Symmetry **2013**, *5*, 47-53; doi:10.3390/sym5010047

OPEN ACCESS

symmetry

ISSN 2073-8994

www.mdpi.com/journal/symmetry

Note

A Note on Lower Bounds for Colourful Simplicial Depth

Antoine Deza [1,*], Tamon Stephen [2] and Feng Xie [1]

[1] Advanced Optimization Laboratory, Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4K1, Canada; E-Mail: xief@mcmaster.ca

[2] Department of Mathematics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada; E-Mail: tamon@sfu.ca

* Author to whom correspondence should be addressed; E-Mail: deza@mcmaster.ca; Tel.: +1-905-525-9140 (ext. 23750).

Received: 18 October 2012; in revised form: 18 December 2012 / Accepted: 31 December 2012 / Published: 7 January 2013

Abstract: The colourful simplicial depth problem in dimension d is to find a configuration of $(d+1)$ sets of $(d+1)$ points such that the origin is contained in the convex hull of each set, or colour, but contained in a minimal number of *colourful* simplices generated by taking one point from each set. A construction attaining d^2+1 simplices is known, and is conjectured to be minimal. This has been confirmed up to $d=3$, however the best known lower bound for $d \geq 4$ is $\lceil \frac{(d+1)^2}{2} \rceil$. In this note, we use a branching strategy to improve the lower bound in dimension 4 from 13 to 14.

Keywords: colourful simplicial depth; Colourful Carathéodory Theorem; discrete geometry; polyhedra; combinatorial symmetry

Classification: MSC 52A35, 05D05, 52B05

A *colourful configuration* is the union of $(d+1)$ sets, or colours, $\mathbf{S}_0, \mathbf{S}_1, \ldots, \mathbf{S}_d$ of $(d+1)$ points in $\mathbb{R}^d$. Let $\mathbf{S} = \cup_{i=0}^{d} \mathbf{S}_i$. Without loss of generality we assume that the points in $\mathbf{S} \cup \{\mathbf{0}\}$ are in general position. We are interested in the *colourful simplices* formed by taking the convex hull of a set containing one point of each colour. The *colourful simplicial depth* problem is to find a colourful configuration, with each $\mathbf{S}_i$ containing the origin $\mathbf{0}$ in the interior of its convex hull, minimizing the number of colourful simplices containing $\mathbf{0}$. We denote this minimum by $\mu(d)$. We take the simplices to be closed and remark that the minimum should be attained.

Computing $\mu(d)$ can be viewed as refining Bárány's Colourful Carathéodory Theorem [1] whose original version gives $\mu(d) \geq 1$, and $\mu(d) \geq d+1$ when strengthened to show that *every* point of the configuration generates at least one such simplex. The question of computing $\mu(d)$ was studied in Deza *et al.* [2], which showed $\mu(2) = 5$, that $2d \leq \mu(d) \leq d^2+1$ for $d \geq 3$ and that $\mu(d)$ is even when d is odd. The lower bound has since been improved by Bárány and Matoušek [3] (who verified the conjecture for $d = 3$), Stephen and Thomas [4] and Deza *et al.* [5], which includes the current strongest bound of $\mu(d) \geq \lceil \frac{(d+1)^2}{2} \rceil$ for $d \geq 4$.

One motivation for colourful simplicial depth is to establish bounds on ordinary simplicial depth. A point $p \in \mathbb{R}^d$ 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 [6] as a statistical measure of how representative p is of S. See [7–10] for recent progress on this problem. We remark also that the colourful simplicial depth of a point is the number of solutions to a colourful linear program in the sense of [11] and [12].

Octahedral Systems

Call a $(d+1)$-uniform hypergraph on $\mathbf{S} = \cup_{i=1}^{d+1} \mathbf{S}_i$ a *colourful hypergraph*. A colourful configuration defines a colourful hypergraph by taking hyperedges corresponding to colourful simplices containing $\mathbf{0}$ in their interior. We will call a colourful hypergraph that arises from a colourful configuration with $\mathbf{0} \in \cap_{i=1}^{d+1} \operatorname{conv}(\mathbf{S}_i)$ a *configuration hypergraph*. Our strategy, following [13], is to show that a particular configuration hypergraph whose hyperedges correspond to the colourful simplices containing $\mathbf{0}$ in a configuration cannot exist. The Colourful Carathéodory Theorem gives that any configuration hypergraph must satisfy:

Property 1. *Every vertex of a configuration hypergraph belongs to at least one of its hyperedges.*

Fix a colour i. We call a set t of d points that contains exactly one point from each $\mathbf{S}_j$ other than $\mathbf{S}_i$ an *i-transversal*. That is to say, an i-transversal t has $t \cap \mathbf{S}_i = \emptyset$ and $|t \cap \mathbf{S}_j| = 1$ for $i \neq j$. We call any pair of disjoint i-transversals an *i-octahedron*; these may or may not generate a cross-polytope, *i.e.*, a d-dimensional octahedron, in the geometric sense that their convex hull is a cross-polytope with same coloured points never adjacent in the skeleton of the polytope.

A key property of colourful configurations is that for a fixed i-octahedron $\mathbf{\Omega}$, the parity of the number of colourful simplices containing $\mathbf{0}$ formed using points from $\mathbf{\Omega}$ and a point of colour i does not depend on which point of colour i is chosen. This is a topological fact that corresponds to the fact that $\mathbf{0}$ is either *inside* or *outside* the octahedron, see the *Octahedron Lemma* of [3] for a proof. Figure 1 illustrates this in a two-dimensional case where $\mathbf{0}$ is at the centre of a circle that contains points of the three colours.

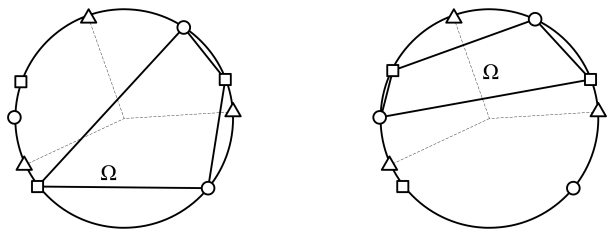
We carry the definitions of *i-transversals* and *i-octahedra* over to the hypergraph setting. Then any configuration hypergraph must satisfy:

Property 2. *For any octahedron $\mathbf{\Omega}$ of a hypergraph, the parity of the set of hyperedges using points from $\mathbf{\Omega}$ and a fixed point s_i for the ith coordinate is the same for all choices of s_i.*

Consider a colourful hypergraph whose vertices are $\mathbf{S} = \cup_{i=0}^{d} \mathbf{S}_i$ and whose hyperedges have exactly one element from each set. If the hypergraph satisfies Property 2 we call it an *octahedral system*, if it additionally satisfies Property 1 we call it an *octahedral system without isolated vertex*. A colourful

configuration with $\mathbf{0} \in \cap_{i=1}^{d+1} \mathrm{conv}(\mathbf{S}_i)$ and k colourful simplices containing $\mathbf{0}$ has a configuration hypergraph that is an octahedral system without isolated vertex with k hyperedges. Let $\nu(d)$ be the minimum number of hyperedges in an octahedral system without isolated vertex with $(d+1)$ colours. Then $\nu(d) \leq \mu(d)$. It is an interesting question whether there are any octahedral systems without isolated vertex not arising from any colourful configurations, and if not, whether $\nu(d) < \mu(d)$ for some d. This purely combinatorial approach was originally suggested by Bárány [14].

Figure 1. Two-dimensional cross-polytopes $\mathbf{\Omega}$ containing and not containing $\mathbf{0}$.



Octahedral systems have the advantage of being combinatorial and finite. In principle, for any particular d and k we can check if there exists an octahedral system without isolated vertex on $\mathbf{S} = \cup_{i=0}^{d} \mathbf{S}_i$ with up to k hyperedges by generating all the—finitely many—hypergraphs with up to k hyperedges, each containing one element from each $\mathbf{S}_i$ and then testing if they satisfy Properties 1 and 2. The difficulty lies in the sheer number of such hypergraphs, and in verifying Property 2 efficiently.

We obtain lower bounds for $\nu(d)$ by trying to build an octahedral system without isolated vertex by adding one hyperedge at a time. We can reduce the search space by exploiting the many combinatorial symmetries in such hypergraphs and considering only configurations that satisfy certain normalizations. However, this alone is not sufficient to improve the known lower bounds even for $d = 4$. We thus turn our attention to how to use Property 2 effectively.

We use two strategies. The first is to look at a particular subset of parity conditions that are relatively independent. The second is to use the following lemma, proved in [15]. Call a hyperedge e of a colourful hypergraph *isolated* if there is no other hyperedge that differs from e only in a single coordinate. Then:

Lemma 1. *An octahedral system with d^2 or fewer hyperedges must not contain any isolated hyperedges.*

Enumeration Details

We begin by fixing an arbitrary colour as colour 0 and an arbitrary 0-transversal. We can label the points in each set from 0 to d and, without loss of generality, take the transversal to contain the 0 point of each set. For convenience we write hyperedges as a string of $(d+1)$ numbers and transversals as string of d numbers with $*$ corresponding to the omitted colour. Thus the 0-transversal considered is $t_0 := *00\ldots0$.

Consider the d octahedra generated by transversals $t_k := *kk\ldots k$, for $k = 1, 2, \ldots d$. Note that the initial numberings are arbitrary, and we may fix them as part of our search algorithm. Given a colourful hypergraph, we can form a $d \times (d+1)$ binary table by writing down for each (r, s) the parity of the number of edges using vertices from the octahedron formed by t_0 and t_r with initial coordinate s. We call this the *parity table*. If a colourful hypergraph satisfies Property 2, its parity table has constant rows.

The advantage of focusing on this table is that the entries are relatively independent. Only hyperedges of the form $x00\ldots0$ can change more than one entry of this table. After accounting for such hyperedges, each entry can only be affected by the $2^d - 1$ hyperedges that are on the relevant octahedron with the given initial coordinate.

We now use the results of [5] to break the problem into several cases based on ℓ, the number of hyperedges containing t_0, b, the number of the parity table octahedra that have odd parity, and j, the minimum number of transversals covering any point of colour 0.

It is clear that for any octahedral system without isolated vertex and with d^2 or fewer hyperedges we must have $1 \leq \ell, b, j \leq d$ and that the number of hyperedges is at least $j(d+1)$. Further, [5] shows that we must have $j + b \geq d + 1$, and that the number of hyperedges must be at least $(b+\ell)(d+1) - 2b\ell$, as well as at least $d\ell + 1$ assuming that the colour 0 is chosen to minimize ℓ and that $\ell \geq \frac{d+2}{2}$. This last fact allows us to assume that $\ell \leq d - 1$.

To rule out possible octahedral systems without isolated vertex of size 13, it is sufficient to consider cases where $j = 1$ or $j = 2$, which in turn means $b = 3$ or $b = 4$. In the case $b = 3$, we have at least $15 - \ell$ simplices, so $\ell = 2$ or $\ell = 3$, and in the case $b = 4$, we have $20 - 3\ell$ so $\ell = 3$. In summary, we need to rule out systems where the triple (ℓ, b, j) is one of $(3, 4, 2), (3, 4, 1), (3, 3, 2), (2, 3, 2)$.

By reordering the points of colour 0, we can take the hyperedges $x0000$ to be in the system for $0 \leq x \leq \ell - 1$, and not in the system for $\ell \leq x \leq 4$. Consider the parity table after including these hyperedges with $\ell = 3$, illustrated in Table 1.

Table 1. The parity table with $\ell = 3$.

	0	1	2	3	4
$*1111$	1	1	1	0	0
$*2222$	1	1	1	0	0
$*3333$	1	1	1	0	0
$*4444$	1	1	1	0	0

Now if $b = 4$, then we are requiring that the parity table be comprised entirely of 1's. So in this case the entries in the first three columns are correct, while the entries in the last two columns are incorrect.

For $(\ell, b, j) = (3, 4, 2)$ we proceed to enumerate configurations as follows. Since $\ell = 3$, we include initial hyperedges $00000, 10000, 20000$. We then add hyperedges to correct each of the eight entries of Table 1, which must be fixed to get the correct parity table for $b = 4$. As previously remarked, adding any hyperedge not of the form $x0000$ will change only a single entry in the parity table. For instance, the entry in the first row and fourth column can be changed only by a hyperedge of the form $3abcd$ where $a, b, c, d \in \{0, 1\}$. Given that that 30000 cannot be added to the configuration without changing ℓ, there remain only 15 possible hyperedges that change the entry, and one must be in our configuration. In fact, by reordering the colours we can take it to be one of 31000, 31100, 31110 and 31111.

We could continue to exploit symmetries in this way—for instance depending on which of the previous 4 hyperedges is chosen, the next hyperedge could be one of 4 to 7 hyperedges fixing the next table entry. However, we did not do this so as to avoid extensive case analysis. Instead, we began branching on all 15 possible hyperedges that switch a given table entry until the table is correct and the partial configuration has 11 hyperedges.

As we branch we check two simple predictors that may indicate that the configuration requires several more hyperedges. First, we look for points that are not currently included in any hyperedge. If some colour still has k uncovered points, then we require k additional hyperedges. Second, since any vertex of colour 0 must be covered by at least j hyperedges, we examine which points of colour 0 are not contained in sufficiently many hyperedges, and get a score k' by summing up the undercounts. At the same time, we may find that all vertices of colour 0 are already covered by more than j hyperedges (especially when $j = 1$), in which case the partial configuration no longer belongs to this subcase and can be excluded. Again, we require k' additional hyperedges. If either k or k' is sufficiently large (in this case 3), then the current partial configuration cannot extend to an octahedral system without isolated vertex with less than 14 hyperedges and is abandoned.

Otherwise, we examine the configuration to see if it has an isolated hyperedge. If it contains an isolated hyperedge e, then by Lemma 1, if the configuration is to extend to an octahedral system without isolated vertex with less than 17 hyperedges, it must include a hyperedge adjacent to e. That is, it must contain e' differing from e only in a single coordinate. There are only 20 such hyperedges so we can branch on them. We then repeat the process of applying predictors and looking for an isolated hyperedge until we either find an octahedral system without isolated vertex with less than 14 hyperedges, or all partial configurations with fewer hyperedges are exhausted.

If we do arrive at a partial configuration with no isolated hyperedges, then as a last resort we may have to branch on all possible hyperedges. However, this happens infrequently enough that the enumeration ends in a reasonable time.

The remaining cases, where (ℓ, b, j) is $(3, 4, 1), (3, 3, 2)$ or $(2, 3, 2)$ are similar. Having exhausted all these cases, we conclude that $\nu(4) \geq 14$, and hence $\mu(d) \geq 14$.

Final remarks

This strategy was implemented by Xie [16] in Python version 2.6 on an AMD Opteron Processor 8356 core (2.3G Hz) and is able to prove that $\nu(4) \geq 14$ in about 30 days of CPU time. This improves by 1 the bound of Deza *et al.* [5], from $\mu(4) \geq 13$ to $\mu(4) \geq 14$. Since this article was written, Deza *et al.* [17] have introduced a different approach that shows $\mu(4) = 17$ and improves the bounds in higher dimension as well.

We note that are $\binom{5^5}{13} \approx 4.25 \times 10^{35}$ colourful hypergraphs on 5 points in each of 5 colours with 13 edges that we need to exclude. In our search strategy, after choosing (l, b, j), the first l hyperedges are determined, and the next $5b + 4l - 2bl$ hyperedges are chosen to fix entries in the parity table. Without considering isolated edges, this leaves a search space of size $15^{5b+4l-2bl} 5^{5^{13-5b-5l+2bl}}$; in our detailed example with $(l, b, j) = (3, 4, 2)$, this is $15^8 (5^5)^2 \approx 2.50 \times 10^{16}$. A space of this size is still slightly beyond the modest computational resources we used. Considering isolated edges further reduces the space substantially, allowing each case to be solved in a few days.

We conclude by mentioning that many aspects of colourful simplices are just beginning to be explored. For instance, the combinatorial complexity of a system of colour simplices is analyzed in [18]. As far as we know, the algorithmic question of computing colourful simplicial depth is untouched, even for $d = 2$ where several interesting algorithms for computing the monochrome simplicial depth have been developed. See for instance the survey [19].

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 second author thanks the University of Cantabria for hospitality while working on this paper. The authors would like to thank the referees for comments that improved the paper and Egon Schulte for editing this special issue.

References

1. Bárány, I. A generalization of Carathéodory's theorem. *Discret. Math.* **1982**, *40*, 141–152.
2. Deza, A.; Huang, S.; Stephen, T.; Terlaky, T. Colourful simplicial depth. *Discret. Comput. Geom.* **2006**, *35*, 597–604.
3. Bárány, I.; Matoušek, J. Quadratically many colorful simplices. *SIAM J. Discret. Math.* **2007**, *21*, 191–198.
4. Stephen, T.; Thomas, H. A quadratic lower bound for colourful simplicial depth. *J. Comb. Opt.* **2008**, *16*, 324–327.
5. Deza, A.; Stephen, T.; Xie, F. More colourful simplices. *Discret. Comput. Geom.* **2011**, *45*, 272–278.
6. Liu, R.Y. On a notion of data depth based on random simplices. *Ann. Statist.* **1990**, *18*, 405–414.
7. Gromov, M. Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry. *Geom. Funct. Anal.* **2010**, *20*, 416–526.
8. Matoušek, J.; Wagner, U. On Gromov's method of selecting heavily covered points. **2012**, Available online: http://arxiv.org/abs/1102.3515 (accessed on 31 December 2012).
9. Karasev, R. A simpler proof of the Boros-Füredi-Bárány-Pach-Gromov theorem. *Discret. Comput. Geom.* **2012**, *47*, 492–495.
10. Král′, D.; Mach, L.; Sereni, J.S. A new lower bound based on Gromov's method of selecting heavily covered points. *Discret. Comput. Geom.* **2012**, *48*, 487–498.
11. Bárány, I.; Onn, S. Colourful linear programming and its relatives. *Math. Oper. Res.* **1997**, *22*, 550–567.
12. Deza, A.; Huang, S.; Stephen, T.; Terlaky, T. The colourful feasibility problem. *Discret. Appl. Math.* **2008**, *156*, 2166–2177.
13. Custard, G.; Deza, A.; Stephen, T.; Xie, F. Small octahedral systems. In *Proceedings of the 23rd Annual Canadian Conference on Computational Geometry*, Toronto, Ontario, Canada, 10–12 August 2011; pp. 267–272.
14. Bárány, I. Hungarian Academy of Sciences, Budapest, Hungary. Personal communication, 2010.
15. Deza, A.; Stephen, T.; Xie, F. Computational lower bounds for colourful simplicial depth. Available online: http://arxiv.org/abs/1210.7621 (accessed on 31 December 2012).
16. Xie, F. Python code for octrahedral system computation. Available online: http://optlab.mcmaster.ca/om/csd/ (accessed on 31 December 2012).
17. Deza, A.; Meunier F.; Sarrabezolles P. A combinatorial approach to colourful simplicial depth. Available online: http://arxiv.org/abs//1212.4720 (accessed on 31 December 2012).

18. Schulz, A.; Tóth, C.D. The union of colorful simplices spanned by a colored point set. *Comput. Geom.* **2013**, in press.
19. Aloupis, G. Geometric measures of data depth. In *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*; American Mathematical Society: Providence, RI, USA, 2006; Volume 72, pp. 147–158.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).