. Mark the $(+--)$-maximal points of P_L. Color the marked points by applying (i), which requires a palette of $\widetilde{O}(n^{1-\alpha})$ colors. Recursively color the unmarked points in P_L and recursively color the points in P_R, using a common new palette. The number of colors satisfies the recurrence $G(n) = G(n/2) + \widetilde{O}(n^{1-\alpha})$, which gives $\widetilde{O}(n^{1-\alpha})$ as desired.

To prove correctness of this coloring, consider a 4-sided range s in $\mathcal{B}^{3,4}$ that contains exactly two points p, q of P. If both p and q are marked, they must have different colors. If one is marked and the other is unmarked, they obviously have different colors. So, assume both are unmarked. If $p, q \in P_L$ or $p, q \in P_R$, they would have different colors by induction. If $p \in P_L$ and $q \in P_R$, then s must contain a $(+--)$-maximal point of P_L; this point must be different from p and q since p is unmarked—a contradiction.

(iii) We describe a similar simple divide-and-conquer algorithm to color $\mathrm{DG}(P, \mathcal{B}^{3,4})$ with $O(n^{1-\alpha})$ colors. Use a vertical line to divide P into a left point set P_L and a right point set P_R, each of size $n/2$. Mark the $(---)$-maximal points of P_R. Color the marked points by applying (ii), which requires a palette of $\widetilde{O}(n^{1-\alpha})$ colors. Recursively color the points in P_L and recursively color the unmarked points in P_R, using a common new palette. The number of colors satisfies the recurrence $G'(n) = G'(n/2) + \widetilde{O}(n^{1-\alpha})$, which gives $\widetilde{O}(n^{1-\alpha})$ again. Correctness follows by the same argument as in (ii). $\qquad\square$

Lemma 3.2 addresses only the $k = 2$ case. We next show that the $k > 2$ case can be directly reduced to the $k = 2$ case for 3-dimensional 4-sided boxes, which implies the same for 2-dimensional rectangles as well. The key is the following proposition. Essentially the same proof as below was given by Ene, Har-Peled, and Raichel [14] in the special case of 2-dimensional rectangles.

Proposition 3.3 *Let P be any n-point set in $\mathbb{R}^3$ and k be a given value. Then we can find a subset Q with $\Omega(n/(k \log n))$ points such that every range $s \in \mathcal{B}^{3,4}$ with $|P \cap s| \leq k$ contains at most 2 points of Q.*

Proof: We describe a divide-and-conquer algorithm to construct a graph $G(P)$ over vertex set P: Use a vertical line to divide P into a left point set P_L and a right point set P_R, each of size $n/2$. Recursively add edges from $G(P_L)$ and $G(P_R)$. Also add edges from the order-k Delaunay graph of P_L w.r.t. $(+--)$-orthants, and the order-k Delaunay graph of P_R w.r.t. $(---)$-orthants. Note that these two Delaunay graphs have $O(nk)$ edges by Proposition 3.1(i). The total number of edges thus satisfies $E(n) = 2E(n/2) + O(nk)$, which gives $E(n) = O(nk \log n)$.

We return an independent set Q of $G(P)$. Since the average degree in $G(P)$ is $O(k \log n)$, we can find such a set of size $\Omega(n/(k \log n))$ by Turán's theorem.

To prove that Q satisfies the desired property, first associate each node of the recursion tree with a slab, where the root slab is the entire plane and the slab of a node is partitioned into the slabs at the two children using the vertical dividing line at the node. Consider a 4-sided range s in $\mathcal{B}^{3,4}$ with $|P \cap s| \leq k$. Find the node v in the recursion tree where s is contained in v's slab τ but not in the left or right child's slab τ_L or τ_R. Since s is equivalent to a $(+--)$-orthant when restricted inside τ_L, and to a $(---)$-orthant when restricted inside τ_R, we know that $|Q \cap s \cap \tau_L| \leq 1$ and $|Q \cap s \cap \tau_R| \leq 1$. Thus, $|Q \cap s| \leq 2$. $\qquad\square$

Corollary 3.4 *Suppose that $\mathrm{DG\text{-}INDEP}(n, \mathcal{B}^2) = \widetilde{\Omega}(n^\alpha)$. Then*

(i) $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^{3,4}) = \widetilde{\Omega}([n/k]^\alpha)$;

(ii) $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^2) = \widetilde{\Omega}([n/k]^\alpha)$.

Proof: To prove (i), we first apply Proposition 3.3 to obtain a subset Q of the given n-point set and then apply Lemma 3.2(iii) to output a DG-independent set of Q of size $\widetilde{\Omega}(|Q|^\alpha) = \widetilde{\Omega}([n/(k \log n)]^\alpha)$.

(ii) is a special case of (i), since after the transformation $(a, b) \mapsto (a, b, -b)$, a 2-dimensional rectangle $[x', x] \times [y', y]$ corresponds to a 3-dimensional 4-sided box $[x', x] \times (-\infty, y] \times (-\infty, -y']$. □

Remarks. Note that since 3-dimensional 4-sided boxes are at least as expressive as 2-dimensional rectangles, as the above transformation illustrates, the bound in Lemma 3.2(iii) is essentially tight.

Proposition 3.3 prompts a more general definition: Given a set P of elements in a range space and $j < k$, define a $\mathrm{DG}_{k,j}$-*independent set* of P to be a subset Q of P such that for every range s, $|P \cap s| \le k$ implies $|Q \cap s| \le j$. Our previous definition corresponds to the $j = 1$ case, whereas Proposition 3.3 considers $j = 2$. A related generalization of CF-coloring was considered in Har-Peled and Smorodinksy's paper [19], where it was shown that bounds improve sharply as j increases for any range spaces with bounded VC dimension. Combinatorial bounds on the $\mathrm{DG}_{k,j}$-independent set problem can be applied to the analysis of approximation algorithms for a generalization of discrete independent set where the "at most one" constraint in Problem 1.7 is replaced by "at most j".

3.2 Points w.r.t. Boxes in Higher Dimensions (Proof of Theorem 1.6(i))

The case of higher-dimensional boxes can be handled by repeated application of the Erdős–Szekeres theorem, as shown by Har-Peled and Smorodinsky [19]. We notice that the bound can be slightly improved using our new result for 2-dimensional rectangles as the base case:

Lemma 3.5 *If* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^{d-1}) = \widetilde{\Omega}(n^\beta/k^c)$, *then* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^d) = \widetilde{\Omega}(n^{\beta/2}/k^c)$.

Proof: Let P be a sequence of n points in $\mathbb{R}^d$ sorted according to the first coordinate. By the Erdős–Szekeres theorem [15], there exists a subsequence Q of $\Omega(\sqrt{n})$ points that is monotone increasing or decreasing in the last coordinate.

For any x_d and x'_d, the range $\mathbb{R}^{d-1} \times [x_d, x'_d]$ contains the same points of Q as the range $[x_1, x'_1] \times \mathbb{R}^{d-1}$ for some x_1 and x'_1. Thus, the d-dimensional box $[x''_1, x'''_1] \times [x_2, x'_2] \times \cdots \times [x_d, x'_d]$ contains the same points of Q as the box $[\max\{x_1, x''_1\}, \min\{x'_1, x'''_1\}] \times [x_2, x'_2] \times \cdots \times [x_{d-1}, x'_{d-1}] \times \mathbb{R}$. We can find a DG_k-independent set of Q (and thus P) w.r.t. $\mathcal{B}^d$ by finding a DG_k-independent set w.r.t. $\mathcal{B}^{d-1}$ of the projection of Q along the last coordinate. This independent set has size $\widetilde{\Omega}(|Q|^\beta/k^c) = \widetilde{\Omega}(n^{\beta/2}/k^c)$.
□

By repeated applications of Lemma 3.5 with Corollary 3.4(ii) as the base case, we obtain $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^d) = \widetilde{\Omega}(n^{\alpha/2^{d-2}}/k)$. This proves Theorem 1.6(i) except for a weaker exponent in the k factor, which can be fixed by the following general reduction:

Proposition 3.6 *For a range space $\mathcal{S}$ with a constant VC dimension,* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{S}) = \widetilde{\Omega}(n^\beta/k^{O(1)})$ *implies* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{S}) = \widetilde{\Omega}([n/k]^\beta)$.

Proof: The claim follows from a standard random sampling argument, or as a special case of Theorem 1.8. For the latter, given a set P of n elements, just let S consist of all combinatorial different ranges $s \in \mathcal{S}$ and $|P \cap s| \le k$. By the bounded VC-dimension assumption (Sauer's lemma) [23], S has polynomial size. In the LP for the unweighted setting, OPT equals the size of the maximum independent set of $\mathrm{DG}_k(P, \mathcal{S})$, and opt $\ge n/k$ since we can set all x_p's to $1/k$. Theorem 1.8 implies that $\mathrm{OPT} \ge \widetilde{\Omega}([n/k]^\beta)$. □

Remark. It would be nice if Ajwani et al.'s grid-based method could be directly extended to handle higher-dimensional boxes to yield better bounds, but such an extension does not appear obvious.

3.3 Points w.r.t. Orthants in Higher Dimensions (Proof of Theorem 1.6(ii))

For higher-dimensional orthants, better results can be obtained by applying the Erdős–Szekeres theorem in a slightly more interesting way:

Lemma 3.7 *If* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^{d-1,d}) = \widetilde{\Omega}(n^\beta/k^c)$ *and* $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{O}^{d-1}) = \widetilde{\Omega}(n^\gamma/k^c)$, *then*

(i) $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^{d,d+1}) = \widetilde{\Omega}(n^{\beta/2}/k^c)$, *and*

(ii) $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{O}^d) = \widetilde{\Omega}(n^{\beta\oplus\gamma}/k^c)$ *where* $\beta \oplus \gamma = 1/(1/\beta + 1/\gamma)$.

Proof:

(i) The proof of this part is almost identical to the proof of Lemma 3.5. We just observe that the d-dimensional $(d+1)$-sided box $[x_1'', x_1'''] \times (-\infty, x_2'] \times \cdots \times (-\infty, x_d']$ contains the same points of Q as the box $[\max\{x_1, x_1''\}, \min\{x_1', x_1'''\}] \times (-\infty, x_2'] \times \cdots \times (-\infty, x_{d-1}'] \times \mathbb{R}$ for some x_1 and x_1'. The latter box is in $\mathcal{B}^{d-1,d}$ after projection along the last coordinate.

(ii) The proof of this part is also similar but requires the use of a parameter r. Let P be a sequence of n points in $\mathbb{R}^d$ sorted in increasing order according to the first coordinate. By the Erdős–Szekeres theorem, there exists either (a) a subsequence Q of $\Omega(r)$ points that is monotone increasing in the last coordinate, or (b) a subsequence Q of $\Omega(n/r)$ points that is monotone decreasing in the last coordinate.

In Case (a), for any value x_d, the range $\mathbb{R}^{d-1} \times (-\infty, x_d]$ contains the same points of Q as the range $(-\infty, x_1] \times \mathbb{R}^{d-1}$ for some x_1. Thus, the d-dimensional orthant $(-\infty, x_1'] \times (-\infty, x_2] \times \cdots \times (-\infty, x_d]$ contains the same points of Q as the range $(-\infty, \min\{x_1, x_1'\}] \times (-\infty, x_2] \times \cdots \times (-\infty, x_{d-1}] \times \mathbb{R}$. We can find a DG_k-independent set of Q (and thus P) w.r.t. $\mathcal{O}^d$ by finding a DG_k-independent set w.r.t. $\mathcal{O}^{d-1}$ of the projection of Q along the last coordinate.

In Case (a), for any value x_d, the range $\mathbb{R}^{d-1} \times (-\infty, x_d]$ contains the same points of Q as the range $[x_1, \infty) \times \mathbb{R}^{d-1}$ for some x_1. Thus, the d-dimensional orthant $(-\infty, x_1'] \times (-\infty, x_2] \times \cdots \times (-\infty, x_d]$ contains the same points of Q as the range $[x_1, x_1'] \times (-\infty, x_2] \times \cdots \times (-\infty, x_{d-1}] \times \mathbb{R}$. We can find a DG_k-independent set of Q (and thus P) w.r.t. $\mathcal{O}^d$ by finding a DG_k-independent set w.r.t. $\mathcal{B}^{d-1,d}$ of the projection of Q along the last coordinate.

In any case, we can find an independent set of $\mathrm{DG}_k(P, \mathcal{O}^d)$ of size

$$\widetilde{\Omega}(\min\{r^\gamma, \ (n/r)^\beta\}/k^c).$$

By setting $r = n^{\beta/(\beta+\gamma)}$ to balance the two terms, we obtain a bound of $\widetilde{\Omega}(n^{\beta\oplus\gamma}/k^c)$. $\qquad\square$

By repeated applications of Lemma 3.7(i) with Corollary 3.4(i) as the base case, we have $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{B}^{d-1,d}) = \widetilde{\Omega}(n^{\alpha/2^{d-4}}/k)$.

By repeated applications of Lemma 3.7(ii) with Proposition 3.1(ii) as the base case, we then have $\mathrm{DG\text{-}INDEP}_k(n, \mathcal{O}^d) = \widetilde{\Omega}(n^{\gamma_d}/k)$ with $\gamma_3 = 1$ and $\gamma_d = \alpha/2^{d-4} \oplus \gamma_{d-1}$, i.e.,

$$\gamma_d = \alpha/2^{d-4} \oplus \cdots \oplus \alpha/2 \oplus \alpha \oplus 1 = \frac{1}{2^{d-4}/\alpha + \cdots + 2/\alpha + 1/\alpha + 1} = \frac{1}{(2^{d-3}-1)/\alpha + 1}.$$

The exponents in the k factors can be fixed again by Proposition 3.6. This proves Theorem 1.6(ii).

15

Remark. As mentioned in the introduction, the case of boxes w.r.t. points in $\mathbb{R}^d$ can be reduced to the case of points w.r.t. orthants in $\mathbb{R}^{2d}$, i.e., DG-INDEP$_k(n, (\mathcal{B}^d)^*) = \widetilde{\Omega}(n^{\alpha/(2^{2d-3}-1+\alpha)}/k)$. It would be interesting to find better bounds by a more direct approach for this dual problem. (Note that for $d = 3$, one can get DG-INDEP$_k(n, (\mathcal{B}^3)^*) = \Omega(n/(k\,\mathrm{polylog}\,n))$ by using divide-and-conquer to reduce the number of sides, and Proposition 3.1 for the base case; this is described in Ene, Har-Peled, and Raichel's paper [14].)

4 Discrete Independent Set (Proof of Theorem 1.8)

In this last section, we present an approximation algorithm for discrete independent set by rounding the LP from Theorem 1.8. Let $\{x_p \mid p \in P\}$ be the solution to the LP. Draw a random sample R of the elements where element p is chosen independently with probability x_p. For a fixed range s, consider the random variable $|R \cap s|$, which has expected value $\mu_s = \sum_{p \in s} x_p \leq 1$ and is a sum of independent 0-1 random variables. By the Chernoff bound,

$$\Pr[|R \cap s| > (1 + \delta)\mu_s] \ \leq \ \left[\frac{e^\delta}{(1+\delta)^{1+\delta}}\right]^{\mu_s}$$

for any $\delta > 0$ (possibly large). Choosing[4] $\delta = \Theta((\log m)/\mu_s)$, we have $|R \cap s| = O(\log m)$ with high probability (e.g., probability $1 - o(1/m)$).

Next consider the total weight $w(R)$ of R, which has expected value $\mu = \sum_{p \in s} w_p x_p = \mathrm{opt}$ and variance $\sigma^2 = \sum_{p \in s} w_p^2 (x_p - x_p^2) \leq W \cdot \mathrm{opt}$, where W denotes the maximum weight in P. We may assume that $W \leq \mathrm{opt}/c$ for any fixed constant c, for otherwise we can just return the singleton with weight W and be done. So, $\sigma \leq \mathrm{opt}/\sqrt{c}$. By Chebyshev's inequality,

$$\Pr[|w(R) - \mu| > t\sigma] \ \leq \ 1/t^2$$

for any $t > 0$. Choosing constants c and t appropriately, we have $|w(R)| \geq \Omega(\mathrm{opt})$ with probability $\Omega(1)$. So, by repeating for an expected $O(1)$ number of trials, we can find a subset R satisfying both (a) $|R \cap s| \leq O(\log m)$ for all $s \in S$ at the same time, and (b) $w(R) \geq \Omega(\mathrm{opt})$.

In the unweighted case, we return a DG$_{c' \log m}$-independent set Q of R of size $\widetilde{\Omega}(|R|^\alpha / \log^{O(1)} m)$, which is $\widetilde{\Omega}(\mathrm{opt}^\alpha)$ by (b). We know that Q is a feasible solution for a sufficiently large constant c' by (a) and the definition of DG$_{c' \log m}$-independence.

In the weighted case, we compute a DG$_{c' \log m}$-coloring of R with $\widetilde{O}(|R|^{1-\alpha} \log^{O(1)} m) = \widetilde{O}(n^\alpha)$ colors, and return a color class Q of the largest weight. Then $w(Q) \geq \widetilde{\Omega}(w(R)/n^{1-\alpha}) \geq \widetilde{\Omega}(\mathrm{opt}/n^{1-\alpha})$ by (b). Again, Q is a feasible solution by (a). This completes the proof of Theorem 1.8.

Remark. It is possible to make small improvements in the logarithmic factors by experimenting with a different choice of probability values, e.g., selecting element p with probability $\min\{bx_p, 1\}$ for some parameter $b \geq 1$.

---

[4]Actually $\delta = \Theta((\log m/\log\log m)/\mu_s)$ works too, as noted in [9].

References

[1] D. Ajwani, K. M. Elbassioni, S. Govindarajan, and S. Ray. Conflict-free coloring for rectangle ranges using $O(n^{.382})$ colors. In *19th ACM Sympos. Parallelism in Algorithms and Architectures*, pages 181–187, 2007. *Discrete Comput. Geom.*, to appear.

[2] N. Alon, M. Krivelevich, and B. Sudakov. Coloring graphs with sparse neighborhoods. *J. Combin. Theory Ser. B*, 77:73–82, 1999.

[3] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley-Interscience, 2nd ed., 2000.

[4] G. Aloupis, J. Cardinal, S. Collette, S. Langerman, and S. Smorodinsky. Coloring geometric range spaces. *Discrete Comput. Geom.*, 41:348–362, 2009.

[5] H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:463–479, 1995.

[6] P. Chalermsook and J. Chuzhoy. Maximum independent set of rectangles. In *Proc. 20th ACM–SIAM Sympos. Discrete Algorithms*, pages 892–901, 2009.

[7] T. M. Chan. Polynomial-time approximation schemes for packing and piercing fat objects. *J. Algorithms*, 46(2):178–189, 2003.

[8] T. M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proc. 23rd ACM–SIAM Sympos. Discrete Algorithms*, pages 1576–1585, 2012.

[9] T. M. Chan and S. Har-Peled. Approximation algorithms for maximum independent set of pseudo-disks. *Discrete Comput. Geom.*, to appear. (Preliminary version in *Proc. 25th ACM Sympos. Comput. Geom.*, pages 333–340, 2009.)

[10] X. Chen, J. Pach, M. Szegedy, and G. Tardos. Delaunay graphs of point sets in the plane with respect to axis-parallel rectangles. *Random Struct. Algorithms*, 34:11–23, 2009.

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

[12] K. L. Clarkson and K. R. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37:43–58, 2007.

[13] K. Elbassioni and N. H. Mustafa. Conflict-free colorings of rectangular ranges. In *Proc. 23rd Sympos. Theoret. Aspects Comput. Sci.*, Lect. Notes Comput. Sci., vol. 3884, Springer-Verlag, pages 254–263, 2006.

[14] A. Ene, S. Har-Peled, and B. Raichel. Geometric packing under non-uniform constraints. To appear in *Proc. 28th ACM Sympos. Comput. Geom.*, 2012.

[15] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[16] G. Even, Z. Lotker, D. Ron, and S. Smorodinsky. Conflict-free colorings of simple geometric regions with applications to frequency assignment in cellular networks. *SIAM J. Comput.*, 33:94–136, 2003.

[17] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Inform. Process. Lett.*, 95:358–362, 2005.

[18] J. Fox and J. Pach. Computing the independence number of intersection graphs. In *Proc. 22nd ACM–SIAM Sympos. Discrete Algorithms*, pages 1161–1165, 2011.

[19] S. Har-Peled and S. Smorodinsky. Conflict-free coloring of points and simple regions in the plane. *Discrete Comput. Geom.*, 34:47–70, 2005.

[20] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. American Statistical Assoc.*, 58:13–30, 1963.

[21] P. M. Long. Using the pseudo-dimension to analyze approximation algorithms for integer programming. In *Proc. 7th Algorithms Data Struct. Sympos.*, Lect. Notes Comput. Sci., vol. 2125, Springer-Verlag, pages 26–37, 2001.

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

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

[24] J. Pach and G. Tóth. Conflict free colorings. In *Discrete and Computational Geometry: The Goodman–Pollack Festschrift*, Springer-Verlag, Berlin, pages 665–671, 2003.

[25] S. Smorodinsky. Conflict-free coloring and its applications. http://arxiv.org/abs/1005.3616, 2010.

Appendix

A   Proofs of Propositions 1.4 and 3.1

Proof of Proposition 1.4:   For the only-if direction of (i), we can color a set P of elements as follows: find a DG_k-independent set Q of P of size $\widetilde{\Omega}(|P|^\alpha/k^c)$, give the elements of Q a common new color, remove Q from P, and repeat.

While the size of P remains above $n/2$, the number of iterations is at most $\widetilde{O}(n^{1-\alpha}k^c)$. The total number of iterations until P becomes empty is thus $\widetilde{O}((n^{1-\alpha}+(n/2)^{1-\alpha}+\cdots)k^c) = \widetilde{O}(n^{1-\alpha}k^c)$. This bounds the total number of colors used.

For the only-if direction of (ii), we use exactly the same algorithm (with $k = 2$). To see why the result is a CF-coloring, consider a range s containing at least two elements of P. If s contains an element of $P - Q$, then s contains a unique color by induction. On the other hand, if s contains only elements of Q, then by the monotonicity property, some subrange s' contains exactly two elements of Q, which contradicts the DG-independence of Q.

For the converse to (i), in a DG_k-coloring with $\widetilde{O}(n^{1-\alpha}k^c)$ colors, the largest color class has size $\widetilde{\Omega}(n^\alpha/k^c)$ and is a DG_k-independent set. Similarly, for the converse to (ii), the largest color class in a CF-coloring is clearly a DG-independent set.

Proof of Proposition 3.1:   For each point $p \in P$, let o_p denote the $(---)$-orthant with corner at p. Draw a random sample R where each point $p \in P$ is chosen independently with probability $1/2k$. Let $VD(R)$ be the vertical decomposition of the complement of the union $U(R)$ of $\{o_r \mid r \in R\}$. (The collection of cells $VD(R)$ may be regarded as a *shallow cutting.*) Since the union complexity of orthants in $\mathbb{R}^3$ is linear by Euler's formula (as $U(R)$ is an orthogonal polyhedron), $VD(R)$ consists of $O(|R|)$ cells. For each cell $\Delta \in VD(R)$, define its *conflict list* P_Δ to consist of all points $p \in P$ with o_p intersecting Δ. Let G_Δ denote the set of all pairs in P_Δ and $G = \bigcup_{\Delta \in VD(R)} G_\Delta$. By Clarkson and Shor's technique [11],

$$E[|G|] \;=\; E\left[\sum_{\Delta \in VD(R)} |P_\Delta|^2\right] \;=\; O((n/k) \cdot k^2) \;=\; O(nk).$$

For a fixed edge $p_1p_2 \in DG_k(P, \mathcal{O}^3)$, let $q \in P$ be a point that $(+++)$-dominates at most k points in P and $(+++)$-dominates p_1 and p_2. Then q is not covered by the cells in $VD(R)$ iff q lies inside

18

$U(R)$, iff at least one point $(+++)$-dominated by q is chosen in R—this happens with probability at most $(1/2k)k = 1/2$. If q lies inside a cell $\Delta \in \text{VD}(R)$, then we know that $p_1 p_2 \in G_\Delta \subseteq G$. It follows that $\Pr[p_1 p_2 \in G] \geq 1/2$. Thus,

$$E[|G|] \ \geq \ (1/2)|\text{DG}_k(P, \mathcal{O}^3)|,$$

implying $|\text{DG}_k(P, \mathcal{O}^3)| = O(nk)$. This proves (i).

Part (ii) follows from (i) by Turán's theorem [3], since the average degree in $\text{DG}_k(P, \mathcal{O}^3)$ is $O(k)$.

Remark. Since the DG_k graph family is hereditary, we can alternatively use the standard greedy algorithm for independent set (select a vertex with the largest degree, remove, and repeat) instead of Turán's theorem.

B   Verification of 0.632 in Maple

The following script in Maple outputs a vector $\geq (0.6320002, 0, 0)$:

```
g := proc(a1,b1,c1,a2,b2,c2)
    if (1-(a1+b1-c1)*a2+b2-c2 < 0) then printf("undefined!\n") end if;
    (1,0,0) + (a1*a2-1, b1*a2, c2) / (2-(a1+b1-c1)*a2+b2-c2) end proc;
s := proc(a,b,c) a,c,b end proc;
u := (1,-1,0); v := (0.618,0,0);
w0 := s(u); w1 := g(u,w0); w2 := g(u,w1); w3 := g(u,w2); w4 := g(u,w3); w5 := g(u,w4);
z0 := s(w1); z1 := g(u,z0); z2 := g(u,z1); z3 := g(u,z2); z4 := g(u,z3); z5 := g(u,z4);
for i from 1 to 11 do v := \
g(s(g(g(g(g(g(g(z5,g(z5,g(z4,g(g(u,g(u,g(u,s(g(w1,w0)))))),s(g(g(w2,s(z1)),s(z2)))))))),g(g(z5,g(z4\
,g(g(u,g(u,g(w3,s(g(w1,w0)))))),g(g(u,g(u,g(w3,s(g(w1,s(w2)))))),s(g(g(w2,g(w1,s(w2))),s(g(u,s(g(w1,w\
0))))))))))))),g(g(g(u,g(u,g(u,g(w3,s(g(w1,w0)))))),g(g(u,g(u,g(w3,s(g(w1,s(w2)))))),g(g(w3,g(w3,s\
(g(z1,s(w2))))),s(g(g(w2,g(z1,s(w2))),s(g(u,s(g(w1,s(w2))))))))))),g(g(g(u,g(w3,g(w3,s(g(z1,s(w2\
)))))),g(g(w4,g(z3,g(z2,s(g(w2,s(z1))))),s(g(g(w3,g(w2,s(z1))),s(g(u,((s(g(w1,w0)))))))))),s(g\
(g(g(w3,g(w2,g(z1,s(w2))),g(g(z2,g(z1,s(w2))),s(g(w3,s(g(z1,s(w2)))))),s(g(g(u,g(w3,s(g(z1,s(\
w2)))),s(g(g(w2,g(z1,s(w2))),s(g(u,s(g(w1,s(w2))))))))))))),g(g(g(z5,g(g(u,g(u,g(u,g(w3,s(g(\
w1,w0)))))))),g(g(u,g(u,g(w3,s(g(w1,s(w2)))))),g(g(w4,g(w3,g(z2,s(z1)))),s(g(g(w2,g(z1,s(w2))),s(\
g(u,s(g(w1,s(w2)))))))))))),g(g(g(u,g(w4,g(w4,g(w3,s(g(w1,s(w2)))))),g(g(w4,g(w4,g(w3,g(w2,s(z1\
))))),g(g(w4,g(z3,g(z2,s(g(u,s(z1)))))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1)))))))))),g(g(\
g(w4,g(w4,g(w3,g(z2,s(z1))))),g(g(w4,g(z3,g(z2,s(g(w2,s(z1)))))),s(g(g(z2,s(g(z2,s(z1)))),s(g(z\
3,g(z2,s(g(u,s(z1)))))))))),s(g(g(g(z3,g(z2,s(g(w2,s(z1))))),g(g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2\
,s(z1)))))))),s(g(g(w4,g(z3,g(z2,s(g(u,s(z1)))))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1)))))))))\
))))))),s(g(g(g(g(g(u,g(u,g(w2,z0))),g(g(w3,g(w2,s(z1))),g(g(w2,s(g(w2,z0))),s(g(w3,g(w2,s(z1))\
)))))),g(g(g(w3,g(w2,s(g(w2,z0)))),g(g(z2,g(w2,s(g(u,((w1)))))),s(g(u,((s(g(w1,w0)))))))),s(g(g(\
u,g(u,((g(w2,z0)))))),s(g(g(w2,s(g(w2,z0))),s(g(w3,g(w2,s(z1))))))),s(g(g(u,g(u,g(w3,g(w2,s\
(z1))))),s(g(g(z2,s(g(w2,z0))),s(g(w3,g(w2,s(z1))))))),s(g(g(w3,g(\
w2,s(z1))),g(g(w2,s(g(w2,z0))),s(g(w3,g(w2,s(z1))))))),s(g(g(u,g(w3,g(w2,s(z1)))),s(g(g(w2,s(z1)\
),s(g(u,g(w2,z0)))))))))))))),s(g(g(g(u,g(u,g(u,g(u,s(g(w1,w0)))))),g(g(u,g(u,g(u,s(g(w1,s(w2)))))\
))),g(g(w4,g(w3,g(z2,s(g(w2,z0)))))),s(g(g(g(w3,g(w2,s(z1))),s(g(u,((z0)))))))),g(g(g(w4,g(w4,g(w3\
,g(z2,s(z1))))),g(g(w4,g(z3,g(z2,s(g(u,s(z1)))))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1))))\
))))),s(g(g(g(w3,g(z2,s(g(w2,z0)))),g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1)))))),s(g(g(w4,g(w3,\
g(z2,s(z1)))),s(g(g(w2,s(g(w2,z0))),s(g(w3,g(w2,s(z1)))))))))))))),s(g(g(g(g(g(z3,g(g(u,g(u,\
g(w2,z0))),g(g(w3,g(w2,s(z1))),s(g(z2,s(g(w2,z0)))))),g(g(g(w3,g(w3,g(w2,s(z1)))),g(g(w3,g(w2,\
g(z1,s(w2))))),s(g(g(u,g(w2,z0)),s(g(w2,s(z1)))))),g(g(g(w3,g(w2,g(z1,s(w2)))),g(g(z2,g(z1,s(w2\
)))),s(g(w3,s(g(z1,s(w2))))))),s(g(w3,((s(g(w2,z0),s(z2))))))),g(g(g(w3,g(w2,g(z1,s(w2\
))))),g(g(z2,g(z1,s(w2))),s(g(u,((((z1))))))),s(g(g(u,g(w3,((g(w2,s(z1))))),s(g(g(z2,s(g\
(w2,s(z1)))),s(g(w3,g(z2,s(z1))))))))))),s(g(g(g(w4,g(w4,g(w3,g(z2,s(z1))))),g(g(w4,g(w3,g(w2,s(\
z1)))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1)))))))),s(g(g(g(w3,g(z2,s(g(w2,z0)))),g(g(z2,s\
(g(w2,s(z1)))),s(g(w3,g(z2,s(z1)))))),s(g(g(u,g(w3,g(w2,s(z1)))),s(g(g(z2,s(g(w2,z0))),s(g(w3,g\
(w2,s(z1))))))))))))))),g(g(g(g(g(w3,g(w3,g(z2,s(z1)))),g(g(z3,g(z2,s(g(w2,s(z1)))))),g(g(z2,s(g(\
w2,s(z1)))),s(g(w3,g(z2,s(z1)))))))),g(g(g(z3,g(z2,s(g(w2,s(z1))))),g(g(z2,g(w2,s(g(u,((w1))))))\
),s(g(w3,((s(g(z1,s(w2)))))))))),s(g(g(g(w3,g(w3,((g(w2,s(z1)))))),s(g(g(z2,s(g(z2,s(z1))))),s(g(w3,\
```

```
g(z2,s(z1)))))))))),s(g(g(g(g(w4,g(w4,g(w3,g(z2,s(z1))))),g(g(w4,g(z3,g(z2,s(g(w1,s(z1)))))),s(g(\
g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2,s(z1)))))))),s(g(g(g(z3,g(z2,s(g(w2,s(z1))))),g(g(z2,s(g(w2,s(\
z1)))),s(g(w3,g(z2,s(z1))))))),s(g(g(w4,g(w3,g(z2,s(z1)))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(w2,\
s(z1))))))))))))))))),s(g(g(g(u,g(w4,g(w4,g(w3,s(g(z1,s(w2))))))),g(g(w4,g(w3,g(w3,s(g(z1,s(w2)))))\
),g(g(w4,g(w3,s(g(z1,s(z2))))),s(g(g(z2,g(z1,s(w2))),s(g(w3,s(g(z1,s(w2))))))))))),g(g(g(w4,g(w4\
,g(w3,s(g(z1,s(w2))))))),g(g(w4,g(z3,g(z2,s(g(w2,s(z1)))))),s(g(g(z2,g(z1,s(w2))),s(g(w3,((g(w2,\
s(z1)))))))))),s(g(g(g(z3,g(z2,s(g(w2,s(z1))))),g(g(z2,g(w2,s(g(u,((w1)))))),s(g(w3,((s(g(z1,s(\
w2)))))))))),s(g(g(w3,g(w3,((g(w2,s(z1)))))),s(g(g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2,s(z1)))))))))))\
)))),s(g(g(g(g(g(u,g(u,g(u,g(w3,g(w3,s(z1)))))),g(g(u,g(w4,g(w4,g(w3,s(g(w1,s(w2))))))),g(g(w4,g(\
(w4,g(w3,s(g(z1,s(w2))))))),g(g(w4,g(w3,g(z2,s(z1)))),s(g(g(z2,g(z1,s(w2))),s(g(w3,s(g(z1,s(w2))\
)))))))))))),g(g(g(u,g(w4,g(w3,g(w3,s(g(g(u,((w0))),s(w2)))))))),g(g(w4,g(w4,g(w3,g(w2,s(z1))))),g(\
g(w4,g(z3,g(z2,s(g(w2,s(z1))))))),s(g(g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2,s(z1)))))))))),g(g(g(z4,g(\
w4,g(w3,g(z2,s(g(u,s(z1))))))),g(g(w4,g(z3,g(z2,s(g(w2,s(z1)))))),s(g(g(g(u,g(u,s(z1))),s(g(z2,\
s(g(u,s(z1))))))),s(g(z3,g(z2,s(g(w2,s(z1))))))))))),s(g(g(g(z3,g(z2,s(g(z2,s(z1))))),g(g(z2,s(g(\
z2,s(z1)))),s(g(w3,g(z2,s(z1)))))))),s(g(g(w4,g(w3,g(w2,s(z1)))),s(g(g(z2,s(g(z2,s(z1)))),s(g(w3,\
g(z2,s(z1))))))))))))))))),s(g(g(g(g(g(w3,g(w3,g(w2,s(z1)))),g(g(w3,g(z2,s(g(w2,z0)))),g(g(z2,s(g(w\
2,s(z1)))),s(g(w3,g(z2,s(z1)))))))),g(g(w3,g(z2,s(g(w2,s(z1)))),g(g(z2,g(w2,s(g(u,((w1)))))),\
s(g(u,((s(g(w1,s(w2)))))))))),s(g(g(w3,g(w3,((g(w2,s(z1)))))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(\
z2,s(z1)))))))))),s(g(g(g(w4,g(w4,g(w3,g(z2,s(z1))))),g(g(w4,g(w3,g(w2,s(z1)))),s(g(g(z2,s(g(w2\
,s(z1)))),s(g(w3,g(z2,s(z1))))))))),s(g(g(g(w3,g(z2,s(g(w2,z0)))),g(g(z2,s(g(w2,s(z1)))),s(g(w3,\
g(z2,s(z1)))))))),s(g(g(w4,g(w3,g(z2,s(z1)))),s(g(g(w2,s(g(w2,z0))),s(g(w3,g(w2,s(z1))))))))))))))\
,s(g(g(g(u,g(u,g(w3,g(w3,s(g(g(u,((u))),s(w2)))))))),g(g(u,g(w3,g(w3,s(g(z1,s(w2)))))),g(g(w4,g(\
z3,g(z2,s(g(w2,s(z1))))))),s(g(g(w3,g(w2,s(z1))),s(g(u,((s(g(w1,w0)))))))))))),g(g(g(w4,g(w4,g(w3\
,g(z2,s(z1))))),g(g(w4,g(z3,g(z2,s(g(w2,s(z1))))))),s(g(g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2,s(z1)))\
))))))),s(g(g(g(z3,g(z2,s(g(w2,s(z1))))),g(g(z2,s(g(z2,s(z1)))),s(g(w3,g(z2,s(z1)))))),s(g(g(w4,g\
(z3,g(z2,s(g(u,s(z1)))))),s(g(g(z2,s(g(w2,s(z1)))),s(g(w3,g(z2,s(z1))))))))))))))))))))))),v)),v);\
end do;
```

Remark. In the above expression, **v** only appears twice. Thus, most of the intermediate vectors (α, β, γ) (except the very last two) actually have $\alpha = 1$. The **g** map in (9) simplifies in this case to

$$\mathbf{g}((\beta_1, \gamma_1), (\beta_2, \gamma_2)) \;=\; \tfrac{1}{1-(\beta_1-\gamma_1)+\beta_2-\gamma_2}(\beta_1, \gamma_1),$$

If $\beta_1 - \gamma_1 > \beta_2 - \gamma_2$, we make **g** undefined. It would be beneficial to study the region of all points in $\mathbb{R}^2$ that can be obtained by repeated application of this map **g** and the map $\mathbf{s}(\beta, \gamma) = (\gamma, \beta)$, starting with the initial point $(-1, 0)$. It is not clear, however, if this region can be described in closed form.