MORE COLOURFUL SIMPLICES

ANTOINE DEZA, TAMON STEPHEN, AND FENG XIE

Abstract. We show that any point in the convex hull of each of $(d+1)$ sets of $(d+1)$ points in general position in $\mathbb{R}^d$ is contained in at least $\lceil (d+1)^2/2 \rceil$ simplices with one vertex from each set. This improves the known lower bounds for all $d \geq 4$.

1. Introduction

A point $p \in R^d$ has *simplicial depth* k relative to a set S if it is contained in k closed simplices generated by $(d+1)$ sets of S. This was introduced by Liu [Liu90] as a statistical measure of how representative p is of S, and is a source of challenging problems in computational geometry – see for instance [FR05]. More generally, we consider *colourful simplicial depth*, where the single set S is replaced by $(d+1)$ sets, or colours, $\mathbf{S}_1, \dots, \mathbf{S}_{d+1}$, and the *colourful* simplices containing p are generated by taking one point from each set.

Assuming that the convex hulls of the $\mathbf{S}_i$'s contain p in their interior, Bárány's Colourful Carathéodory Theorem [Bár82] shows that p must be contained in some colourful simplex. We are interested in determining the minimum number of colourful simplices that can contain p for sets satisfying these conditions. That is, we would like to determine $\mu(d)$, the minimum number of colourful simplices drawn from $\mathbf{S}_1, \dots, \mathbf{S}_{d+1}$ that contain $p \in R^d$ given that $p \in \text{int}(\text{conv}(\mathbf{S}_i))$ for each i. Without loss of generality, we assume that the points in $\bigcup_i \mathbf{S}_i \cup \{p\}$ are in general position. Besides intrinsic appeal, $\mu(d)$ represents the minimum number of solutions to the colourful linear programming feasibility problem proposed in [BO97] and discussed in [DHST06].

The quantity $\mu(d)$ was investigated in [DHST06], where it is shown that $2d \leq \mu(d) \leq d^2+1$, that $\mu(d)$ is even for odd d, and that $\mu(2) = 5$. This paper also conjectures that $\mu(d) = d^2+1$ for all $d \geq 1$. Subsequently, [BM07] verified the conjecture for $d = 3$ and provided a lower bound of $\mu(d) \geq \max(3d, \left\lceil \frac{d(d+1)}{5} \right\rceil)$ for $d \geq 3$, while [ST08] independently provided a lower bound of $\mu(d) \geq \left\lfloor \frac{(d+2)^2}{4} \right\rfloor$. In this note we show:

Theorem 1: For $d \geq 1$, we have $\mu(d) \geq \lceil \frac{(d+1)^2}{2} \rceil$.

This strengthens the previously known lower bound for all $d \geq 4$.

2. Preliminaries

Without loss of generality we can take $p = \mathbf{0}$. The sets $\mathbf{S}_1, \dots, \mathbf{S}_{d+1}$ must each contain at least $(d+1)$ points for $\mathbf{0}$ to be in the interior of their convex hulls, and since we are minimizing we can assume they contain no additional points, i.e. that $|\mathbf{S}_i| = d+1$ for

2000 *Mathematics Subject Classification.* 52C45, 52A35.

Key words and phrases. Colourful simplicial depth, colourful Carathéodory theorem, discrete geometry.

each i. We assume that all points are distinct, so no point occurs in two $\mathbf{S}_i$'s, and $\mathbf{0}$ is not in any $\mathbf{S}_i$. We can scale the points of the $\mathbf{S}_i$'s so that they lie on the unit sphere $\mathbb{S}^{d-1}$: $\mathbf{0}$ is in a simplex after scaling if and only if it was in the simplex before scaling.

We call a set of points drawn from the $\mathbf{S}_i$'s *colourful* if it contains at most one point from each $\mathbf{S}_i$. We call a colourful set of d points which misses $\mathbf{S}_i$ an *$\widehat{i}$-transversal*. Note that $\widehat{i}$-transversals generate full dimensional pointed colourful cones; we will say that a transversal *spans* a point if the point is contained in the associated cone. A key observation is that colourful simplices containing $\mathbf{0}$ are generated whenever the antipode of a point of colour i is spanned by an $\widehat{i}$-transversal.

In particular, we look at the combinatorial *octahedra*, or cross polytopes, generated by pairs of disjoint $\widehat{i}$-transversals. We rely on the topological fact that every octahedron Ω either covers all of $\mathbb{S}^{d-1}$ with colourful cones, or, every point $x \in \mathbb{S}^{d-1}$ that is covered by colourful cones from Ω is covered by at least two distinct such cones. In the case where the points of Ω form an octahedron in the geometric sense, these correspond to the cases where $\mathbf{0}$ is inside and outside Ω respectively. For a proof, see for example the *Octahedron Lemma* of [BM07]. We remark that a given octahedron contains 2^d transversals, though we specify only two disjoint ones to generate it.

Our strategy for finding distinct colourful simplices is to begin with a transversal that generates at least one colourful simplex, and get further points from octahedra that include this transversal. We will break into cases based on the number of colourful simplices generated by the initial transversal and how many of the octahedra cover $\mathbb{S}^{d-1}$.

3. Proof of the Theorem 1

We know that at least one colourful simplex contains $\mathbf{0}$. Therefore we have an antipode of colour $(d+1)$ lying in the cone generated by a $\widehat{d+1}$-transversal T. Without loss of generality we can number the points of $\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_d$ so that point $(d+1)$ of S_i is included in T. The remaining points of the S_i's can be numbered arbitrarily. Let T_i be the set that contains the points numbered i from $\mathbf{S}_1, \mathbf{S}_2, \dots \mathbf{S}_{d+1}$. Then each T_i is a $\widehat{d+1}$-transversal and $T_{d+1} = T$. Further, the sets $T_1, T_2, \dots, T_{d+1}$ are pairwise disjoint. Let L be the set of antipodes of colour $(d+1)$ spanned by T_{d+1}, where $|L| = l > 0$.

3.1. **Points from d octahedra that share a transversal.** Now consider the d octahedra $\Omega_1, \Omega_2, \dots, \Omega_d$ given by pairing T_i with T_{d+1} for $i = 1, 2, \dots, d$. Except for the common transversal T_{d+1}, every $\widehat{d+1}$-transversal found among the Ω_i's is distinct. For each i, Ω_i may or may not cover all of $\mathbb{S}^{d-1}$. Suppose that b of the octahedra cover $\mathbb{S}^{d-1}$. There are $(d+1-l)$ antipodes of colour $(d+1)$ that are not spanned by T_{d+1}, and hence must be spanned by a different transversal from each of these octahedra. This gives us a total of $b(d+1-l)$ distinct simplices containing $\mathbf{0}$. Now there remain $(d-b)$ octahedra that do not span all of $\mathbb{S}^{d-1}$. By the Octahedron Lemma, each of the l antipodes spanned by T_{d+1} must also be spanned by a second transversal from the octahedron generated by T_{d+1} and T_i. So we find an additional $(d-b)l$ distinct simplices along with the l simplices generated by the antipodes with T_{d+1} itself. This brings us to a total of: $l + b(d+1-l) + (d-b)l = (d+1)(b+l) - 2bl$ distinct colourful simplices containing $\mathbf{0}$ through this simple argument.

3.2. **Choice of T_{d+1}.** In the above argument, T_{d+1} can be any $\widehat{d+1}$-transversal containing an antipode of colour $(d+1)$. In the construction of previous lower bounds, it was noted that if $\mathbf{depth}(\mathbf{0})$ is low, then there must be a portion of $\mathbb{S}^d$ lightly covered by colourful cones. That is to say, if each antipode of colour $(d+1)$ is spanned by at least j $\widehat{d+1}$-transversals, then $\mathbf{depth}(\mathbf{0}) \geq j(d+1)$. We can take T_{d+1} to be a transversal spanning the least covered antipode. As we move through the possible values of i in the argument of Subsection 3.1, whenever the octahedron fails to cover $\mathbb{S}^{d-1}$ we will see a new cone covering the lightly covered antipode. Hence $(j-1)+b \geq d$. We thus have that $\mathbf{depth}(\mathbf{0})$ is at least $\max[j(d+1), (d+1)(b+l) - 2bl]$ with $j \geq 1, 1 \leq b, l \leq d$, and $j+b \geq d+1$.

As long as $l \leq \frac{d+1}{2}$, this gives the desired result: by taking either $j \geq \frac{d+1}{2}$ or $b \geq \frac{d+2}{2}$ we get $\mathbf{depth}(\mathbf{0}) \geq \frac{d^2+2d+1}{2}$.

3.3. **Single transversals spanning many antipodes.** This leaves only the case where $l \geq \frac{d+2}{2}$. In this situation, we begin with l simplices containing $\mathbf{0}$ differing only in the $(d+1)$st colour. We can repeat this exercise for each colour, in which case we will either find that for each colour i, $l_i \geq \frac{d+2}{2}$, or, for some colour i, $l_i \leq \frac{d+1}{2}$. In the latter case, we apply the analysis above to get at least $\frac{d^2+2d+1}{2}$ distinct simplices containing zero.

If it happens that we get $l_i \geq \frac{d+2}{2}$ for each i, then for each i we have a set L_i of at least l_i antipodes of colour i which lie on a single $\hat{i}$-transversal U_i. These generate $(d+1)$ sets $X_1, X_2, \ldots X_{d+1}$ of at least $l = \min_i(l_i) \geq \frac{d+2}{2}$ colourful simplices. There may be some duplication between sets, but we note that the simplices within each set are distinct and differ only in the ith colour.

We can identify the simplices that make up the X_i's with vectors in $\{1, 2, \ldots, d+1\}^{d+1}$. We find it helpful to consider them as vectors in $\mathbb{R}^{d+1}$ unrelated to the initial configuration.

A simplex α_d belonging to a given X_i is represented by a vector in $\mathbb{R}^{d+1}$ in the following way. The axes correspond to the $d+1$ colours, and the qth coordinate is set to the index in S_q of the point of colour q of α_d. We recall that the index of points in S_q is set by the arbitrary numbering of points of colour q proposed at the beginning of Section 3.

The vectors associated to the simplices from a given X_i lie on a line segment in the ith coordinate direction. If a simplex is in both X_i and X_q, then the associated vector must lie at the intersection of the corresponding line segments.

Lemma: There are at most d *duplicate vectors* in the union of the X_i's, where a vector that is in $k+1$ sets is counted as k duplicate vectors.

Proof: Consider adding the sets iteratively. We will say that two sets are in the same *component* if they contain a common point, and extend this to an equivalence relation. We remark that each component is contained in the topological component formed by taking the union of the line segments associated to the X_i's, but a given topological component will contain multiple components if the points of intersection of the line segments are not included in the corresponding X_i's.

We begin with $c = 0$ components and $k = 0$ duplicate vectors. Each added set either creates a new component or intersects r components, producing r duplicate vectors while reducing the number of components by $(r-1)$ through the equivalence relation. Therefore at each step $c+k$ increases by 1. Upon termination, we will have at least 1 component, and hence at most d duplicate vectors. $\square$

Then the X_i's contain distinct simplices except possibly for up to $d+1-c \leq d$ repeats arising in this construction, where c is the number of components. This gives us a total of $(d+1)l-(d+1-c)=(d+1)(l-1)+c$ distinct simplices containing $\mathbf{0}$.

However, if c is small, we can readily find additional distinct simplices containing $\mathbf{0}$ by observing that for a fixed colour i, for instance one attaining $l=l_i$, we also have $(d+1-l)$ antipodes outside of L_i. Each of these antipodes must generate some colourful simplex containing $\mathbf{0}$. In fact, for each antipode omitted, we could get $\frac{d+1}{2}$ simplices since either l_i or b is this large, but it does not improve our worst case. Call this set of simplices M, and again consider them as vectors in $\mathbb{R}^{d+1}$. They are not included among the vectors associated to simplices in X_i, since they have different values of coordinate i.

The vectors associated to simplices in M could duplicate vectors from components other than the one containing X_i. However, each such component has a fixed value of colour i. If $c-1 \geq d+1-l$ it may be the case that all such simplices are repeats, but our guarantee is $(d+1)(l-1)+c \geq dl+1$. If $c-1 < d+1-l$ we get at least $d+2-l-c$ additional distinct simplices from vertices omitted from the $(d+1)$ sets. This again guarantees us at least $(d+1)(l-1)+c+(d+2-l-c)=dl+1$ distinct simplices.

Now as $l \geq \frac{d+2}{2}$ we get at least $d\frac{d+2}{2}+1=\frac{d^2+2d+2}{2}$ distinct simplices containing $\mathbf{0}$. Thus our overall worst case for this analysis is at $\frac{d^2+2d+1}{2}=\frac{(d+1)^2}{2}$, which can be rounded up to an integer when d is even. This improves the known bounds for $d \geq 4$, in particular from 12 to 13 when $d=4$. We remark that unlike previous general approaches, this analysis gives the tight bound of 5 when $d=2$.

4. A Combinatorial Generalization

The methods in Section 3 rely on the combinatorial structure of the vectors representing the simplices. Indeed, there is a nice generalization of the colourful simplicial depth problem to systems of vectors of in $\{1,2,\ldots,d+1\}^{d+1}$.

Given sets $\mathbf{S}_1,\ldots,\mathbf{S}_{d+1}$ as in Section 1, we form the system of vectors $\mathcal{V}$ where $\mathbf{v}=(s_1,\ldots,s_{d+1})$ is in $\mathcal{V}$ exactly if the colourful simplex described by $\mathbf{v}$ contains $\mathbf{0}$. In this context, $\widehat{i}$*-transversals* are simply vectors with the ith coordinate removed, and *octahedra* are pairs of disjoint $\widehat{i}$-transversals. The system $\mathcal{V}$ has the following two properties:

1. Every element of $\{1,2,\ldots,d+1\}^{d+1}$ is in some $\mathbf{v}\in\mathcal{V}$. This is the combinatorial requirement from Bárány's Colourful Carathéodory Theorem.

2. For any octahedron $\mathcal{O}$, the parity of the set of vectors using points from $\mathcal{O}$ and a fixed point s_i for the ith coordinate is the same for all choices of s_i. For a system $\mathcal{V}$ arising from colourful simplices, the parity is odd when the octahedron $\mathcal{O}$ contains $\mathbf{0}$, and even when it does not. This is a purely combinatorial version of the Octahedron Lemma mentioned in Section 2.

Question 4.1. For a given $d \geq 2$, what is the size $\nu(d)$ of a minimal system $\mathcal{V}$ of vectors in $\{1,2,\ldots,d+1\}^{d+1}$ satisfying properties 1 and 2?

The system corresponding to the conjectured minimal core colourful Carathéodory configuration from [DHST06] satisfies properties 1 and 2 with d^2+1 vectors, so $\nu(d) \leq \mu(d) \leq d^2+1$. Clearly $\nu(d) \geq d+1$. An exhaustive computer search on a laptop shows in a few seconds that $\nu(2) > 4$ and in a few hours that $\nu(3) > 8$. In other words, this approach computationally verifies that $\mu(2) = 5$ and $\mu(3) = 10$ (using the fact that $\mu(3)$ must be even).

5. A Generalized Core

As a final remark, we mention the recent generalization of the Colourful Carathéodory Theorem in [HPT08] and [ABB+09], in which the condition of $\mathbf{0}$ being in the convex hull of each $\mathbf{S}_i$ is relaxed to require $\mathbf{0}$ to only be in the convex hull of $\mathbf{S}_i \cup \mathbf{S}_j$ for each $i \neq j$. It is natural to ask whether the minimum number of colourful simplices containing $\mathbf{0}$ is lower for configurations satisfying these weaker conditions. Call the analogous quantity $\mu^\diamond(d)$.

In fact, the construction of [DHST06] can be modified in this to produce configurations showing that $\mu^\diamond(d) \leq d+1$ by fixing the points of colours $1, 2, \ldots, d$ in the same way and then clustering all antipodes of the final colour in region that is covered by only a single colourful cone from the first d colours. In this case the relaxed conditions are satisfied almost trivially since $\mathbf{0}$ is in $\mathrm{conv}(\mathbf{S}_i)$ for $i = 1, 2, \ldots, d$. We note that in this configuration, each colour from $1, \ldots, d$ has a unique point which is a generator for all $(d+1)$ colourful simplices colourful simplices containing $\mathbf{0}$. In other words, in contrast to the situation when $\mathbf{0}$ is in all the $\mathbf{S}_i$'s, some (in fact, most) points from the $\mathbf{S}_i$ generate no colourful simplices containing $\mathbf{0}$.

The following simple argument shows that $\mu^\diamond(2) = 3$. Using the assumptions of Section 2, we place the points of the first two colours on the unit circle around $\mathbf{0}$. The condition $\mathbf{0} \in \mathrm{conv}(\mathbf{S}_1 \cup \mathbf{S}_2)$ then means that every half-circle contains a point from $\mathbf{S}_1 \cup \mathbf{S}_2$. If the circle is covered by colourful cones, then each antipode of the remaining colour generates a colourful simplex containing $\mathbf{0}$ and we are done. Otherwise, some segment of the circle is not covered by any colourful cone. This segment must be bounded by two points p and p' of the same $\mathbf{S}_i$, say $\mathbf{S}_1$. The three points of $\mathbf{S}_2$ then are on the longer arc between these points, and for each point of $\mathbf{S}_2$, every point on the longer arc is covered by a colourful cone using that point and either p or p'. The condition that $\mathbf{0} \in \mathrm{conv}(\mathbf{S}_2 \cup \mathbf{S}_3)$ forces at least one of the antipodes of $\mathbf{S}_3$ to lie in the arc that spans the three points of $\mathbf{S}_2$.

Finally, we remark that we can generalize $\mu^\diamond(d)$ combinatorially to $\nu^\diamond(d)$ analogously to Section 4. Combinatorial Property 2 must still hold for such configurations, but Property 1 fails in the constructions above. Nevertheless, we can quickly verify computationally that $\nu^\diamond(3) = \mu^\diamond(3) = 4$.

6. Acknowledgments

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and MITACS, and by the Canada Research Chairs program. The authors would like to thank the referees for helpful comments and Imre Bárány for initiating the discussion in Section 4.

References

[ABB+09] Jorge L. Arocha, Imre Bárány, Javier Bracho, Ruy Fabila, and Luis Montejano, *Very colorful theorems*, Discrete Comput. Geom. **42** (2009), no. 2, 142–154.

[Bár82] Imre Bárány, *A generalization of Carathéodory's theorem*, Discrete Math. **40** (1982), no. 2-3, 141–152.

[BM07] Imre Bárány and Jiří Matoušek, *Quadratically many colorful simplices*, SIAM Journal on Discrete Mathematics **21** (2007), no. 1, 191–198.

[BO97] Imre Bárány and Shmuel Onn, *Colourful linear programming and its relatives*, Math. Oper. Res. **22** (1997), no. 3, 550–567.

[DHST06] Antoine Deza, Sui Huang, Tamon Stephen, and Tamás Terlaky, *Colourful simplicial depth*, Discrete Comput. Geom. **35** (2006), no. 4, 597–604.

[FR05] Komei Fukuda and Vera Rosta, *Data depth and maximal feasible subsystems*, Graph Theory and Combinatorial Optimization (David Avis, Alain Hertz, and Odile Marcotte, eds.), Springer-Verlag, New York, 2005, pp. 37–67.

[HPT08] Andreas F. Holmsen, János Pach, and Helge Tverberg, *Points surrounding the origin*, Combinatorica **28** (2008), no. 6, 633–644.

[Liu90] Regina Y. Liu, *On a notion of data depth based on random simplices*, Ann. Statist. **18** (1990), no. 1, 405–414.

[ST08] Tamon Stephen and Hugh Thomas, *A quadratic lower bound for colourful simplicial depth*, J. Comb. Opt. **16** (2008), no. 4, 324–327.

Advanced Optimization Laboratory, Department of Computing and Software, 1280 Main St. West, McMaster University, Hamilton, Ontario, Canada L8S 4K1.
E-mail address: {deza,xief}@mcmaster.ca

Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6 Canada
E-mail address: tamon@sfu.ca