Journal of Machine Learning Research 13 (2012) 1221-1261 Submitted 11/09; Revised 2/12; Published 4/12

A Geometric Approach to Sample Compression

Benjamin I. P. Rubinstein BEN.RUBINSTEIN@MICROSOFT.COM
Microsoft Research
1288 Pear Avenue
Mountain View, CA 94043, USA

J. Hyam Rubinstein RUBIN@MS.UNIMELB.EDU.AU
Department of Mathematics & Statistics
University of Melbourne
Parkville, Victoria 3010, Australia

Editor: Manfred K. Warmuth

Abstract

The Sample Compression Conjecture of Littlestone & Warmuth has remained unsolved for a quarter century. While maximum classes (concept classes meeting Sauer's Lemma with equality) can be compressed, the compression of general concept classes reduces to compressing maximal classes (classes that cannot be expanded without increasing VC dimension). Two promising ways forward are: embedding maximal classes into maximum classes with at most a polynomial increase to VC dimension, and compression via operating on geometric representations. This paper presents positive results on the latter approach and a first negative result on the former, through a systematic investigation of finite maximum classes. Simple arrangements of hyperplanes in hyperbolic space are shown to represent maximum classes, generalizing the corresponding Euclidean result. We show that sweeping a generic hyperplane across such arrangements forms an unlabeled compression scheme of size VC dimension and corresponds to a special case of peeling the one-inclusion graph, resolving a recent conjecture of Kuzmin & Warmuth. A bijection between finite maximum classes and certain arrangements of piecewise-linear (PL) hyperplanes in either a ball or Euclidean space is established. Finally we show that d-maximum classes corresponding to PL-hyperplane arrangements in $\mathbb{R}^d$ have cubical complexes homeomorphic to a d-ball, or equivalently complexes that are manifolds with boundary. A main result is that PL arrangements can be swept by a moving hyperplane to unlabeled d-compress *any* finite maximum class, forming a peeling scheme as conjectured by Kuzmin & Warmuth. A corollary is that some d-maximal classes cannot be embedded into any maximum class of VC-dimension $d+k$, for any constant k. The construction of the PL sweeping involves Pachner moves on the one-inclusion graph, corresponding to moves of a hyperplane across the intersection of d other hyperplanes. This extends the well known Pachner moves for triangulations to cubical complexes.

Keywords: sample compression, hyperplane arrangements, hyperbolic and piecewise-linear geometry, one-inclusion graphs

1. Introduction

Maximum concept classes have the largest cardinality possible for their given VC dimension. Such classes are of particular interest as their special recursive structure underlies all general sample compression schemes known to-date (Floyd, 1989; Warmuth, 2003; Kuzmin and Warmuth, 2007).

It is this structure that admits many elegant geometric and algebraic topological representations upon which this paper focuses.

Littlestone and Warmuth (1986) introduced the study of *sample compression schemes*, defined as a pair of mappings for given concept class C: a *compression function* mapping a C-labeled n-sample to a subsequence of labeled examples and a *reconstruction function* mapping the subsequence to a concept consistent with the entire n-sample. A compression scheme of bounded size—the maximum cardinality of the subsequence image—was shown to imply learnability. The converse—that classes of VC-dimension d admit compression schemes of size d—has become one of the oldest unsolved problems actively pursued within learning theory (Floyd, 1989; Helmbold et al., 1992; Ben-David and Litman, 1998; Warmuth, 2003; Hellerstein, 2006; Kuzmin and Warmuth, 2007; Rubinstein et al., 2007, 2009; Rubinstein and Rubinstein, 2008). Interest in the conjecture has been motivated by its interpretation as the converse to the existence of compression bounds for PAC learnable classes (Littlestone and Warmuth, 1986), the basis of practical machine learning methods on compression schemes (Marchand and Shawe-Taylor, 2003; von Luxburg et al., 2004), and the conjecture's connection to a deeper understanding of the combinatorial properties of concept classes (Rubinstein et al., 2009; Rubinstein and Rubinstein, 2008). Recently Kuzmin and Warmuth (2007) achieved compression of maximum classes without the use of labels. They also conjectured that their elegant min-peeling algorithm constitutes such an unlabeled d-compression scheme for d-maximum classes.

As discussed in our previous work (Rubinstein et al., 2009), maximum classes can be fruitfully viewed as *cubical complexes*. These are also topological spaces, with each cube equipped with a natural topology of open sets from its standard embedding into Euclidean space. We proved that d-maximum classes correspond to *d-contractible complexes*—topological spaces with an identity map homotopic to a constant map—extending the result that 1-maximum classes have trees for one-inclusion graphs. Peeling can be viewed as a special form of contractibility for maximum classes. However, there are many non-maximum contractible cubical complexes that cannot be peeled, which demonstrates that peelability reflects more detailed structure of maximum classes than given by contractibility alone.

In this paper we approach peeling from the direction of simple hyperplane arrangement representations of maximum classes. Kuzmin and Warmuth (2007, Conjecture 1) predicted that d-maximum classes corresponding to simple linear-hyperplane arrangements could be unlabeled d-compressed by sweeping a generic hyperplane across the arrangement, and that concepts are min peeled as their corresponding cell is swept away. We positively resolve the first part of the conjecture and show that sweeping such arrangements corresponds to a new form of *corner peeling*, which we prove is distinct from min peeling. While *min peeling* removes minimum degree concepts from a one-inclusion graph, corner peeling peels vertices that are contained in unique cubes of maximum dimension.

We explore simple hyperplane arrangements in hyperbolic geometry, which we show correspond to a set of maximum classes, properly containing those represented by simple linear Euclidean arrangements. These classes can again be corner peeled by sweeping. Citing the proof of existence of maximum unlabeled compression schemes due to Ben-David and Litman (1998), Kuzmin and Warmuth (2007) ask whether unlabeled compression schemes for infinite classes such as positive half spaces can be constructed explicitly. We present constructions for illustrative but simpler classes, suggesting that there are many interesting infinite maximum classes admitting explicit compression schemes, and under appropriate conditions, sweeping infinite Euclidean, hyperbolic or PL arrangements corresponds to compression by corner peeling.

Next we prove that all maximum classes in $\{0,1\}^n$ are represented as simple arrangements of piecewise-linear (PL) hyperplanes in the n-ball. This extends previous work by Gärtner and Welzl (1994) on viewing simple PL-hyperplane arrangements as maximum classes. The close relationship between such arrangements and their hyperbolic versions suggests that they could be equivalent. Resolving the main problem left open in the preliminary version of this paper (Rubinstein and Rubinstein, 2008), we show that sweeping of d-contractible PL arrangements does compress all finite maximum classes by corner peeling, completing (Kuzmin and Warmuth, 2007, Conjecture 1).

We show that a one-inclusion graph Γ can be represented by a d-contractible PL-hyperplane arrangement if and only if Γ is a strongly contractible cubical complex. This motivates the nomenclature of d-contractible for the class of arrangements of PL hyperplanes. Note then that these one-inclusion graphs admit a corner-peeling scheme of the same size d as the largest dimension of a cube in Γ. Moreover if such a graph Γ admits a corner-peeling scheme, then it is a contractible cubical complex. We give a simple example to show that there are one-inclusion graphs which admit corner-peeling schemes but are not strongly contractible and so are not represented by a d-contractible PL-hyperplane arrangement.

Compressing *maximal classes*—classes which cannot be grown without an increase to their VC dimension—is sufficient for compressing all classes, as embedded classes trivially inherit compression schemes of their super-classes. This reasoning motivates the attempt to embed d-maximal classes into $O(d)$-maximum classes (Kuzmin and Warmuth, 2007, Open Problem 3). We present non-embeddability results following from our earlier counter-examples to Kuzmin & Warmuth's minimum degree conjecture (Rubinstein et al., 2009), and our new results on corner peeling. We explore with examples, maximal classes that can be compressed but not peeled, and classes that are not strongly contractible but can be compressed.

Finally, we investigate algebraic topological properties of maximum classes. Most notably we characterize d-maximum classes, corresponding to simple linear Euclidean arrangements, as cubical complexes homeomorphic to the d-ball. The result that such classes' boundaries are homeomorphic to the $(d-1)$-sphere begins the study of the boundaries of maximum classes, which are closely related to peeling. We conclude with several open problems.

2. Background

We begin by presenting relevant background material on algebraic topology, computational learning theory, and sample compression.

2.1 Algebraic Topology

Definition 1 *A* homeomorphism *is a one-to-one and onto map f between topological spaces such that both f and f^{-1} are continuous. Spaces X and Y are said to be* homeomorphic *if there exists a homeomorphism $f : X \to Y$.*

Definition 2 *A* homotopy *is a continuous map $F : X \times [0,1] \to Y$. The* initial map *is F restricted to $X \times \{0\}$ and the* final map *is F restricted to $X \times \{1\}$. We say that the initial and final maps are* homotopic. *A* homotopy equivalence *between spaces X and Y is a pair of maps $f : X \to Y$ and $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are homotopic to the identity maps on Y and X respectively. We* *say that X and Y have the* same homotopy type *if there is a homotopy equivalence between them. A* deformation retraction *is a special homotopy equivalence between a space X and a subspace $A \subseteq X$. It is a continuous map $r : X \to X$ with the properties that the restriction of r to A is the identity map on A, r has range A and r is homotopic to the identity map on X.*

Definition 3 *A* cubical complex *is a union of solid cubes of the form $[a_1,b_1] \times \ldots \times [a_m,b_m]$, for bounded $m \in \mathbb{N}$, such that the intersection of any two cubes in the complex is either a cubical face of both cubes or the empty-set.*

Definition 4 *A* contractible cubical complex *X is one which has the same homotopy type as a one point space $\{p\}$. X is contractible if and only if the constant map from X to p is a homotopy equivalence.*

Definition 5 *A* simplicial complex *is a union of simplices, each of which is affinely equivalent*[1] *to the convex hull of $k+1$ points $(0,0,\ldots,0),(1,0,\ldots,0),\ldots(0,0,\ldots,1)$ in $\mathbb{R}^k$, for some k. The intersection of any two simplices in the complex is either a face of both simplices or the empty-set. A map $f : X \to Y$ is called* simplicial *if X,Y are simplicial complexes and f maps each simplex of X to a simplex of Y so that vertices are mapped to vertices and the map is affine linear. A* subdivision *of a simplicial complex is a new simplicial complex with the same underlying point-set obtained by cutting up the original simplices into smaller simplices.*

For a more formal treatment of simplicial complexes see (Rourke and Sanderson, 1982). We will need the concepts of piecewise-linear (PL) manifolds and maps.

Definition 6 *A mapping $f : X \to Y$ is called* piecewise linear *(PL) if X,Y are simplicial complexes and there are subdivisions $X^\star,Y^\star$ of the respective complexes, so that $f : X^\star \to Y^\star$ is simplicial. A PL homeomorphism $f : X \to Y$ is a bijection so that both f, f^{-1} are PL maps. A PL manifold M is a space which is covered by open sets U_α for $\alpha \in I$ some index set, together with bijections $\phi_\alpha : U_\alpha \to V_\alpha$, where V_α is an open set in $\mathbb{R}^n$. Moreover when $U_\alpha \cap U_\beta \neq \emptyset$, then the transition function $\phi_\beta \circ {\phi_\alpha}^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is a PL homeomorphism. A pair (U_α, ϕ_α) is called a* chart *for M.*

2.2 Pachner Moves

Pachner (1987) showed that triangulations of manifolds which are combinatorially equivalent after subdivision are also equivalent by a series of moves which are now referred to as Pachner moves. For the main result of this paper, we need a version of Pachner moves for cubical structures rather than simplicial ones. The main idea of Pachner moves remains the same.

A *Pachner move* replaces a topological d-ball U divided into d-cubes, with another ball U' with the same $(d-1)$-cubical boundary but with a different interior cubical structure. In dimension $d = 2$, for example, such an initial ball U can be constructed by taking three 2-cubes forming a hexagonal disk and in dimension $d = 3$, four 3-cubes forming a rhombic dodecahedron, which is a polyhedron U with 12 quadrilateral faces in its boundary. The set U' of d-cubes is attached to the same boundary as for U, that is, $\partial U = \partial U'$, as cubical complexes homeomorphic to the $(d-1)$-sphere. Moreover, U' and U are isomorphic cubical complexes, but the gluing between their boundaries produces

1. The simplices are related via an affine bijection.

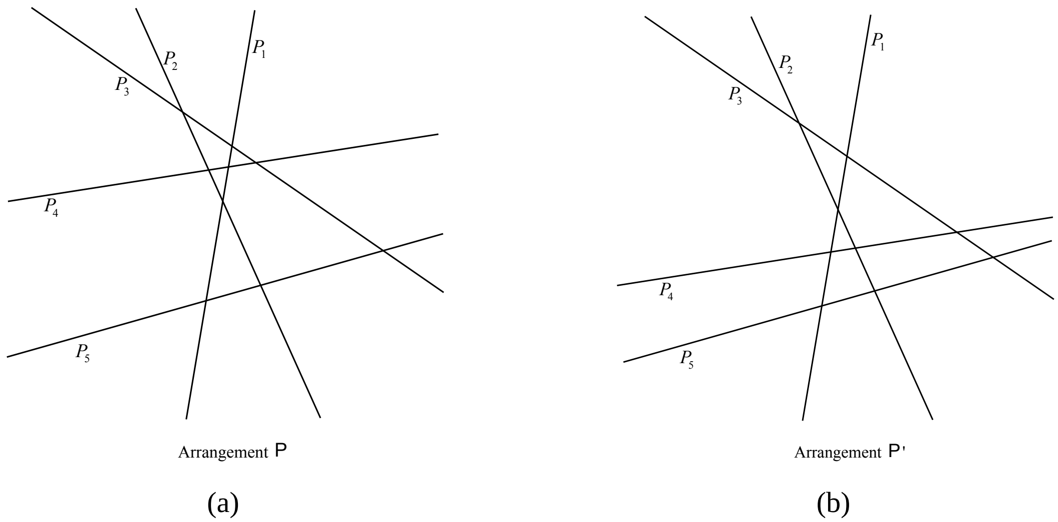


Figure 1: (a) An example linear-hyperplane arrangement $\mathcal{P}$ and (b) the result of a Pachner move of hyperplane P_4 on $\mathcal{P}$.

the boundary of the 3- or 4-cube, as a 2- or 3-dimensional cubical structure on the 2- or 3-sphere respectively.

To better understand this move, consider the cubical face structure of the boundary V of the $(d+1)$-cube. This is a d-sphere containing $2d+2$ cubes, each of dimension d. There are many embeddings of the $(d-1)$-sphere as a cubical subcomplex into V, dividing it into a pair of d-balls. One ball is combinatorially identical to U and the other to U'.

There are a whole series of Pachner moves in each dimension d, but we are only interested in the ones where the pair of balls U, U' have the same numbers of d-cubes. In Figure 1 a change in a hyperplane arrangement is shown, which corresponds to a Pachner move on the corresponding one-inclusion graph (considered as a cubical complex).

2.3 Concept Classes and their Learnability

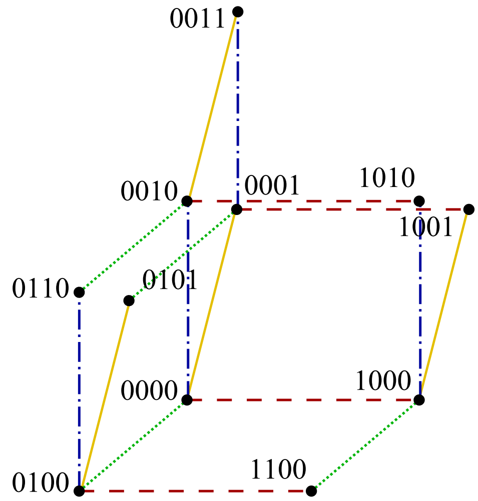
A *concept class* C on *domain* X, is a subset of the power set of set X or equivalently $C \subseteq \{0,1\}^X$. We primarily consider finite domains and so will write $C \subseteq \{0,1\}^n$ in the sequel, where it is understood that $n = |X|$ and the n dimensions or *colors* are identified with an ordering $\{x_i\}_{i=1}^n = X$.

The *one-inclusion graph* $\mathcal{G}(C)$ of $C \subseteq \{0,1\}^n$ is the graph with vertex-set C and edge-set containing $\{u,v\} \subseteq C$ iff u and v differ on exactly one component (Haussler et al., 1994); $\mathcal{G}(C)$ forms the basis of a prediction strategy with essentially-optimal worst-case expected risk. $\mathcal{G}(C)$ can be viewed as a simplicial complex in $\mathbb{R}^n$ by filling in each face with a product of continuous intervals (Rubinstein et al., 2009). Each edge $\{u,v\}$ in $\mathcal{G}(C)$ is labeled by the component on which the two vertices u, v differ.

Example 1 *An example concept class in* $\{0,1\}^4$ *is enumerated in Figure 2(a). The corresponding one-inclusion graph is visualized in Figure 2(b), making immediately apparent the interpretation of*

	x_1	x_2	x_3	x_4
v_0	0	0	0	0
v_1	1	0	0	0
v_2	0	1	0	0
v_3	0	0	1	0
v_4	1	0	1	0
v_5	1	1	0	0
v_6	0	1	1	0
v_7	0	0	0	1
v_8	1	0	0	1
v_9	0	1	0	1
v_{10}	0	0	1	1

(a)



(b)

Figure 2: (a) A concept class in $\{0,1\}^4$ that is maximum with VC-dim 2 and (b) the one-inclusion graph of the concept class.

the object as a simplicial complex: in this case the concepts form vertices which are connected by edges; these edges bound 2-cubes.

Probably Approximately Correct learnability of a concept class $C \subseteq \{0,1\}^X$ is characterized by the finiteness of the Vapnik-Chervonenkis (VC) dimension of C (Blumer et al., 1989). One key to all such results is Sauer's Lemma.

Definition 7 *The* VC dimension *of concept class* $C \subseteq \{0,1\}^X$ *is defined as* $\mathrm{VC}(C) = \sup\left\{n \,\middle|\, \exists Y \in \binom{X}{n}, \Pi_Y(C) = \{0,1\}^n\right\}$ *where* $\Pi_Y(C) = \{(c(x_1),\dots,c(x_n)) \mid c \in C\} \subseteq \{0,1\}^n$ *is the* projection *of* C *on sequence* $Y = (x_1,\dots,x_n)$.

Lemma 8 (Vapnik and Chervonenkis, 1971; Sauer, 1972; Shelah, 1972) *The cardinality of any concept classes* $C \subseteq \{0,1\}^n$ *is bounded by* $|C| \leq \sum_{i=1}^{\mathrm{VC}(C)} \binom{n}{i}$.

Motivated by maximizing concept class cardinality under a fixed VC dimension, which is related to constructing general sample compression schemes (see Section 2.4), Welzl (1987) defined the following special classes.

Definition 9 *Concept class* $C \subseteq \{0,1\}^X$ *is called* maximal *if* $\mathrm{VC}(C \cup \{c\}) > \mathrm{VC}(C)$ *for all* $c \in \{0,1\}^X \backslash C$. *Furthermore if* $\Pi_Y(C)$ *satisfies Sauer's Lemma with equality for each* $Y \in \binom{X}{n}$, *for every* $n \in \mathbb{N}$, *then* C *is termed* maximum. *If* $C \subseteq \{0,1\}^n$ *then* C *is maximum (and hence maximal) if* C *meets Sauer's Lemma with equality.*

Example 2 *The concept class of Example 1 has VC-dimension* 2 *as witnessed by projecting onto any two of the four available axes. Moreover its cardinality of* 11 *exactly meets Sauer's Lemma with equality, so the class is also maximum.*

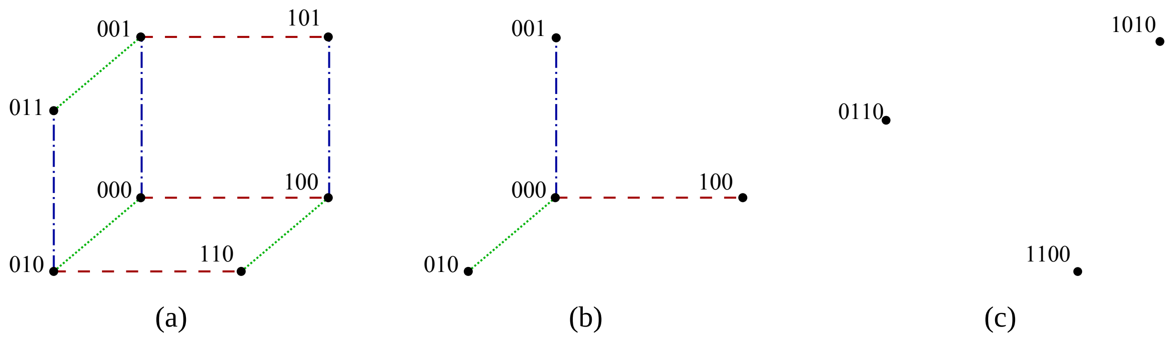


Figure 3: The (a) projection (b) reduction and (c) tail of the concept class of Figure 2 with respect to projecting on to the first three coordinates (i.e., projecting out the fourth coordinate).

The *reduction* of $C \subseteq \{0,1\}^n$ with respect to $i \in [n] = \{1,\ldots,n\}$ is the class $C^i = \Pi_{[n]\backslash\{i\}}\left(\left\{c \in C \mid i \in I_{\mathcal{G}(C)}(c)\right\}\right)$ where $I_{\mathcal{G}(C)}(c) \subseteq [n]$ denotes the labels of the edges incident to vertex c; a *multiple reduction* is the result of performing several reductions in sequence. The *tail* of class C is $\mathrm{tail}_i(C) = \left\{c \in C \mid i \notin I_{\mathcal{G}(C)}(c)\right\}$. Welzl showed that if C is d-maximum, then $\Pi_{[n]\backslash\{i\}}(C)$ and C^i are maximum of VC-dimensions d and $d-1$ respectively.

Example 3 *A projection, reduction and tail of the concept class of Figure 2 are shown in Figures 3(a)—3(c) respectively, when projecting onto coordinates* $\{1,2,3\}$. *In particular note that the reduction, like the projection, is a class in the smaller* 3*-cube while the tail is in the original* 4*-cube. Moreover note that the projection and reduction and maximum with VC-dimensions* 2 *and* 1 *respectively.*

The results presented below relate to other geometric and topological representations of maximum classes existing in the literature. Under the guise of 'forbidden labels', Floyd (1989) showed that maximum $C \subseteq \{0,1\}^n$ of VC-dim d is the union of a maximally overlapping *d-complete collection of cubes* (Rubinstein et al., 2009)—defined as a collection of $\binom{n}{d}$ d-cubes which uniquely project onto all $\binom{n}{d}$ possible sets of d coordinate directions. (An alternative proof was developed by Neylon 2006.) It has long been known that VC-1 maximum classes have one-inclusion graphs that are trees (Dudley, 1985); we previously extended this result by showing that when viewed as complexes, d-maximum classes are contractible d-cubical complexes (Rubinstein et al., 2009). Finally the cells of a simple linear arrangement of n hyperplanes in $\mathbb{R}^d$ form a VC-d maximum class in the n-cube (Edelsbrunner, 1987), but not all finite maximum classes correspond to such Euclidean arrangements (Floyd, 1989).

Example 4 *It is immediately clear from visual inspection that the* 2*-maximum concept classes of Figures 2 and 3(a) are composed of complete collections of* 2*-cubes. Similarly the* 1*-maximum class of Figure 3(c) is a tree with one edge of each color.*

2.4 Sample Compression Schemes

Littlestone and Warmuth (1986) showed that the existence of a compression scheme of finite size is sufficient for learnability of C, and conjectured the converse, that $\mathrm{VC}(C) = d < \infty$ implies a compression scheme of size d. Later Warmuth (2003) weakened the conjectured size to $O(d)$. To-date it is only known that maximum classes can be d-compressed (Floyd, 1989). Unlabeled compression was first explored by Ben-David and Litman (1998); Kuzmin and Warmuth (2007) defined unlabeled compression as follows, and explicitly constructed schemes of size d for maximum classes.

Definition 10 *Let C be a d-maximum class on a finite domain X. A mapping r is called a* representation mapping *of C if it satisfies the following conditions:*

1. *r is a bijection between C and subsets of X of size at most d; and*
2. *[*non-clashing*] :*[2] $\Pi_{r(c)\cup r(c')}(c) \neq \Pi_{r(c)\cup r(c')}(c')$ *for all $c,c' \in C$, $c \neq c'$.*

As with all previously published labeled schemes, all previously known unlabeled compression schemes for maximum classes exploit their special recursive projection-reduction structure and so it is doubtful that such schemes will generalize. Kuzmin and Warmuth (2007, Conjecture 2) conjectured that their *min-peeling* algorithm constitutes an unlabeled d-compression scheme for maximum classes; it iteratively removes minimum degree vertices from $\mathcal{G}(C)$, representing the corresponding concepts by the remaining incident dimensions in the graph. The authors also conjectured that sweeping a hyperplane in general position across a simple linear arrangement forms a compression scheme that corresponds to min peeling the associated maximum class (Kuzmin and Warmuth, 2007, Conjecture 1). A particularly promising approach to compressing general classes is via their maximum-embeddings: a class C embedded in class C' trivially inherits any compression scheme for C', and so an important open problem is to embed maximal classes into maximum classes with at most a linear increase in VC dimension (Kuzmin and Warmuth, 2007, Open Problem 3).

3. Preliminaries

A first step towards characterizing and compressing maximum classes is a process of building them. After describing this process of *lifting* we discuss compressing maximum classes by peeling, and properties of the boundaries of maximum classes.

3.1 Constructing All Maximum Classes

The aim in this section is to describe an algorithm for constructing all maximum classes of VC-dimension d in the n-cube. This process can be viewed as the inverse of mapping a maximum class to its d-maximum projection on $[n]\backslash\{i\}$ and the corresponding $(d-1)$-maximum reduction.

Definition 11 *Let $C,C' \subseteq \{0,1\}^n$ be maximum classes of VC-dimensions $d,d-1$ respectively, so that $C' \subset C$, and let $C_1,C_2 \subset C$ be d-cubes, that is, d-faces of the n-cube $\{0,1\}^n$.*

1. *C_1,C_2 are* connected *if there exists a path in the one-inclusion graph $\mathcal{G}(C)$ with end-points in C_1 and C_2; and*
2. *C_1,C_2 are said to be C'-*connected *if there exists such a connecting path that further does not intersect C'.*

*The C'-*connected components *of C are the equivalence classes of the d-cubes of C under the C'-connectedness relation.*

2. We abuse notation slightly by applying projections, originally defined to operate on concept classes in Definition 7, to concepts.

Algorithm 1 MAXIMUMCLASSES(n,d)

Given: $n \in \mathbb{N}, d \in [n]$

Returns: the set of d-maximum classes in $\{0,1\}^n$

1. **if** $d = 0$ **then return** $\{\{\mathbf{v}\} \mid \mathbf{v} \in \{0,1\}^n\}$;
2. **if** $d = n$ **then return** $\{0,1\}^n$;
3. $\mathcal{M} \leftarrow \emptyset$;
   for each $C \in$ MAXIMUMCLASSES$(n-1,d)$,
   $C' \in$ MAXIMUMCLASSES$(n-1,d-1)$ s.t. $C' \subset C$ **do**
4. $\{C_1,\ldots,C_k\} \leftarrow C'$-connected components of C ;
5. $\mathcal{M} \leftarrow \mathcal{M} \cup \bigcup_{\mathbf{p}\in\{0,1\}^k} \left\{(C' \times \{0,1\}) \cup \bigcup_{q\in[k]} C_q \times \{p_q\}\right\}$;
   done
6. **return** $\mathcal{M}$;

The recursive algorithm for constructing all maximum classes of VC-dimension d in the n-cube, detailed as Algorithm 1, considers each possible d-maximum class C in the $(n-1)$-cube and each possible $(d-1)$-maximum subclass C' of C as the projection and reduction of a d-maximum class in the n-cube, respectively. The algorithm *lifts* C and C' to all possible maximum classes in the n-cube. Then $C' \times \{0,1\}$ is contained in each lifted class; so all that remains is to find the tails from the complement of the reduction in the projection. It turns out that each C'-connected component C_i of C can be lifted to either $C_i \times \{0\}$ or $C_i \times \{1\}$ arbitrarily and independently of how the other C'-connected components are lifted. The set of lifts equates to the set of d-maximum classes in the n-cube that project-reduce to (C,C').

Lemma 12 MAXIMUMCLASSES(n,d) *(cf. Algorithm 1) returns the set of maximum classes of VC-dimension d in the n-cube for all $n \in \mathbb{N}, d \in [n]$.*

Proof We proceed by induction on n and d. The base cases correspond to $n \in \mathbb{N}, d \in \{0,n\}$ for which all maximum classes, enumerated as singletons in the n-cube and the n-cube itself respectively, are correctly produced by the algorithm. For the inductive step we assume that for $n \in \mathbb{N}, d \in [n-1]$ all maximum classes of VC-dimension d and $d-1$ in the $(n-1)$-cube are already known by recursive calls to the algorithm. Given this, we will show that MAXIMUMCLASSES(n,d) returns only d-maximum classes in the n-cube, and that all such classes are produced by the algorithm.

Let classes $C \in$ MAXIMUMCLASSES$(n-1,d)$ and $C' \in$ MAXIMUMCLASSES$(n-1,d-1)$ be such that $C' \subset C$. Then C is the union of a d-complete collection and C' is the union of a $(d-1)$-complete collection of cubes that are faces of the cubes of C. Consider a concept class $C^\star$ formed from C and C' by Algorithm 1. The algorithm partitions C into C'-connected components $C_1,\ldots,C_k$ each of which is a union of d-cubes. While C' is lifted to $C' \times \{0,1\}$, some subset of the components $\{C_i\}_{i\in S_0}$ are lifted to $\{C_i \times \{0\}\}_{i\in S_0}$ while the remaining components are lifted to $\{C_i \times \{1\}\}_{i\notin S_0}$. Here S_0 ranges over all subsets of $[k]$, selecting which components are lifted to 0; the complement of S_0 specifies those components lifted to 1. By definition $C^\star$ is a d-complete collection of cubes with cardinality equal to $\binom{n}{\leq d}$ since $|C^\star| = |C'| + |C|$ (Kuzmin and Warmuth, 2007). So $C^\star$ is d-maximum (Rubinstein et al., 2009, Theorem 34).

If we now consider any d-maximum class $C^\star \subseteq \{0,1\}^n$, its projection on $[n]\backslash\{i\}$ is a d-maximum class $C \subseteq \{0,1\}^{n-1}$ and $C^{\star i}$ is the $(d-1)$-maximum projection $C' \subset C$ of all the d-cubes in $C^\star$

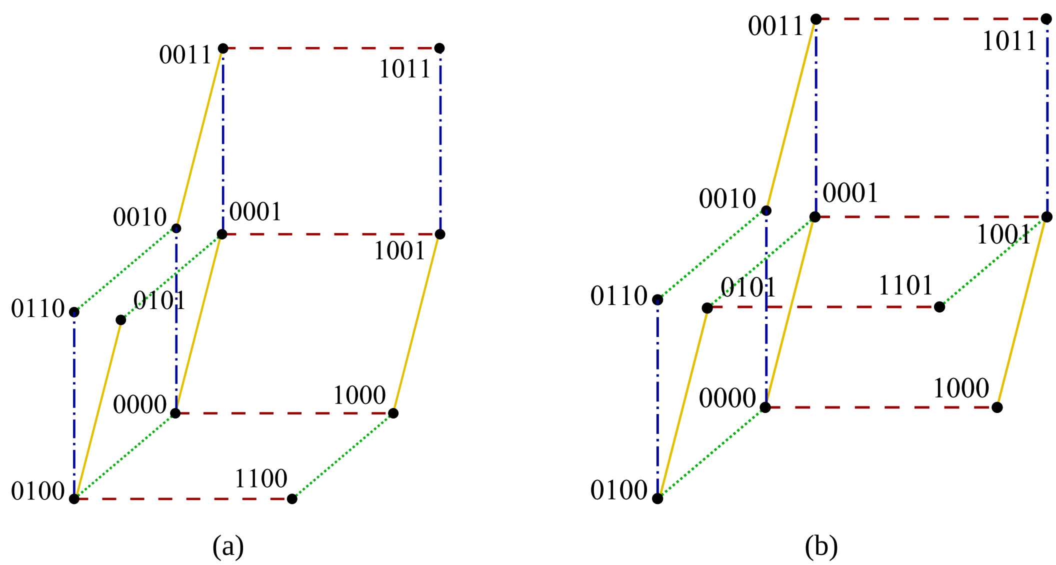


Figure 4: 2-maximum concept classes in $\{0,1\}^4$ constructed by lifting concept class Figure 3(a) as the projection, and concept class Figure 3(b) as the reduction.

which contain color i. It is thus clear that $C^\star$ must be obtained by lifting parts of the C'-connected components of C to the 1 level and the remainder to the 0 level, and C' to $C' \times \{0,1\}$. We will now show that if the vertices of each component are not lifted to the same levels, then while the number of vertices in the lift match that of a d-maximum class in the n-cube, the number of edges are too few for such a maximum class. Define a lifting operator on C as $\ell(v) = \{v\} \times \ell_v$, where $\ell_v \subseteq \{0,1\}$ and

$$|\ell_v| \quad = \quad \begin{cases} 2\ , & \text{if } v \in C' \\ 1, & \text{if } v \in C \backslash C' \end{cases} .$$

Consider now an edge $\{u,v\}$ in $\mathcal{G}(C)$. By the definition of a C'-connected component there exists some C_j such that either $u,v \in C_j \backslash C'$, $u,v \in C'$ or WLOG $u \in C_j \backslash C', v \in C'$. In the first case $\ell(u) \cup \ell(v)$ is an edge in the lifted graph iff $\ell_u = \ell_v$. In the second case $\ell(u) \cup \ell(v)$ contains four edges and in the last it contains a single edge. Furthermore, it is clear that this accounts for all edges in the lifted graph by considering the projection of an edge in the lifted product. Thus any lift other than those produced by Algorithm 1 induces strictly too few edges for a d-maximum class in the n-cube (cf. Kuzmin and Warmuth, 2007, Corollary 7.5). ■

Example 5 *Let C and C' refer to the* 2- *and* 1-*maximum concept classes in Figures 3(a) and 3(b) respectively. Then Figures 4(a), 4(b) and 2 make up all possible* 2-*maximum classes (up to symmetry) resulting from lifting projection C and reduction C'. Figure 2 corresponds to lifting no C'-connected components of C; Figure 4(a) corresponds to lifting just one component; and Figure 4(b) corresponds to lifting two components. (Note that Figure 4(a) and Figure 4(b) are actually equivalent after a symmetry.)*

3.2 Corner Peeling

Kuzmin and Warmuth (2007, Conjecture 2) conjectured that their simple *min-peeling* procedure is a valid unlabeled compression scheme for maximum classes. Beginning with a concept class $C_0 = C \subseteq \{0,1\}^n$, min peeling operates by iteratively removing a vertex v_t of minimum-degree in $\mathcal{G}(C_t)$ to produce the peeled class $C_{t+1} = C_t \backslash \{v_t\}$. The concept class corresponding to v_t is then represented by the dimensions of the edges incident to v_t in $\mathcal{G}(C_t)$, $I_{\mathcal{G}(C_t)}(v_t) \subseteq [n]$. Providing that no-clashing holds for the algorithm, the size of the min-peeling scheme is the largest degree encountered during peeling. Kuzmin and Warmuth predicted that this size is always at most d for d-maximum classes. We explore these questions for a related special case of peeling, where we prescribe which vertex to peel at step t as follows.

Definition 13 *We say that $C \subseteq \{0,1\}^n$ can be* corner peeled *if there exists an ordering $v_1, \ldots, v_{|C|}$ of the vertices of C such that, for each $t \in [|C|]$ where $C_0 = C$,*

1. *$v_t \in C_{t-1}$ and $C_t = C_{t-1} \backslash \{v_t\}$;*
2. *There exists a unique cube C'_{t-1} of maximum dimension over all cubes in C_{t-1} containing v_t;*
3. *The neighbors $\Gamma(v_t)$ of v_t in $\mathcal{G}(C_{t-1})$ satisfy $\Gamma(v_t) \subseteq C'_{t-1}$; and*
4. *$C_{|C|} = \emptyset$.*

The v_t are termed the corner vertices *of C_{t-1} respectively. If d is the maximum degree of each v_t in $\mathcal{G}(C_{t-1})$, then C is d* corner peeled.

Note that we do not constrain the cubes C'_t to be of non-increasing dimension. It turns out that an important property of maximum classes is invariant to this kind of peeling.

Definition 14 *We call a class $C \subseteq \{0,1\}^n$* shortest-path closed *if for any $u, v \in C$, $\mathcal{G}(C)$ contains a path connecting u, v of length $\|u - v\|_1$.*

Lemma 15 *If $C \subseteq \{0,1\}^n$ is shortest-path closed and $v \in C$ is a corner vertex of C, then $C \backslash \{v\}$ is shortest-path closed.*

Proof Consider a shortest-path closed $C \subseteq \{0,1\}^n$. Let c be a corner vertex of C, and denote the cube of maximum dimension in C, containing c, by C'. Consider $\{u, v\} \subseteq C \backslash \{c\}$. By assumption there exists a u-v-path p of length $\|u - v\|_1$ contained in C. If c is not in p then p is contained in the peeled product $C \backslash \{c\}$. If c is in p then p must cross C' such that there is another path of the same length which avoids c, and thus $C \backslash \{c\}$ is shortest-path closed. ■

3.2.1 CORNER PEELING IMPLIES COMPRESSION

Theorem 16 *If a maximum class C can be corner peeled then C can be d-unlabeled compressed.*

Proof The invariance of the shortest-path closed property under corner peeling is key. The corner-peeling unlabeled compression scheme represents each $v_t \in C$ by $r(v_t) = I_{\mathcal{G}(C_{t-1})}(v_t)$, the colors of

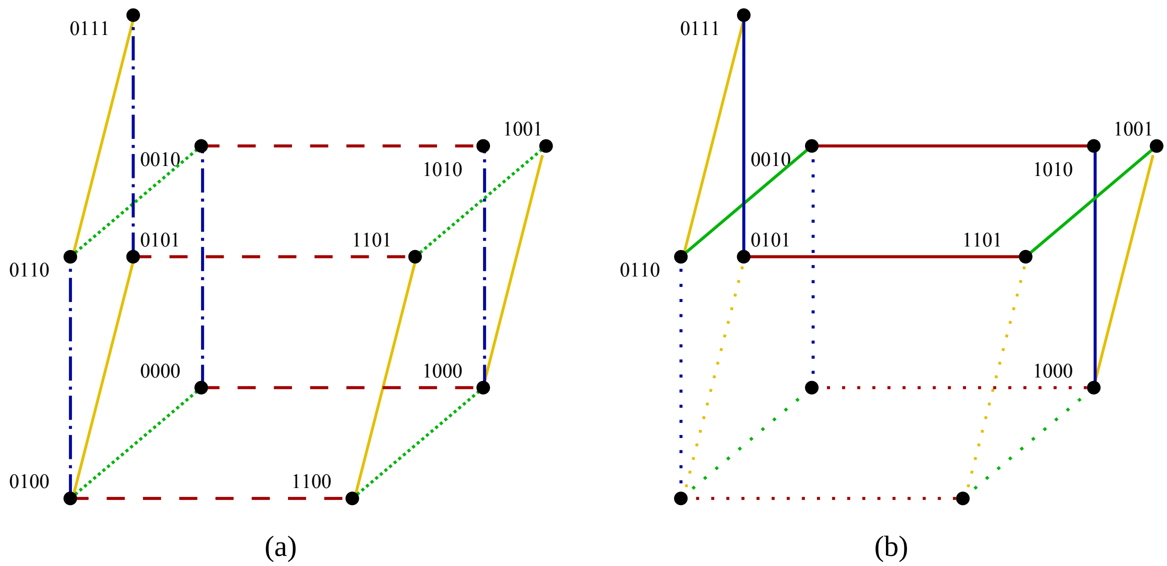


Figure 5: (a) A 2-maximum class in the 4-cube and (b) its boundary highlighted by solid lines.

the cube C'_{t-1} which is deleted from C_{t-1} when v_t is corner peeled. We claim that any two vertices $v_s, v_t \in C$ have non-clashing representatives. WLOG, suppose that $s < t$. The class C_{s-1} must contain a shortest v_s-v_t-path p. Let i be the color of the single edge contained in p that is incident to v_s. Color i appears once in p, and is contained in $r(v_s)$. This implies that $v_{s,i} \neq v_{t,i}$ and that $i \in r(v_s) \cup r(v_t)$, and so $v_s | (r(v_s) \cup r(v_t)) \neq v_t | (r(v_s) \cup r(v_t))$. By construction, $r(\cdot)$ is a bijection between C and all subsets of $[n]$ of cardinality $\leq \text{VC}(C)$. ■

If the oriented one-inclusion graph, with each edge directed away from the incident vertex represented by the edge's color, has no cycles, then that representation's compression scheme is termed *acyclic* (Floyd, 1989; Ben-David and Litman, 1998; Kuzmin and Warmuth, 2007).

Proposition 17 *All corner-peeling unlabeled compression schemes are acyclic.*

Proof We follow the proof that the min-peeling algorithm is acyclic (Kuzmin and Warmuth, 2007). Let $v_1, \ldots, v_{|C|}$ be a corner vertex ordering of C. As a corner vertex v_t is peeled, its unoriented incident edges are oriented away from v_t. Thus all edges incident to v_1 are oriented away from v_1 and so the vertex cannot take part in any cycle. For $t > 1$ assume $V_t = \{v_s \mid s < t\}$ is disjoint from all cycles. Then v_t cannot be contained in a cycle, as all incoming edges into v_t are incident to some vertex in V_t. Thus the oriented $\mathcal{G}(C)$ is indeed acyclic. ■

3.3 Boundaries of Maximum Classes

We now turn to the geometric boundaries of maximum classes, which are closely related to corner peeling.

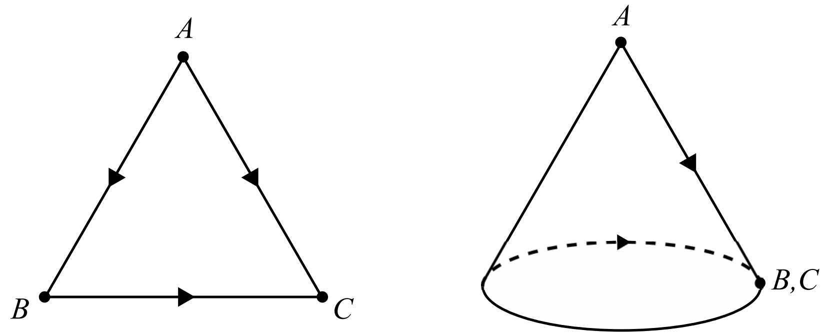


Figure 6: The first steps of building the dunce hat in Example 7.

Definition 18 *The boundary ∂C of a d-maximum class C is defined as all the $(d-1)$-subcubes which are the faces of a single d-cube in C.*

Maximum classes, when viewed as cubical complexes, are analogous to soap films (an example of a minimal energy surface encountered in nature), which are obtained when a wire frame is dipped into a soap solution. Under this analogy the boundary corresponds to the wire frame and the number of d-cubes can be considered the area of the soap film. An important property of the boundary of a maximum class is that all lifted reductions meet the boundary multiple times.

Theorem 19 *Every d-maximum class has boundary containing at least two $(d-1)$-cubes of every combination of $d-1$ colors, for all $d>1$.*

Proof We use the lifting construction of Section 3.1. Let $C^\star \subseteq \{0,1\}^n$ be a 2-maximum class and consider color $i \in [n]$. Then the reduction $C^{\star i}$ is an unrooted tree with at least two leaves, each of which lifts to an i-colored edge in $C^\star$. Since the leaves are of degree 1 in $C^{\star i}$, the corresponding lifted edges belong to exactly one 2-cube in $C^\star$ and so lie in $\partial C^\star$. Consider now a d-maximum class $C^\star \subseteq \{0,1\}^n$ for $d>2$, and make the inductive assumption that the projection $C = \Pi_{[n-1]}(C^\star)$ has two of each type of $(d-1)$-cube, and that the reduction $C' = C^{\star n}$ has two of each type of $(d-2)$-cube, in their boundaries. Pick $d-1$ colors $I \subseteq [n]$. If $n \in I$ then consider two $(d-2)$-cubes colored by $I \backslash \{x_n\}$ in $\partial C'$. By the same argument as in the base case, these lift to two I-colored cubes in $\partial C^\star$. If $n \notin I$ then ∂C contains two I-colored $(d-1)$-cubes. For each cube, if the cube is contained in C' then it has two lifts one of which is contained in $\partial C^\star$, otherwise its unique lift is contained in $\partial C^\star$. Therefore $\partial C^\star$ contains at least two I-colored cubes. ■

Example 6 *The one-inclusion graph of a 2-maximum concept class in the 4-cube is depicted in Figure 5(a), along with its boundary of edges in Figure 5(b). Note that all four colors are represented by exactly two boundary edges in this case.*

Having a large boundary is an important property of maximum classes that does not follow from contractibility.

Example 7 *Take a* 2*-simplex with vertices* A,B,C*. Glue the edges AB to AC to form a cone. Next glue the end loop BC to the edge AB . The result is a complex D with a single vertex, edge and* 2*-simplex, which is classically known as the* dunce hat *(cf. Figure 6). It is not hard to verify that D is contractible, but has no (geometric) boundary.*

Although Theorem 19 will not be explicitly used in the sequel, we return to boundaries of maximum complexes later.

4. Euclidean Arrangements

Definition 20 *A* linear arrangement *is a collection of* $n \geq d$ *oriented hyperplanes in* $\mathbb{R}^d$*. Each region or cell in the complement of the arrangement is naturally associated with a concept in* $\{0,1\}^n$*; the side of the* i^{th} *hyperplane on which a cell falls determines the concept's* i^{th} *component. A* simple arrangement *is a linear arrangement in which any subset of d planes has a unique point in common and all subsets of* $d+1$ *planes have an empty mutual intersection. Moreover any subset of* $k < d$ *planes meet in a plane of dimension* $d-k$*. Such a collection of n planes is also said to be in* general position*.*

Many of the familiar operations on concept classes in the n-cube have elegant analogues on arrangements.

- Projection on $[n]\backslash\{i\}$ corresponds to removing the i^{th} plane;
- The reduction C^i is the new arrangement given by the intersection of C's arrangement with the i^{th} plane; and
- The corresponding lifted reduction is the collection of cells in the arrangement that adjoin the i^{th} plane.

A k-cube in the one-inclusion graph corresponds to a collection of 2^k cells, all having a common $(d-k)$-face, which is contained in the intersection of k planes, and an edge corresponds to a pair of cells which have a common face on a single plane. The following result is due to Edelsbrunner (1987).

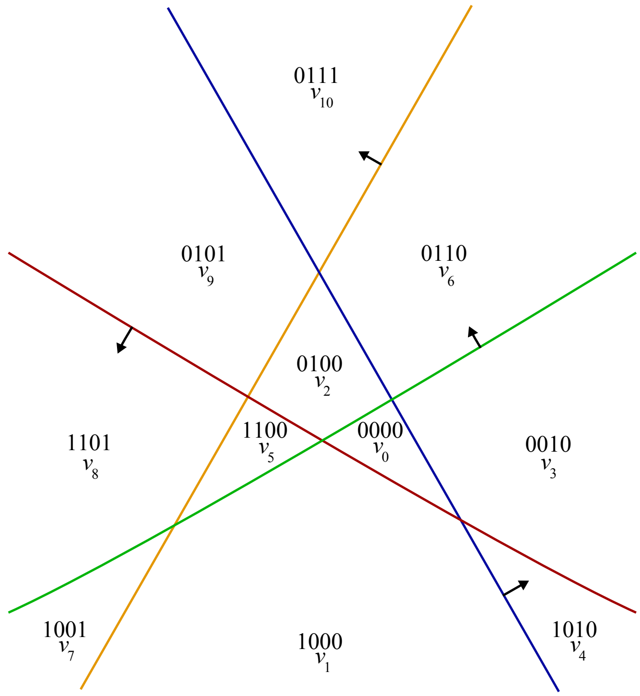
Lemma 21 *The concept class* $C \subseteq \{0,1\}^n$ *induced by a simple linear arrangement of n planes in* $\mathbb{R}^d$ *is d-maximum.*

Proof Note that C has VC dimension at most d, since general position is invariant to projection, that is, no $d+1$ planes are shattered. Since C is the union of a d-complete collection of cubes (every cell contains d-intersection points in its boundary) it follows that C is d-maximum (Rubinstein et al., 2009). ■

Example 8 *Consider the simple linear arrangement in* $\mathbb{R}^2$ *shown in Figure 7(b). The given labeling of its cells map to the concept class in the* 4*-cube enumerated in Figure 7(a) with one-inclusion graph shown in Figure 5(a). This class is maximum with VC-dimension* 2*.*

	x_1	x_2	x_3	x_4
v_0	0	0	0	0
v_1	1	0	0	0
v_2	0	1	0	0
v_3	0	0	1	0
v_4	1	0	1	0
v_5	1	1	0	0
v_6	0	1	1	0
v_7	1	0	0	1
v_8	1	1	0	1
v_9	0	1	0	1
v_{10}	0	1	1	1

(a)



(b)

Figure 7: (a) The enumeration of the 2-maximum class in $\{0,1\}^4$ in Figure 5(a) and (b) a simple linear line arrangement corresponding to the class, with each cell corresponding to a unique vertex.

Corollary 22 *Let A be a simple linear arrangement of n hyperplanes in $\mathbb{R}^d$ with corresponding d-maximum $C \subseteq \{0,1\}^n$. The intersection of A with a generic hyperplane corresponds to a $(d-1)$-maximum class $C' \subseteq C$. In particular if all d-intersection points of A lie to one side of the generic hyperplane, then C' lies on the boundary of C; and ∂C is the disjoint union of two $(d-1)$-maximum sub-classes.*

Proof The intersection of A with a generic hyperplane is again a simple arrangement of n hyperplanes but now in $\mathbb{R}^{d-1}$. Hence by Lemma 21 C' is a $(d-1)$-maximum class in the n-cube. $C' \subseteq C$ since the adjacency relationships on the cells of the intersection are inherited from those of A.

Suppose that all d-intersections in A lie in one half-space of the generic hyperplane. C' is the union of a $(d-1)$-complete collection. We claim that each of these $(d-1)$-cubes is a face of exactly one d-cube in C and is thus in ∂C. A $(d-1)$-cube in C' corresponds to a line in A where $d-1$ planes mutually intersect. The $(d-1)$-cube is a face of a d-cube in C iff this line is further intersected by a d^{th} plane. This occurs for exactly one plane, which is closest to the generic hyperplane along this intersection line. For once the d-intersection point is reached, when following along the line away from the generic plane, a new cell is entered. This verifies the second part of the result.

Consider two parallel generic hyperplanes h_1, h_2 such that all d-intersection points of A lie in between them. We claim that each $(d-1)$-cube in ∂C is in exactly one of the concept classes induced by the intersection of A with h_1 and A with h_2. Consider an arbitrary $(d-1)$-cube in ∂C. As before this cube corresponds to a region of a line formed by a mutual intersection of $d-1$ planes. Moreover this region is a ray, with one end-point at a d-intersection. Because the ray begins at a point between the generic hyperplanes h_1, h_2, it follows that the ray must cross exactly one of these. ■

Example 9 *To illustrate, consider the* 2*-maximum class in Figure 5(a) that corresponds to the simple linear arrangement in Figure 7(b). The boundary, shown in Figure 5(b) is clearly a disjoint union of two* 1*-maximum classes—in this case sticks.*

Corollary 23 *Let A be a simple linear arrangement of n hyperplanes in $\mathbb{R}^d$ and let $C \subseteq \{0,1\}^n$ be the corresponding d-maximum class. Then C considered as a cubical complex is homeomorphic to the d-ball B^d; and ∂C considered as a $(d-1)$-cubical complex is homeomorphic to the $(d-1)$-sphere S^{d-1}.*

Proof We construct a Voronoi cell decomposition corresponding to the set of d-intersection points inside a very large ball in Euclidean space. By induction on d, we claim that this is a cubical complex and the vertices and edges correspond to the class C. By induction, on each hyperplane, the induced arrangement has a Voronoi cell decomposition which is a $(d-1)$-cubical complex with edges and vertices matching the one-inclusion graph for the tail of C corresponding to the label associated with the hyperplane. It is not hard to see that the Voronoi cell defined by a d-intersection point p on this hyperplane is a d-cube. In fact, its $(d-1)$-faces correspond to the Voronoi cells for p, on each of the d hyperplanes passing through p. We also see that this d-cube has a single vertex in the interior of each of the 2^d cells of the arrangement adjacent to p. In this way, it follows that the vertices of this Voronoi cell decomposition are in bijective correspondence to the cells of the hyperplane arrangement. Finally the edges of the Voronoi cells pass through the faces in the hyperplanes. So these correspond bijectively to the edges of C, as there is one edge for each face of the hyperplanes. Using a very large ball, containing all the d-intersection points, the boundary faces become spherical cells. In fact, these form a spherical Voronoi cell decomposition, so it is easy to replace these by linear ones by taking the convex hull of their vertices. So a piecewise linear cubical complex $\mathbf{C}$ is constructed, which has one-skeleton (graph consisting of all vertices and edges) isomorphic to the one-inclusion graph for C.

Finally we want to prove that $\mathbf{C}$ is homeomorphic to B^d. This is quite easy by construction. For we see that $\mathbf{C}$ is obtained by dividing up B^d into Voronoi cells and replacing the spherical boundary cells by linear ones, using convex hulls of the boundary vertices. This process is clearly given by a homeomorphism by projection. In fact, the homeomorphism preserves the PL-structure so is a PL homeomorphism. ■

Example 10 *Consider again the one-inclusion graph in Figure 5(a) corresponding to a* 2*-maximum concept class in the* 4*-cube. It is trivial to see via inspection that this class, when viewed as a simplicial complex, is homeomorphic to a disc; similarly its boundary, highlighted in Figure 5(b), is homeomorphic to a circle.*

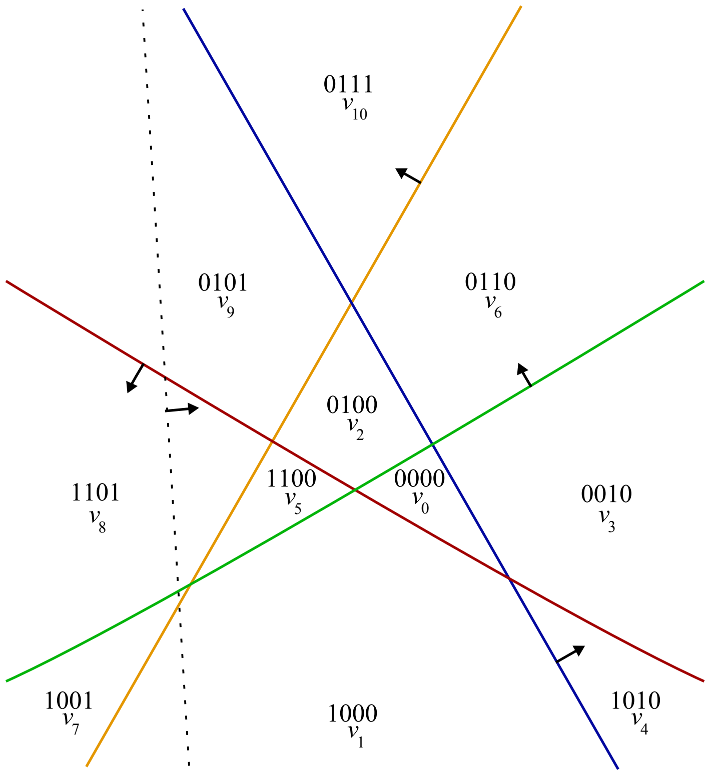


Figure 8: The simple linear line arrangement from Figure 7(b) corresponding to the concept class enumerated in Figure 7(a) and visualized in Figure 5(a). The arrangement is in the process of being swept by the dashed line.

The following example demonstrates that not all maximum classes of VC-dimension d are homeomorphic to the d-ball. The key to such examples is branching.

Example 11 *A simple linear arrangement in* $\mathbb{R}$ *corresponds to points on the line—cells are simply intervals between these points and so corresponding* 1*-maximum classes are sticks. Any tree that is not a stick can therefore not be represented as a simple linear arrangement in* $\mathbb{R}$ *and is also not homeomorphic to the* 1*-ball which is simply the interval* $[-1,1]$.

As Kuzmin and Warmuth (2007) did previously, consider a generic hyperplane h sweeping across a simple linear arrangement A. h begins with all d-intersection points of A lying in its positive half-space $\mathcal{H}_+$. The concept corresponding to cell c is peeled from C when $|\mathcal{H}_+ \cap c| = 1$, that is, h crosses the last d-intersection point adjoining c. At any step in the process, the result of peeling j vertices from C to reach C_j, is captured by the arrangement $\mathcal{H}_+ \cap A$ for the appropriate h.

Example 12 *Figure 7(a) enumerates the 11 vertices of a* 2*-maximum class in the* 4*-cube. Figures 8 and 5(a) display a hyperplane arrangement in Euclidean space and its Voronoi cell decomposition, corresponding to this maximum class. In this case, sweeping the vertical dashed line across the arrangement corresponds to a partial corner peeling of the concept class with peeling sequence* v_7, v_8, v_5, v_9, v_2, v_0. *What remains is the* 1*-maximum stick* $\{v_1, v_3, v_4, v_6, v_{10}\}$.

Next we resolve the first half of Kuzmin and Warmuth (2007, Conjecture 1).

Theorem 24 *Any d-maximum class $C \subseteq \{0,1\}^n$ corresponding to a simple linear arrangement A can be corner peeled by sweeping A, and this process is a valid unlabeled compression scheme for C of size d.*

Proof We must show that as the j^{th} d-intersection point p_j is crossed, there is a corner vertex of C_{j-1} peeled away. It then follows that sweeping a generic hyperplane h across A corresponds to corner peeling C to a $(d-1)$-maximum sub-class $C' \subseteq \partial C$ by Corollary 22. Moreover C' corresponds to a simple linear arrangement of n hyperplanes in $\mathbb{R}^{d-1}$.

We proceed by induction on d, noting that for $d = 1$ corner peeling is trivial. Consider h as it approaches the j^{th} d-intersection point p_j. The d planes defining this point intersect h in a simple arrangement of hyperplanes on h. There is a compact cell Δ for the arrangement on h, which is a d-simplex[3] and shrinks to a point as h passes through p_j. We claim that the cell c for the arrangement A, whose intersection with h is Δ, is a corner vertex v_j of C_{j-1}. Consider the lines formed by intersections of $d-1$ of the d hyperplanes, passing through p_j. Each is a segment starting at p_j and ending at h without passing through any other d-intersection points. So all faces of hyperplanes adjacent to c meet h in faces of Δ. Thus, there are no edges in C_{j-1} starting at the vertex corresponding to p_j, except for those in the cube C'_{j-1}, which consists of all cells adjacent to p_j in the arrangement A. So c corresponds to a corner vertex v_j of the d-cube C'_{j-1} in C_{j-1}. Finally, just after the simplex is a point, c is no longer in $\mathcal{H}_+$ and so v_j is corner peeled from C_{j-1}.

Theorem 16 completes the proof that this corner peeling of C constitutes unlabeled compression. ■

Corollary 25 *The sequence of cubes $C'_0, \ldots, C'_{|C|}$, removed when corner peeling by sweeping simple linear arrangements, is of non-increasing dimension. In fact, there are $\binom{n}{d}$ cubes of dimension d, then $\binom{n}{d-1}$ cubes of dimension $d-1$, etc.*

While corner peeling and min peeling share some properties in common, they are distinct procedures. Notice that sweeping produces a monotonic corner-peeling sequence, as cubes are removed in order of non-increasing dimensions.

Example 13 *Consider sweeping a simple linear arrangement corresponding to a* 2*-maximum class. After all but one* 2*-intersection point has been swept, the corresponding corner-peeled class C_t is the union of a single* 2*-cube with a* 1*-maximum stick. Min peeling applied to C_t would first peel a leaf, while sweeping must peel the* 2*-cube next.*

A second example is the class in a 3*-cube which consists of six vertices, so that two opposite vertices, for example,* 000 *and* 111 *are not included. This class cannot be corner peeled as the one-inclusion graph consists of six edges forming a single cycle. On the other hand, it has many min-peeling schemes.*

An interesting question is if a class has a corner-peeling scheme, does it always have a min-peeling scheme which is also a corner-peeling scheme? This is given as Question 50 below.

3. Δ is a topological simplex—the convex hull of $d+1$ affinely independent points in $\mathbb{R}^d$.

The next result follows from our counter-examples to Kuzmin & Warmuth's minimum degree conjecture (Rubinstein et al., 2009).

Corollary 26 *There is no constant c so that all maximal classes of VC-dimension d can be embedded into maximum classes corresponding to simple hyperplane arrangements of dimension $d+c$.*

5. Hyperbolic Arrangements

To motivate the introduction of hyperbolic arrangements, note that linear-hyperplane arrangements can be efficiently described, since each hyperplane is determined by its unit normal and distance from the origin. Similarly, a hyperbolic hyperplane is a hypersphere. So it can be parametrized by its center—a point on the ideal sphere at infinity—and its radius.[4]

However the family of hyperbolic hyperplanes has more flexibility than linear hyperplanes since there are many disjoint hyperbolic hyperplanes, whereas in the linear case only parallel hyperplanes do not meet. Thus we turn to hyperbolic arrangements to represent a larger collection of concept classes than those represented by simple linear arrangements.

We briefly discuss the Klein model of hyperbolic geometry (Ratcliffe, 1994, pg. 7). Consider the open unit ball $\mathbb{H}^k$ in $\mathbb{R}^k$. Geodesics (lines of shortest length in the geometry) are given by intersections of straight lines in $\mathbb{R}^k$ with the unit ball. Similarly planes of any dimension between 2 and $k-1$ are given by intersections of such planes in $\mathbb{R}^k$ with the unit ball. Note that such planes are completely determined by their spheres of intersection with the unit sphere S^{k-1}, which is called the ideal boundary of hyperbolic space $\mathbb{H}^k$. Note that in the appropriate metric, the ideal boundary consists of points which are infinitely far from all points in the interior of the unit ball.

We can now see immediately that a simple hyperplane arrangement in $\mathbb{H}^k$ can be described by taking a simple hyperplane arrangement in $\mathbb{R}^k$ and intersecting it with the unit ball. However we require an important additional property to mimic the Euclidean case. Namely we add the constraint that every subcollection of d of the hyperplanes in $\mathbb{H}^k$ has mutual intersection points inside $\mathbb{H}^k$, and that no $(d+1)$-intersection point lies in $\mathbb{H}^k$. We need this requirement to obtain that the resulting class is maximum.

Definition 27 *A* simple hyperbolic d-arrangement *is a collection of n hyperplanes in $\mathbb{H}^k$ with the property that every sub-collection of d hyperplanes mutually intersect in a $(k-d)$-dimensional hyperbolic plane, and that every sub-collection of $d+1$ hyperplanes mutually intersect as the empty set.*

Corollary 28 *The concept class C corresponding to a simple d-arrangement of hyperbolic hyperplanes in $\mathbb{H}^k$ is d-maximum in the k-cube.*

Proof The result follows by the same argument as before. Projection cannot shatter any $(d+1)$-cube and the class is a complete union of d-cubes, so is d-maximum. ■

The key to why hyperbolic arrangements represent many new maximum classes is that they allow flexibility of choosing d and k independently. This is significant because the unit ball can be

4. Note also that hyperbolic hyperplanes are 'linear' in the sense that they are filled by a family of geodesics, which are shortest paths or lines in the hyperbolic metric.

chosen to miss much of the intersections of the hyperplanes in Euclidean space. Note that the new maximum classes are embedded in maximum classes induced by arrangements of linear hyperplanes in Euclidean space.

A simple example is any 1-maximum class. It is easy to see that this can be realized in the hyperbolic plane by choosing an appropriate family of lines and the unit ball in the appropriate position. In fact, we can choose sets of pairs of points on the unit circle, which will be the intersections with our lines. So long as these pairs of points have the property that the smaller arcs of the circle between them are disjoint, the lines will not cross inside the disk and the desired 1-maximum class will be represented.

Corner-peeling maximum classes represented by hyperbolic-hyperplane arrangements proceeds by sweeping, just as in the Euclidean case. Note first that intersections of the hyperplanes of the arrangement with the moving hyperplane appear precisely when there is a first intersection at the ideal boundary. Thus it is necessary to slightly perturb the collection of hyperplanes to ensure that only one new intersection with the moving hyperplane occurs at any time. Note also that new intersections of the sweeping hyperplane with the various lower dimensional planes of intersection between the hyperplanes appear similarly at the ideal boundary. The important claim to check is that the intersection at the ideal boundary between the moving hyperplane and a lower dimensional plane, consisting entirely of d intersection points, corresponds to a corner-peeling move. We include two examples to illustrate the validity of this claim.

Example 14 *In the case of a* 1*-maximum class coming from disjoint lines in* $\mathbb{H}^2$*, a cell can disappear when the sweeping hyperplane meets a line at an ideal point. This cell is indeed a vertex of the tree, that is, a corner-vertex.*

Example 15 *Assume that we have a family of* 2*-planes in the unit* 3*-ball which meet in pairs in single lines, but there are no triple points of intersection, corresponding to a* 2*-maximum class. A corner-peeling move occurs when a region bounded by two half disks and an interval disappears, in the positive half space bounded by the sweeping hyperplane. Such a region can be visualized by taking a slice out of an orange. Note that the final point of contact between the hyperplane and the region is at the end of a line of intersection between two planes on the ideal boundary.*

We next observe that sweeping by generic hyperbolic hyperplanes induces corner peeling of the corresponding maximum class, extending Theorem 24. As the generic hyperplane sweeps across hyperbolic space, not only do swept cells correspond to corners of d-cubes but also to corners of lower dimensional cubes as well. Moreover, the order of the dimensions of the cubes which are corner peeled can be arbitrary—lower dimensional cubes may be corner peeled before all the higher dimensional cubes are corner peeled. This is in contrast to Euclidean sweepouts (cf. Corollary 25). Similar to Euclidean sweepouts, hyperbolic sweepouts correspond to corner peeling and not min peeling.

Theorem 29 *Any d-maximum class $C \subseteq \{0,1\}^n$ corresponding to a simple hyperbolic d-arrangement A can be corner peeled by sweeping A with a generic hyperbolic hyperplane.*

Proof We follow the same strategy of the proof of Theorem 24. For sweeping in hyperbolic space $\mathbb{H}^k$, the generic hyperplane h is initialized as tangent to $\mathbb{H}^k$. As h is swept across $\mathbb{H}^k$, new intersections appear with A just after h meets the non-empty intersection of a subset of hyperplanes of A with

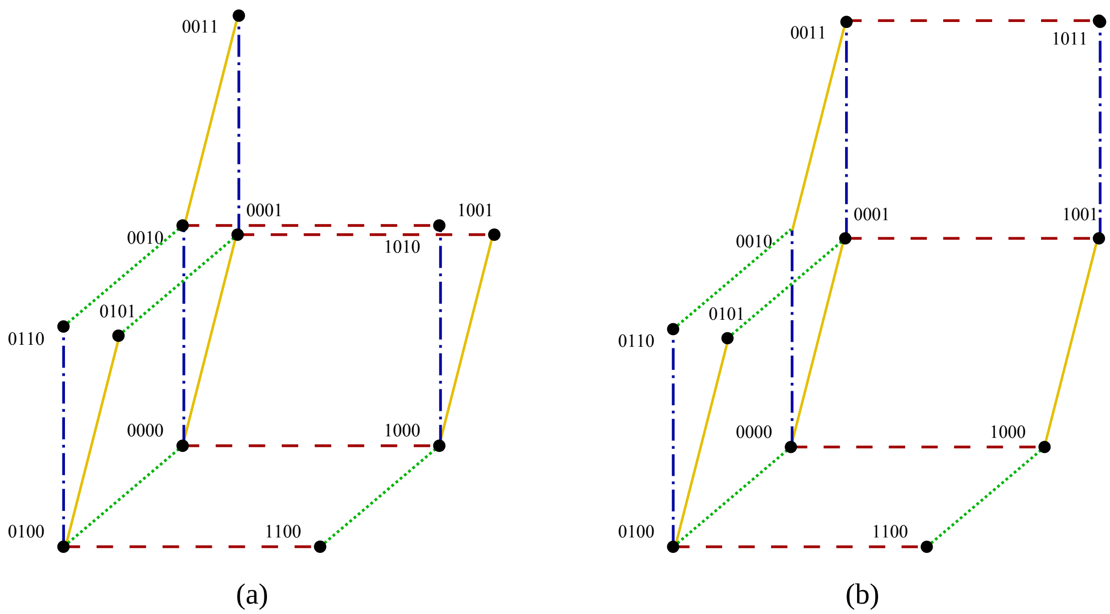
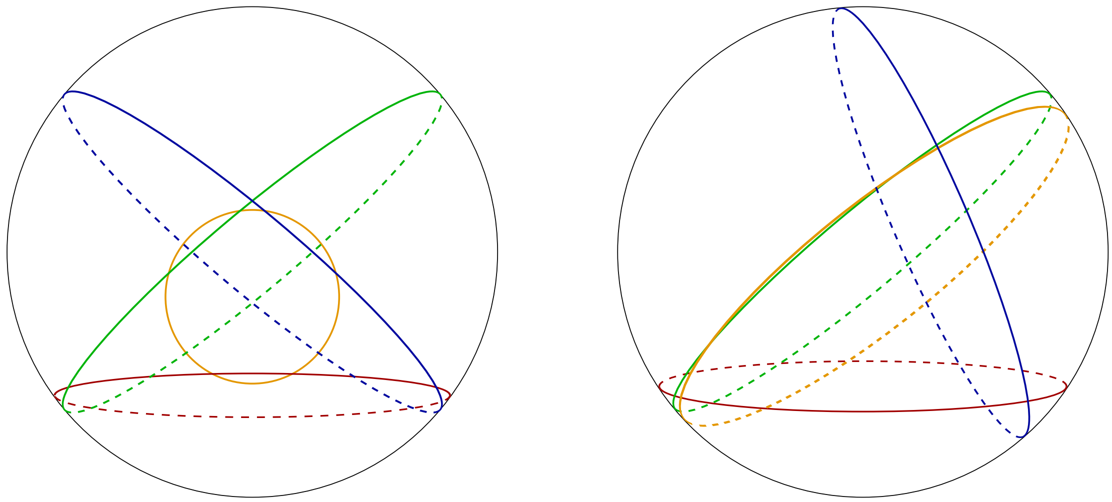


Figure 9: 2-maximum classes in $\{0,1\}^4$ that can be represented as hyperbolic arrangements but not as Euclidean arrangements.

the ideal boundary. Each d-cube C' in C still corresponds to the cells adjacent to the intersection $I_{C'}$ of d hyperplanes. But now $I_{C'}$ is a $(k-d)$-dimensional hyperbolic hyperplane. A cell c adjacent to $I_{C'}$ is corner peeled precisely when h last intersects c at a point of $I_{C'}$ at the ideal boundary. As for simple linear arrangements, the general position of $A \cup \{h\}$ ensures that corner-peeling events never occur simultaneously. For the case $k = d+1$, as for the simple linear arrangements just prior to the corner peeling of c, $\mathcal{H}_+ \cap c$ is homeomorphic to a $(d+1)$-simplex with a missing face on the ideal boundary. And so as in the simple linear case, this d-intersection point corresponds to a corner d-cube. In the case $k > d+1$, $\mathcal{H}_+ \cap c$ becomes a $(d+1)$-simplex (as before) multiplied by $\mathbb{R}^{k-d-1}$. If $k = d$, then the main difference is just before corner peeling of c, $\mathcal{H}_+ \cap c$ is homeomorphic to a k-simplex which may be either closed (hence in the interior of $\mathbb{H}^k$) or with a missing face on the ideal boundary. The rest of the argument remains the same, except for one important observation.

Although swept corners in hyperbolic arrangements can be of cubes of differing dimensions, these dimensions never exceed d and so the proof that sweeping simple linear arrangements induces d-compression schemes is still valid. ■

Example 16 *Constructed with lifting, Figure 9 completes the enumeration, up to symmetry, of the* 2*-maximum classes in* $\{0,1\}^4$ *begun with Example 12. These cases cannot be represented as simple Euclidean linear arrangements, since their boundaries do not satisfy the condition of Corollary 23 but can be represented as hyperbolic arrangements as in Figure 10. Figures 11(a) and 11(b) display the sweeping of a general hyperplane across the former arrangement and the corresponding corner peeling. Notice that the corner-peeled cubes' dimensions decrease and then increase.*



(a) (b)

Figure 10: Hyperbolic-hyperplane arrangements corresponding to the classes in Figure 9. In both cases the four hyperbolic planes meet in 6 straight line segments (not shown). The planes' colors correspond to the edges' colors in Figure 9.

Corollary 30 *There is no constant c so that all maximal classes of VC-dimension d can be embedded into maximum classes corresponding to simple hyperbolic-hyperplane arrangements of VC-dimension $d+c$.*

This result follows from our counter-examples to Kuzmin & Warmuth's minimum degree conjecture (Rubinstein et al., 2009).

Corollary 28 gives a proper superset of simple linear-hyperplane arrangement-induced maximum classes as hyperbolic arrangements. We will prove in Section 7 that all maximum classes can be represented as PL-hyperplane arrangements in a ball. These are the topological analogue of hyperbolic-hyperplane arrangements. If the boundary of the ball is removed, then we obtain an arrangement of PL hyperplanes in Euclidean space.

6. Infinite Euclidean and Hyperbolic Arrangements

We consider a simple example of an infinite maximum class which admits corner peeling and a compression scheme analogous to those of previous sections.

Example 17 *Let $\mathcal{L}$ be the set of lines in the plane of the form $L_{2m} = \{(x,y) \mid x = m\}$ and $L_{2n+1} = \{(x,y) \mid y = n\}$ for $m,n \in \mathbb{N}$. Let v_{00}, v_{0n}, v_{m0}, and v_{mn} be the cells bounded by the lines $\{L_2, L_3\}$, $\{L_2, L_{2n+1}, L_{2n+3}\}$, $\{L_{2m}, L_{2m+2}, L_3\}$, and $\{L_{2m}, L_{2m+2}, L_{2n+1}, L_{2n+3}\}$, respectively. Then the cubical complex C, with vertices v_{mn}, can be corner peeled and hence compressed, using a sweepout by the lines $\{(x,y) \mid x + (1+\varepsilon)y = t\}$ for $t \geq 0$ and any small fixed irrational $\varepsilon > 0$. C is a 2-maximum class and the unlabeled compression scheme is also of size 2.*

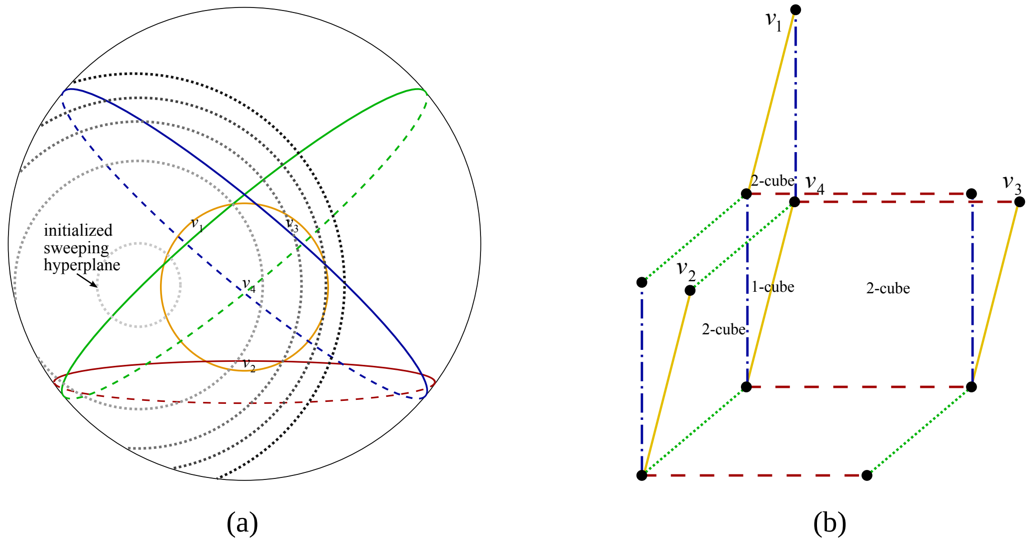


Figure 11: (a) The simple hyperbolic arrangement corresponding to the 2-maximum class in $\{0,1\}^4$ of Figure 9(a)—as shown in Figure 10(a)—with a generic sweeping hyperplane shown in several positions before and after it sweeps past four cells; and (b) the class with the first four corner-vertices peeled by the hyperbolic arrangement sweeping. Notice that three 2-cubes are peeled, then a 1-cube (all shown) followed by 2-cubes.

To verify the properties of this example, notice that sweeping as specified corresponds to corner peeling the vertex v_{00}, then the vertices v_{10}, v_{01}, then the remaining vertices v_{mn}. The lines $x + (1 + \varepsilon)y = t$ are generic as they pass through only one intersection point of $\mathcal{L}$ at a time. Additionally, representing v_{00} by $\emptyset$, v_{0n} by $\{L_{2n+1}\}$, v_{m0} by $\{L_{2m}\}$ and v_{mn} by $\{L_{2m}, L_{2n+1}\}$ constitutes a valid unlabeled compression scheme. Note that the compression scheme is associated with sweeping across the arrangement in the direction of decreasing t. This is necessary to pick up the boundary vertices of C last in the sweepout process, so that they have either singleton representatives or the empty set. In this way, similar to Kuzmin and Warmuth (2007), we obtain a compression scheme so that every labeled sample of size 2 is associated with a unique concept in C, which is consistent with the sample. On the other hand to obtain corner peeling, we need the sweepout to proceed with t increasing so that we can begin at the boundary vertices of C.

In concluding this brief discussion, we note that many infinite collections of simple hyperbolic hyperplanes and Euclidean hyperplanes can also be corner peeled and compressed, even if intersection points and cells accumulate. However a key requirement in the Euclidean case is that the concept class C has a non-empty boundary, when considered as a cubical complex. An easy approach is to assume that all the d-intersections of the arrangement lie in a half-space. Moreover, since the boundary must also admit corner peeling, we require more conditions, similar to having all the intersection points lying in an octant.

Example 18 *In $\mathbb{R}^3$, choose the family of planes $\mathcal{P}$ of the form $P_{3n+i} = \{\mathbf{x} \in \mathbb{R}^3 \mid x_{i+1} = 1 - 1/n\}$ for $n \geq 1$ and $i \in \{0,1,2\}$. A corner-peeling scheme is induced by sweeping a generic plane $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + \alpha x_2 + \beta x_3 = t\}$ across the arrangement, where t is a parameter and $1, \alpha, \beta$ are algebraically* *independent (in particular, no integral linear combination is rational) and* α, β *are both close to* 1. *This example has similar properties to Example 17: the compression scheme is again given by decreasing* t *whereas corner peeling corresponds to increasing* t. *Note that cells shrink to points, as* $\mathbf{x} \to \mathbf{1}$ *and the volume of cells converge to zero as* $n \to \infty$, *or equivalently any* $x_i \to 1$.

Example 19 *In the hyperbolic plane* $\mathbb{H}^2$, *represented as the unit circle centered at the origin in* $\mathbb{R}^2$, *choose the family of lines* $\mathcal{L}$ *given by* $L_{2n} = \{(x,y) \mid x = 1 - 1/n\}$ *and* $L_{2n+1} = \{(x,y) \mid x + ny = 1\}$, *for* $n \geq 1$. *This arrangement has corner peeling and compression schemes given by sweeping across* $\mathcal{L}$ *using the generic line* $\{y = t\}$.

7. Piecewise-Linear Arrangements

PL hyperplanes have the advantage that they can be easily manipulated, by cutting and pasting or isotoping part of a hyperplane to a new position, keeping the rest of the hyperplane fixed. However a disadvantage is that there is no simple way of describing a PL hyperplane, similar to the parametrizations of either linear or hyperbolic hyperplanes. The methods of proof of our main results about representing maximum classes and corner peeling, require PL-hyperplane arrangements. We conjecture that PL-hyperplane arrangements are equivalent to hyperbolic ones. This would give an interesting geometric approach of forming all maximum classes as simple hyperbolic arrangements.

A *PL hyperplane* is the image of a proper piecewise-linear homeomorphism from the $(k-1)$-ball B^{k-1} into B^k, that is, the inverse image of the boundary S^{k-1} of the k-ball is S^{k-2} (Rourke and Sanderson, 1982). A *simple PL d-arrangement* is an arrangement of n PL hyperplanes such that every subcollection of j hyperplanes meet transversely in a $(k-j)$-dimensional PL plane for $2 \leq j \leq d$ and every subcollection of $d+1$ hyperplanes are disjoint.

Corollary 31 *The concept class* C *corresponding to a simple* d*-arrangement of PL hyperplanes in* B^k *is* d*-maximum in the* k*-cube.*

Proof The result follows by the same argument as in the linear or hyperbolic cases. Projection cannot shatter any $(d+1)$-cube and the class is a complete union of d-cubes, so is d-maximum. ∎

7.1 Maximum Classes are Represented by Simple PL-Hyperplane Arrangements

Our aim is to prove the following theorem by a series of steps.

Theorem 32 *Every* d*-maximum class* $C \subseteq \{0,1\}^n$ *can be represented by a simple arrangement of PL hyperplanes in an* n*-ball. Moreover the corresponding simple arrangement of PL hyperspheres in the* $(n-1)$*-sphere also represents* C*, so long as* $n > d+1$.

7.1.1 EMBEDDING A d-MAXIMUM CUBICAL COMPLEX IN THE n-CUBE INTO AN n-BALL

We begin with a d-maximum cubical complex $C \subseteq \{0,1\}^n$ embedded into $[0,1]^n$. This gives a natural embedding of C into $\mathbb{R}^n$. Take a small regular neighborhood $\mathcal{N}$ of C so that the boundary $\partial\mathcal{N}$ of $\mathcal{N}$ will be a closed manifold of dimension $n-1$. Note that $\mathcal{N}$ is contractible because it has a deformation retraction onto C and so $\partial\mathcal{N}$ is a homology $(n-1)$-sphere (by a standard, well-known

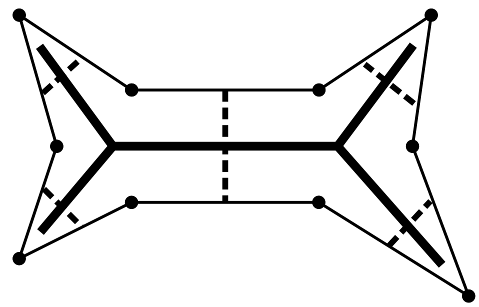


Figure 12: A 1-maximum class (thick solid lines) with its fattening (thin solid lines with points), bisecting sets (dashed lines) and induced complementary cells.

argument from topology due to Mazur 1961). Our aim is to prove that $\partial\mathcal{N}$ is an $(n-1)$-sphere and $\mathcal{N}$ is an n-ball. There are two ways of proving this: show that $\partial\mathcal{N}$ is simply connected and invoke the well-known solution to the generalized Poincaré conjecture (Smale, 1961), or use the cubical structure of the n-cube and C to directly prove the result. We adopt the latter approach, although the former works fine. The advantage of the latter is that it produces the required hyperplane arrangement, not just the structures of $\partial\mathcal{N}$ and $\mathcal{N}$.

7.1.2 BISECTING SETS

For each color i, there is a hyperplane P_i in $\mathbb{R}^n$ consisting of all vectors with i^{th} coordinate equal to $1/2$. We can easily arrange the choice of regular neighborhood $\mathcal{N}$ of C so that $\mathcal{N}_i = P_i \cap \mathcal{N}$ is a regular neighborhood of $C \cap P_i$ in P_i. (We call $\mathcal{N}_i$ a *bisecting set* as it intersects C along the 'center' of the reduction in the i^{th} coordinate direction, see Figure 12.) But then since $C \cap P_i$ is a cubical complex corresponding to the reduction C^i, by induction on n, we can assert that $\mathcal{N}_i$ is an $(n-1)$-ball. Similarly the intersections $\mathcal{N}_i \cap \mathcal{N}_j$ can be arranged to be regular neighborhoods of $(d-2)$-maximum classes and are also balls of dimension $n-2$, etc. In this way, we see that if we can show that $\mathcal{N}$ is an n-ball, then the induction step will be satisfied and we will have produced a PL-hyperplane arrangement (the system of $\mathcal{N}_i$ in $\mathcal{N}$) in a ball.

7.1.3 SHIFTING

To complete the induction step, we use the technique of shifting (Alon, 1983; Frankl, 1983; Haussler, 1995). In our situation, this can be viewed as the converse of lifting. Namely if a color i is chosen, then the cubical complex C has a lifted reduction C' consisting of all d-cubes containing the i^{th} color. By shifting, we can move down any of the lifted components, obtained by splitting C open along C', from the level $x_i = 1$ to the level $x_i = 0$, to form a new cubical complex $C^\star$. We claim that the regular neighborhood of C is a ball if and only if the same is true for $C^\star$. But this is quite straightforward, since the operation of shifting can be thought of as sliding components of C, split open along C', continuously from level $x_i = 1$ to $x_i = 0$. So there is an isotopy of the attaching maps of the components onto the lifted reduction, using the product structure of the latter. It is easy then to check that this does not affect the homeomorphism type of the regular neighborhood and so the claim of shift invariance is proved.

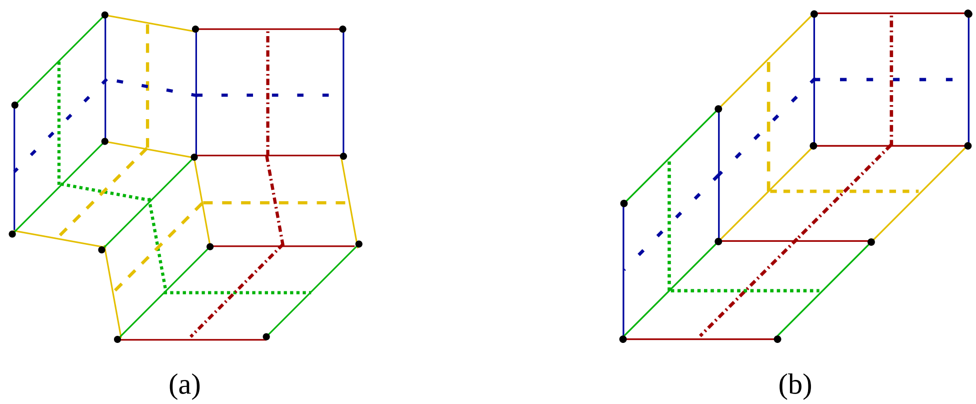


Figure 13: The (a) top and (b) bottom of Figure 9(b) (i.e., the 2-cubes seen from above and below, respectively) both give part of the boundary of a regular neighborhood in $\mathbb{R}^3$.

But repeated shifting finishes with the downwards closed maximum class consisting of all vertices in the n-cube with at most d coordinates being one and the remaining coordinates all being zero. It is easy to see that the corresponding cubical complex $\tilde{C}$ is star-like, that is, contains all the straight line segments from the origin to any point in $\tilde{C}$. If we choose a regular neighborhood $\tilde{\mathcal{N}}$ to also be star-like, then it is obvious that $\tilde{\mathcal{N}}$ is an n-ball, using radial projection. Hence our induction is complete and we have shown that any d-maximum class in $\{0,1\}^n$ can be represented by a family of PL hyperplanes in the n-ball.

7.1.4 IDEAL BOUNDARY

To complete the proof of Theorem 32, let $\partial\mathcal{N} = S^{n-1}$ denote the boundary of the n-ball $\mathcal{N}$ constructed above (cf. Figure 13). Each PL hyperplane intersects this sphere in a PL hypersphere of dimension $n-2$. It remains to show this arrangement of hyperspheres gives the same cubical complex as C, unless $n = d+1$.

Suppose that $n > d+1$. Then it is easy to see that each cell c in the complement of the PL-hyperplane arrangement in $\mathcal{N}$ has part of its boundary on the ideal boundary $\partial\mathcal{N}$. Let $\partial c = \partial c_+ \cup \partial c_-$, where ∂c_+ is the intersection of c with the ideal boundary and ∂c_- is the closure of $\partial c \setminus \partial c_+$.

It is now straightforward to verify that the face structure of ∂c_+ is equivalent to the face structure of ∂c_-. Note that any family of at most d PL hyperplanes meet in a PL ball properly embedded in $\mathcal{N}$. Since $n > d+1$, the smallest dimension of such a ball is two, and hence its boundary is connected. Then ∂c_- has faces which are PL balls obtained in this way of dimension varying between $n-d$ and $n-1$. Each of these faces has boundary a PL sphere which is a face of ∂c_+. So this establishes a bijection between the faces of ∂c_+ and those of ∂c_-. It is easy to check that the cubical complexes corresponding to the PL hyperplanes and to the PL hyperspheres are the same.

Note that if $n = d+1$, then any d-maximum class $C \subseteq \{0,1\}^{d+1}$ is obtained by taking all the d-faces of the $(d+1)$-cube which contain a particular vertex. So C is a d-ball and the ideal boundary of $\mathcal{N}$ is a d-sphere. The cubical complex associated with the ideal boundary is the double $2C$ of C, that is, two copies of C glued together along their boundaries. The proof of Theorem 32 is now complete.

Example 20 *Consider the unique bounded below* 2*-maximum class* $\tilde{C} \subseteq \{0,1\}^5$*. We claim that* $\tilde{C}$ *cannot be realized as an arrangement of PL hyperplanes in the* 3*-ball* B^3*. Note that our method gives* $\tilde{C}$ *as an arrangement in* B^5 *and this example shows that* B^4 *is the best one might hope for in terms of dimension of the hyperplane arrangement.*

For suppose that $\tilde{C}$ *could be realized by any PL-hyperplane arrangement in* B^3*. Then clearly we can also embed* $\tilde{C}$ *into* B^3*. The vertex* $v_0 = \{0\}^5$ *has link given by the complete graph* K *on* 5 *vertices in* $\tilde{C}$*. (By* link*, we mean the intersection of the boundary of a small ball in* B^3 *centered at* v_0 *with* $\tilde{C}$*.) But as is well known,* K *is not planar, that is, cannot be embedded into the plane or* 2*-sphere. This contradiction shows that no such arrangement is possible.*

7.2 Maximum Classes with Manifold Cubical Complexes

We prove a partial converse to Corollary 23: if a d-maximum class has a ball as cubical complex, then it can always be realized by a simple PL-hyperplane arrangement in $\mathbb{R}^d$.

Theorem 33 *Suppose that* $C \subseteq \{0,1\}^n$ *is a* d*-maximum class. Then the following properties of* C*, considered as a cubical complex, are equivalent:*

(i) There is a simple arrangement A *of* n *PL hyperplanes in* $\mathbb{R}^d$ *which represents* C*.*

(ii) C *is homeomorphic to the* d*-ball.*

(iii) C *is a* d*-manifold with boundary.*

Proof To prove (i) implies (ii), we can use exactly the same argument as Corollary 23. Next (ii) trivially implies (iii). So it remains to show that (iii) implies (i). The proof proceeds by double induction on n, d. The initial cases where either $d = 1$ or $n = 1$ are very easy.

Assume that C is a manifold. Let p denote the i^{th} coordinate projection. Then $p(C)$ is obtained by collapsing $C^i \times [0,1]$ onto C^i, where C^i is the reduction. As before, let P_i be the linear hyperplane in $\mathbb{R}^n$, where the i^{th} coordinate takes value $1/2$. Viewing C as a manifold embedded in the n-cube, since P_i intersects C transversely, we see that $C^i \times \{1/2\}$ is a proper submanifold of C. But it is easy to check that collapsing $C^i \times [0,1]$ to C^i in C produces a new manifold which is again homeomorphic to C. (The product region $C^i \times [0,1]$ in C can be expanded to a larger product region $C^i \times [-\varepsilon, 1+\varepsilon]$ and so collapsing shrinks the larger region to one of the same homeomorphism type, namely $C^i \times [-\varepsilon, \varepsilon]$). So we conclude that the projection $p(C)$ is also a manifold. By induction on n, it follows that there is a PL-hyperplane arrangement A, consisting of $n-1$ PL hyperplanes in B^d, which represents $p(C)$.

Next, observe that the reduction C^i can be viewed as a properly embedded submanifold M in B^d, where M is a union of some of the $(d-1)$-dimensional faces of the Voronoi cell decomposition corresponding to A, described in Corollary 23. By induction on d, we conclude that C^i is also represented by n PL hyperplanes in B^{d-1}. But then since condition (i) implies (ii), it follows that M is PL homeomorphic to B^{d-1}, since the underlying cubical complex for C^i is a $(d-1)$-ball. So it follows that $A \cup \{M\}$ is a PL-hyperplane arrangement in B^d representing C. This completes the proof that condition (iii) implies (i). ∎

8. Corner Peeling 2-Maximum Classes

We give a separate treatment for the case of 2-maximum classes, since it is simpler than the general case and shows by a direct geometric argument, that representation by a simple family of PL hyperplanes or PL hyperspheres implies a corner-peeling scheme.

Theorem 34 *Every* 2*-maximum class can be corner peeled.*

Proof By Theorem 32, we can represent any 2-maximum class $C \subseteq \{0,1\}^n$ by a simple family of PL hyperspheres $\{S_i\}$ in S^{n-1}. Every pair of hyperspheres S_i, S_j intersects in an $(n-3)$-sphere S_{ij} and there are no intersection points between any three of these hyperspheres. Consider the family of spheres S_{ij}, for i fixed. These are disjoint hyperspheres in S_i so we can choose an innermost one S_{ik} which bounds an $(n-2)$-ball B_1 in S_i not containing any other of these spheres. Moreover there are two balls B_2, B_3 bounded by S_{ik} on S_k. We call the two $(n-1)$-balls Q_2, Q_3 bounded by $B_1 \cup B_2$, $B_1 \cup B_3$ respectively in S^{n-1}, which intersect only along B_1, *quadrants*.

Assume B_2 is innermost on S_k. Then the quadrant Q_2 has both faces B_1, B_2 innermost. It is easy to see that such a quadrant corresponds to a corner vertex in C which can be peeled. Moreover, after peeling, we still have a family of PL hyperspheres which give an arrangement corresponding to the new peeled class. The only difference is that cell Q_2 disappears, by interchanging B_1, B_2 on the corresponding spheres S_i, S_k and then slightly pulling the faces apart. (If $n = 3$, we can visualize a pair of disks on two intersecting spheres with a common boundary circle. Then peeling can be viewed as moving these two disks until they coincide and then pulling the first past the second). So it is clear that if we can repeatedly show that a quadrant can be found with two innermost faces, until all the intersections between the hyperspheres have been removed, then we will have corner peeled C to a 1-maximum class, that is, a tree. So peeling will be established.

Suppose neither of the two quadrants Q_2, Q_3 has both faces innermost. Consider Q_2 say and let $\{S_\alpha\}$ be the family of spheres intersecting the interior of the face B_2. Amongst these spheres, there is clearly at least one S_β so that the intersection $S_{k\beta}$ is innermost on S_k. But then $S_{k\beta}$ bounds an innermost ball B_4 in S_k whose interior is disjoint from all the spheres $\{S_\alpha\}$. Similarly, we see that $S_{k\beta}$ bounds a ball B_5 which is the intersection of the sphere S_β with the quadrant Q_2. We get a new quadrant bounded by $B_4 \cup B_5$ which is strictly smaller than Q_2 and has at least one innermost face. But clearly this process must terminate—we cannot keep finding smaller and smaller quadrants and so a smallest one must have both faces innermost. ■

9. Corner Peeling Finite Maximum Classes

Above, simple PL-hyperplane arrangements in the n-ball B^n are defined. For the purposes of this section, we study a slightly more general class of arrangements. Every simple arrangement is in this larger class, but the latter class has many good properties. In Example 23, a maximal class is represented by a 2-contractible hyperbolic-hyperplane arrangement. By contrast, simple hyperplane arrangements always represent maximum classes.

Definition 35 *Suppose that a finite arrangement $\mathcal{P}$ of PL hyperplanes $\{P_\alpha\}$, each properly embedded in an n-ball B^n, satisfies the following conditions:*

i. *Each k-subcollection of hyperplanes either intersects transversely in a PL plane of dimension $n-k$, or has an empty intersection; and*

ii. *The maximum number of hyperplanes which mutually intersect is $d \leq n$.*

*Then we say that the arrangement $\mathcal{P}$ is d-*contractible.

The arrangements in Definition 35 are called d-contractible because we prove later that their corresponding one-inclusion graphs are strongly contractible cubical complexes of dimension d. Moreover we now prove that the corresponding one-inclusion graphs have VC dimension exactly d.

Lemma 36 *The one-inclusion graph Γ corresponding to a d-contractible arrangement $\mathcal{P}$ has VC-dimension d.*

Proof We observe first of all, that since $\mathcal{P}$ has a subcollection of d hyperplanes which mutually intersect, the corresponding one-inclusion graph Γ has a d-subcube, when considered as a cubical complex. But then the VC dimension of Γ is clearly at least d. On the other hand, suppose that the VC dimension of Γ was greater than d. Then there is a projection of Γ which shatters some $(d+1)$-cube. But this projection can be viewed as deleting all the hyperplanes of $\mathcal{P}$ except for a subcollection of $d+1$ hyperplanes. However, by assumption, such a collection cannot have any mutual intersection points. It is easy to see that any such an arrangement has at most $2^{d+1}-1$ complementary regions and hence cannot represent the $(d+1)$-cube. This completes the proof. ■

Definition 37 *A one-inclusion graph Γ is* strongly contractible *if it is contractible as a cubical complex and moreover, all reductions and multiple reductions of Γ are also contractible.*

Definition 38 *The* complexity *of a PL-hyperplane arrangement $\mathcal{P}$ is the lexicographically ordered pair (r,s), where r is the number of regions in the complement of $\mathcal{P}$, and s is the smallest number of regions in any half space on one side of an individual hyperplane in $\mathcal{P}$.*

We allow several different hyperplanes to be used for a single sweeping process. So a hyperplane P may start sweeping across an arrangement $\mathcal{P}$. One of the half spaces defined by P can become a new ball B_+ with a new arrangement $\mathcal{P}_+$ defined by restriction of $\mathcal{P}$ to the half space B_+. Then a second generic hyperplane P' can start sweeping across this new arrangement $\mathcal{P}_+$. This process may occur several times. It is easy to see that sweeping a single generic hyperplane as in Theorem 29, applies to such a multi-hyperplane process. Below we show that a suitable multiple sweeping of a PL-hyperplane arrangement $\mathcal{P}$ gives a corner-peeling sequence of all finite maximum classes.

The following states our main theorem.

Theorem 39 *Assume that $\mathcal{P}$ is a d-contractible PL-hyperplane arrangement in the n-ball B^n. Then there is a d-corner-peeling scheme for this collection $\mathcal{P}$.*

Corollary 40 *There is no constant k so that every finite maximal class of VC-dimension d can be embedded into a maximum class of VC-dimension $d+k$.*

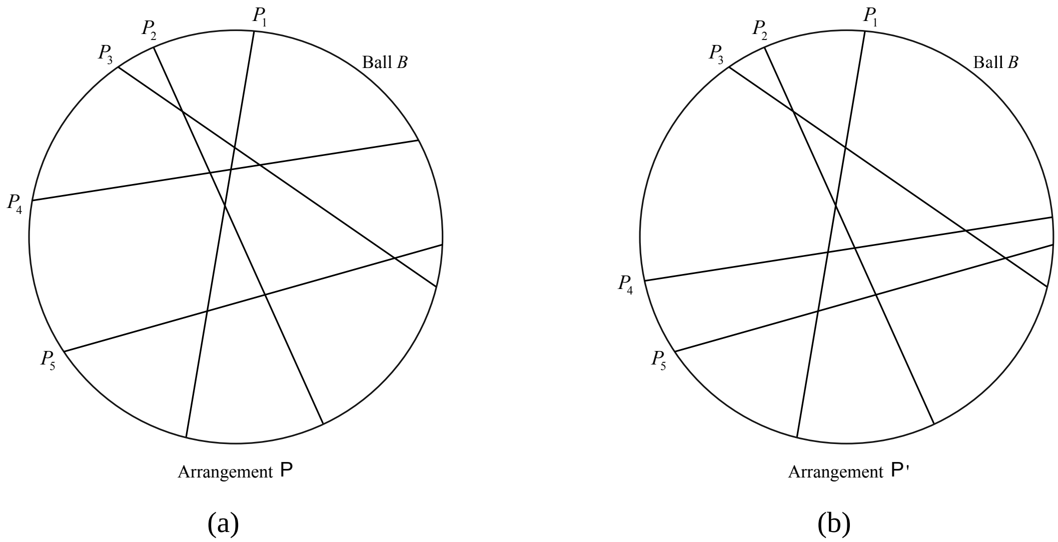


Figure 14: (a) An example PL-hyperplane arrangement $\mathcal{P}$ and (b) the result of a Pachner move of hyperplane P_4 on $\mathcal{P}$.

Proof By Theorem 39, every maximum class has a peeling scheme which successively removes vertices from the one-inclusion graph, so that the vertices being discarded never have degree more than d. But Rubinstein et al. (2007) gave examples of maximal classes of VC-dimension d which have a core of the one-inclusion graph of size $d+k$ for any constant k. Recall that a core is a subgraph and its size is the minimum degree of all the vertices. Having a peeling scheme gives an upper bound on the size of any core and so the result follows. ∎

9.1 Proof of Main Theorem

The proof is by induction on the complexity of $\mathcal{P}$. Since we are dealing with the class of d-contractible PL-hyperplane arrangements, it is easy to see that if any such $\mathcal{P}$ is *split open* along some fixed hyperplane P_1 in the arrangement (see Figures 14–15), then the result is two new arrangements $\mathcal{P}_+, \mathcal{P}_-$ each of which contains fewer hyperplanes and also fewer complementary regions than the initial one. The new arrangements have smaller complexity than $\mathcal{P}$ and are $k-, k'$-contractible for some $k, k' \leq d$. This is the key idea of the construction.

To examine this splitting process in detail, first note that each hyperplane P_α of $\mathcal{P}$ is either disjoint from P_1 or splits along P_1 into two hyperplanes P_α^+, P_α^-. We can now construct the new PL-hyperplane arrangements $\mathcal{P}_+, \mathcal{P}_-$ in the balls B_+, B_- obtained by splitting B along P_1. Note that $\partial B_+ = P_1 \cup D_+$ and $\partial B_- = P_1 \cup D_-$ where D_+, D_- are balls of dimension $n-1$ which have a common boundary with P_1. It is easy to verify that $\mathcal{P}_+, \mathcal{P}_-$ satisfy similar hypotheses to the original arrangement. Observe that the maximum number of mutually intersecting hyperplanes in $\mathcal{P}_+, \mathcal{P}_-$ may decrease relative to this number for $\mathcal{P}$, after the splitting operation. The reason is that the hyperplane P_1 'disappears' after splitting and so if all maximum subcollections of $\mathcal{P}$ which

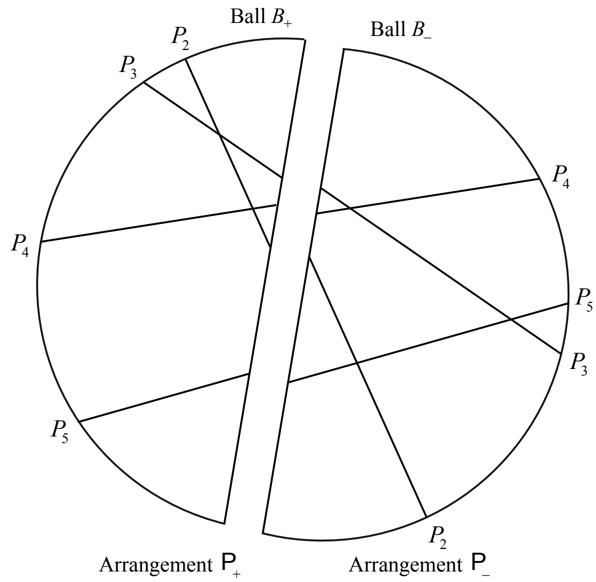


Figure 15: Result of splitting $\mathcal{P}$ in Figure 14(a) along hyperplane P_1.

mutually intersect, all contain P_1, then this number is smaller for $\mathcal{P}_+, \mathcal{P}_-$ as compared to the initial arrangement $\mathcal{P}$. This number shows that $\mathcal{P}_+, \mathcal{P}_-$ can be k- or k'-contractible, for $k, k' < d$ as well as the cases where $k, k' = d$.

Start the induction with any arrangement with one hyperplane. This gives two regions and complexity $(2,1)$. The corresponding graph has one edge and two vertices and obviously can be corner peeled.

We now describe the inductive step. There are two cases. In the first, assume the arrangement has complexity $(r,1)$. The corresponding graph has a vertex which belongs to only one edge, so can be corner peeled. This gives an arrangement with fewer hyperplanes and clearly the complexity has decreased to $(r-1,s)$ for some s. This completes the inductive step for the first case.

For the second case, assume that all d-contractible hyperplane arrangements with complexity smaller than (r,s) have corner-peeling sequences and $s > 1$. Choose any d-contractible hyperplane arrangement $\mathcal{P}$ with complexity (r,s). Select a hyperplane P_1 which splits the arrangement into two smaller arrangements $\mathcal{P}_+, \mathcal{P}_-$ in the balls B_+, B_-. By our definition of complexity, it is easy to see that however we choose P_1, the complexity of each of $\mathcal{P}_+, \mathcal{P}_-$ will be less than that of $\mathcal{P}$. However, a key requirement for the proof will be that we select P_1 so that it has precisely s complementary regions for $\mathcal{P}_+$, that is, P_1 has fewest complementary regions in one of its halfspaces, amongst all hyperplanes in the arrangement.

Since $\mathcal{P}_+$ has smaller complexity than (r,s), by our inductive hypothesis, it can be corner peeled (cf. Figure 16). If any of the corner-peeling moves of $\mathcal{P}_+$ is a corner-peeling move for $\mathcal{P}$, then the argument follows. For any corner-peeling move of $\mathcal{P}$ gives a PL-hyperplane arrangement with fewer complementary cells than $\mathcal{P}$ and thus smaller complexity than (r,s). Hence by the inductive hypothesis, it follows that $\mathcal{P}$ can be corner peeled.

Next, suppose that no corner-peeling move of $\mathcal{P}_+$ is a corner-peeling move for $\mathcal{P}$. In particular, the first corner-peeling move for $\mathcal{P}_+$ must occur for a cell R_+ in the complement of $\mathcal{P}$, which is

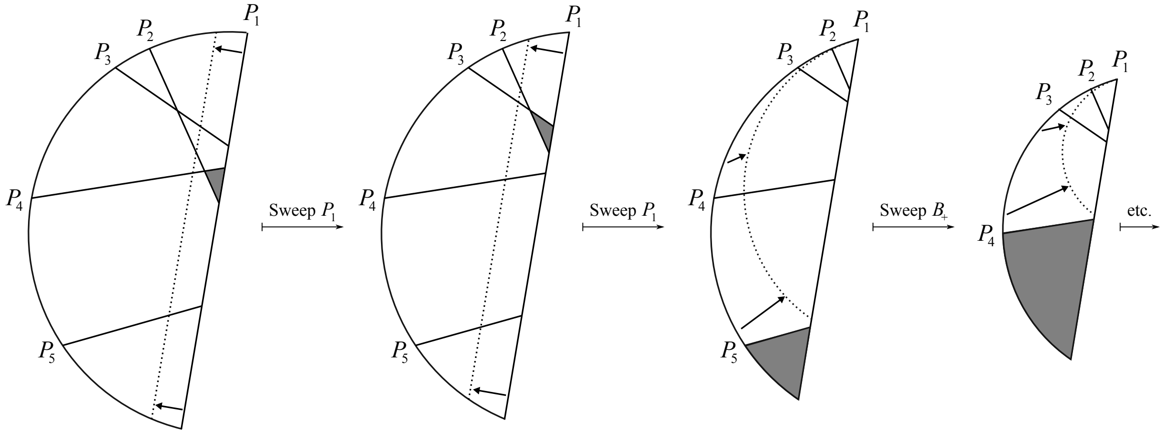


Figure 16: Partial corner-peeling sequence for the (B_+, P_+) arrangement split from the arrangement of Figure 15, in the proof of Theorem 39.

adjacent to P_1. (Clearly any corner-peeling move for $\mathcal{P}_+$, which occurs at a region R_1 with a face on D_+, will be a corner-peeling move for $\mathcal{P}$.) R_+ must be a product of a d'-simplex Δ with a copy of $\mathbb{R}^{n-d'}$, with one face on P_1 and the other faces on planes of $\mathcal{P}$. This is because a corner-peeling move can only occur at a cell with this type of face structure, as described in Theorem 29. The corresponding effect on the one-inclusion graph is peeling of a vertex which is a corner of a d'-cube in the binary class corresponding to the arrangement $\mathcal{P}_+$, where $d' \leq d$.

Now even though such a cell R_+ does not give a corner-peeling move for $\mathcal{P}$, we can push P_1 across R_+. The effect of this is to move the complementary cell R_+ from B^+ to B^-. Moreover, since we assumed that the hyperplane P_1 satisfies $\mathcal{P}^+$ has a minimum number s of complementary regions, it follows that the move pushing P_1 across R_+ produces a new arrangement $\mathcal{P}^\star$ with smaller complexity $(r, s-1)$ than the original arrangement $\mathcal{P}$. Hence by our inductive assumption, $\mathcal{P}^\star$ admits a corner-peeling sequence.

To complete the proof, we need to show that existence of a corner-peeling sequence for $\mathcal{P}^\star$ implies that the original arrangement $\mathcal{P}$ has at least one corner-peeling move. Recall that R_+ has face structure given by $\Delta \times \mathbb{R}^{n-d'}$, with one face on P_1 and the other faces on planes of $\mathcal{P}$. Consider the subcomplex U of the one-inclusion graph consisting of all the regions sharing a vertex or face of dimension k for $1 \leq k \leq n-1$ with R_+. It is not difficult to see that U is a d'-ball consisting of $d'+1$ cubes, each of dimension d'. (As examples, if $d' = 2$, U consists of 3 2-cubes forming a hexagon and if $d' = 3$, U consists of 4 3-cubes with boundary a rhombic dodecahedron.)

Consider the first corner-peeling move on the arrangement $\mathcal{P}^\star$. Note that the one-inclusion graphs of $\mathcal{P}^\star$ and $\mathcal{P}$ differ precisely by replacing U with U', that is, by a Pachner move. Hence this first corner-peeling move must occur at a vertex v_1 whose degree is affected by this replacement, since otherwise, the corner-peeling move would also apply to $\mathcal{P}$ and the proof would be complete. In fact, if v_1 has the same number of adjacent edges before and after the Pachner move, then it must belong to the same single maximum dimension cube before and after the Pachner move. (The only cubes altered by the Pachner move are the ones in U.) It is easy to see that, v_1 must belong to $\partial U = \partial U'$ and must have degree d' in $\mathcal{P}^\star$. So v_1 is a corner of a single d'-cube for U' and does

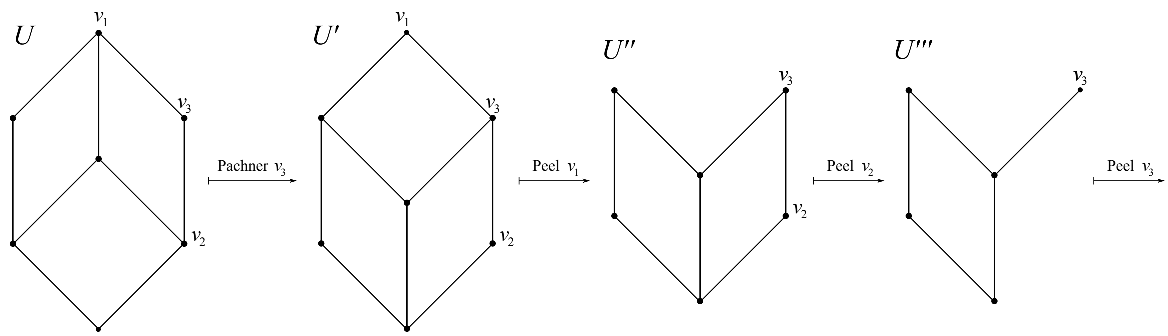


Figure 17: A 2-maximum complex in the 3-cube. After a Pachner move vertices v_1, v_2, v_3, etc. can be corner-peeled.

not belong to any other edges or cubes of the one inclusion graph for $\mathcal{P}^\star$. In U (and hence also in $\mathcal{P}$), v_1 belongs to d'-cubes of dimension d' and so has degree $d'+1$. After peeling away v_1 and its corresponding d'-cube, we still have a d'-ball with only d'-cubes, (cf. Figure 17).

Consider the next corner-peeling move. We claim that it must again be at a vertex v_2 belonging to $\partial U'$. The reason is that only vertices belonging to U' have degree reduced by our first corner-peeling move. So if this second move did not occur at a vertex of U', then it could be used as a corner-peeling move of our initial arrangement $\mathcal{P}$. There may be several choices for v_2. For example, if $d' = 2$, then U' is a hexagonal disk and removing one 2-cube from U' gives a choice which could be either of the two vertices which are corners of a single 2-cube in U', (cf. Figure 17). Note that a vertex which is a corner of a single cube in U' remains so after corner peeling at v_1. Note also that v_2 cannot belong to any edges of the one-inclusion graph which are not in U', as for v_1, if v_2 can be used for corner peeling.

We can continue examining corner-peeling moves of $\mathcal{P}^\star$ and find that all must occur at vertices in $\partial U'$, until the unique interior vertex is ready to be peeled, that is, belongs to a single cube. (See Figure 17.) The key to understanding this is that firstly, when we initially peel only vertices in $\partial U'$, these are not adjacent to any vertices of the one-inclusion graph outside U' and so cannot produce any new opportunities for corner peeling of vertices not in U'. Secondly, if the unique interior vertex v of U' can be corner peeled, after sufficiently many vertices in $\partial U'$ have been peeled, then new vertices in $\partial U'$ become candidates for peeling. For although these latter vertices may be adjacent to vertices outside U', after v has been peeled, they may become a corner vertex of a unique maximal cube.

But now a final careful examination of this situation shows that there must be at least one vertex of U which belongs to a single d'-cube in U and to no other edges in $\mathcal{P}$. So this will give our initial corner-peeling move of $\mathcal{P}$.

To elaborate, we can describe U as the set of d'-cubes which share the vertex $(0,0,\dots,0)$ in the $(d'+1)$-cube $\{0,1\}^{d'+1}$. Then U' consists of all the d'-cubes in $\{0,1\}^{d'+1}$ which contain the vertex $(1,1,\dots,1)$. Now assume that an initial sequence of corner peeling of vertices in $\partial U'$ allows the next step to be corner peeling of the unique interior vertex v. Note that in the notation above, v corresponds to the vertex $(1,1,\dots,1)$.

As in Figure 17, we may assume that after the corner peeling corresponding to the initial sequence of vertices in $\partial U'$, that there is a single d'-cube left in U' containing v. Without loss of generality, suppose this is the cube with vertices with $x_1 = 1$ where the coordinates are $x_1, x_2, \ldots, x_{d'+1}$ in $\{0,1\}^{d'+1}$. But then, it follows that there are no vertices outside U' adjacent to any of the initial sequence of vertices, which are all the vertices in $\{0,1\}^{d'+1}$ with $x_1 = 0$, except for $(0,0,\ldots,0)$. But now the vertex $(0,1,\ldots,1)$ has the property that we want - it is contained in a unique d'-cube in U and is adjacent to no other vertices outside U. This completes the proof.

9.2 Peeling Classes with Generic Linear or Generic Hyperbolic Arrangements

In this subsection, we study a special class of d-contractible arrangements. If a collection of hyperplanes in an n-manifold is in general position, then they have the property in the following definition. Then a key idea in differential or PL topology is that any collection can be slightly perturbed to be in general position. See Rourke and Sanderson (1982) for a discussion of these issues in the PL case.

Definition 41 *A linear or hyperbolic-hyperplane arrangement $\mathcal{P}$ in $\mathbb{R}^n$ or $\mathbb{H}^n$ respectively, is called* generic, *if any subcollection of k hyperplanes of $\mathcal{P}$, for $2 \le k \le n$ has the property that there are no intersection points or the subcollection intersects transversely in a plane of dimension $n-k$.*

Corollary 42 *Suppose $\mathcal{P}$ is a generic linear or hyperbolic-hyperplane arrangement in $\mathbb{R}^n$ or $\mathbb{H}^n$ and amongst all subcollections of $\mathcal{P}$, the largest with an intersection point in common, has d hyperplanes. Then $\mathcal{P}$ admits a d-corner-peeling scheme.*

Remark 43 *The proof of Corollary 42 is immediate since it is obvious that any generic linear or hyperbolic-hyperplane arrangement is a d-contractible PL-hyperplane arrangement, where d is the cardinality of the largest subcollection of hyperplanes which mutually intersect. Note that many generic linear, hyperbolic or d-contractible PL-hyperplane arrangements do not embed in any simple linear, hyperbolic or PL-hyperplane arrangement. For if there are two hyperplanes in $\mathcal{P}$ which are disjoint, then this is an obstruction to enlarging the arrangement by adding additional hyperplanes to obtain a simple arrangement. Hence this shows that Theorem 39 produces compression schemes, by corner peeling, for a considerably larger class of one-inclusion graphs than just maximum one-inclusion graphs. However it seems possible that d-contractible PL hyperplanes always embed in d-maximum classes, by 'undoing' the operation of sweeping and corner peeling, which pulls apart the hyperplanes.*

10. Peeling Infinite Maximum Classes with Finite-Dimensional Arrangements

We seek infinite classes represented by arrangements satisfying the same conditions as above. Note that any finite subclass of such an infinite class then satisfies these conditions and so can be corner peeled. Hence any such a finite subclass has a complementary region R which has face structure of the product of a d'-simplex with a copy of $\mathbb{R}^{n-d'}$ with one face on the boundary of B^n. To find such a region in the complement of our infinite collection $\mathcal{P}$, we must impose some conditions.

One convenient condition (cf. the proof of Theorem 39) is that a hyperplane P_α in $\mathcal{P}$ can be found which splits B^n into pieces B_+, B_- so that one, say B_+ gives a new arrangement for which the maximum number of mutually intersecting hyperplanes is strictly less than that for $\mathcal{P}$. Assume that the new arrangement satisfies a similar condition, and we can keep splitting until we get to disjoint hyperplanes.

It is not hard to prove that such arrangements always have peeling sequences. Moreover the peeling sequence does give a compression scheme. This sketch establishes the following.

Theorem 44 *Suppose that a countably infinite collection $\mathcal{P}$ of PL hyperplanes $\{P_\alpha\}$, each properly embedded in an n-ball B^n, satisfies the following conditions:*

i. $\mathcal{P}$ satisfies the conditions of d-contractible arrangements as in Definition 35 and

ii. There is an ordering of the planes in $\mathcal{P}$ so that if we split B^n successively along the planes, then at each stage, at least one of the two resulting balls has an arrangement with a smaller maximum number of planes which mutually intersect.

Then there is a d-corner-peeling scheme for $\mathcal{P}$, and this provides a d-unlabeled compression scheme.

Example 21 *Rubinstein and Rubinstein (2008) give an example that satisfies the assumptions of Theorem 44. Namely in $\mathbb{R}^n$ choose the positive octant $O = \{(x_1, x_2, \ldots x_n) : x_i \geq 0\}$. Inside O choose the collection of hyperplanes given by $x_i = m$ for all $1 \leq i \leq n$ and $m \geq 1$ a positive integer. There are many more examples, we present only a very simple model here. Take a graph inside the unit disk D with a single vertex of degree 3 and the three end vertices on ∂D. Now choose a collection of disjoint embedded arcs representing hyperplanes with ends on ∂D and meeting one of the edges of the graph in a single point. We choose finitely many such arcs along two of the graph edges and an infinite collection along one arc. This gives a very simple family of hyperplanes satisfying the hypotheses of Theorem 44. Higher dimensional examples with intersecting hyperplanes based on arbitrary trees can be constructed in a similar manner.*

11. Contractibility, Peeling and Arrangements

In this section, we characterize the concept classes which have one-inclusion graphs representable by d-contractible PL-hyperplane arrangements.

Theorem 45 *Assume that $\mathcal{C}$ is a concept class in the binary n-cube and d is the largest dimension of embedded cubes in its one-inclusion graph Γ. The following are equivalent.*

i. Γ is a strongly contractible cubical complex.

ii. There is a d-contractible PL-hyperplane arrangement $\mathcal{P}$ in an n-ball which represents Γ.

Proof To prove that *i* implies *ii*, we use some important ideas in the topology of manifolds. The cubical complex $\mathcal{C}$ is naturally embedded into the binary n-cube, which can be considered as an n-ball B^n. A regular neighborhood N of $\mathcal{C}$ homotopy retracts onto $\mathcal{C}$ and so is contractible. Now we can use a standard argument from algebraic and geometric topology to prove that N is a ball. Firstly, ∂N is simply connected, assuming that $n-d>2$. For given a loop in ∂N, it bounds a disk in N by contractibility. Since $\mathcal{C}$ is a d-dimensional complex and $n-d>2$ it follows that this disk can be pushed off $\mathcal{C}$ by transversality and then pushed into ∂N. But now we can follow a standard argument using the solution of the Poincaré conjecture in all dimensions (Perelman, 2002; Freedman, 1982; Smale, 1961). By duality, it follows that ∂N is a homotopy $(n-1)$-sphere and so by the Poincaré conjecture, ∂N is an $(n-1)$-sphere. Another application of the Poincaré conjecture shows that N is an n-ball.

Next, the bisecting planes of the binary n-cube meet the n-ball N in neighborhoods of reductions. Hence the assumption that each reduction is contractible enables us to conclude that these intersections are also PL hyperplanes in N. Therefore the PL-hyperplane arrangement has been constructed which represents Γ. It is easy to see that this arrangement is indeed d-contractible, since strong contractibility implies that all multiple reductions are contractible and so intersections of subfamilies of PL hyperplanes are either empty or are contractible and hence planes, by the same argument as the previous paragraph. (Note that such intersections correspond to multiple reductions of Γ.)

Finally to show that *ii* implies *i*, by Theorem 39, a d-contractible PL-hyperplane arrangement $\mathcal{P}$ has a peeling sequence and so the corresponding one-inclusion graph Γ is contractible. This follows since a corner-peeling move can be viewed as a homotopy retraction. But then reductions and multiple reductions are also represented by d'-contractible hyperplane arrangements, since these correspond to the restriction of $\mathcal{P}$ to the intersection of a finite subfamily of hyperplanes of $\mathcal{P}$. It is straightforward to check that these new arrangements are d'-contractible, completing the proof. ■

Remark 46 *Note that any one-inclusion graph Γ which satisfies the hypotheses of Theorem 45 admits a corner-peeling sequence. From the proof above, Γ must be contractible if it has a peeling sequence. However Γ does not have to be strongly contractible. A simple example can be found in the binary 3-cube, with coordinate directions x,y,z. Define Γ to be the union of four edges, labeled x,y,z,x. It is easy to see that Γ has a peeling sequence and is contractible but not strongly contractible. For the bisecting hyperplane transverse to the x direction meets Γ in two points, so the reduction Γ^x is a pair of vertices, which is not contractible.*

Note that all maximum classes are strongly contractible, as are also all linear and hyperbolic arrangements, by Corollary 42 and Theorem 45.

12. Future Directions: Compression Schemes for Maximal Classes

In this section, we compare two maximal classes of VC-dimension 2 in the binary 4-cube. For the first, we show that the one-inclusion graph is not contractible and therefore there is no peeling or corner-peeling scheme. There is an unlabeled compression scheme, but this is not associated with either peeling or a hyperplane arrangement. For the second, the one-inclusion graph is contractible but not strongly contractible. However there are simple corner-peeling schemes and a related compression scheme. Note that the relation between the compression scheme and the corner-peeling scheme is not as straightforward as in our main result above. Finally for the second example, there is a non simple hyperplane arrangement consisting of lines in the hyperbolic plane which represents the class. It would be interesting to know if there are many maximal classes which admit such non simple representations and if there is a general procedure to find associated compression schemes.

Example 22 *Let $\mathcal{C}$ be the maximal class of VC-dimension 2 in the 4-cube with concepts and labels shown in Figure 18(a). This forms an unlabeled compression scheme. Note that the one-inclusion graph is not connected, consisting of four 2-cubes with common vertex at the origin 0000 and an isolated vertex at 1111. So since a contractible complex is connected, the one-inclusion graph*

Concept	Label
0000	$\emptyset$
1000	x_1
0100	x_2
0010	x_3
0001	x_4
1100	x_1x_2
0011	x_3x_4
0110	x_2x_3
1001	x_1x_4
1111	x_1x_3, x_2x_4

(a)

Concept	Label
0000	$\emptyset$
1000	x_1
0100	x_2
0010	x_3
1100	x_1x_2
0110	x_2x_3
1010	x_1x_3
1011	x_2x_4
1101	x_3x_4
0111	x_1x_4

(b)

Figure 18: VC-2 maximal classes from (a) Example 22 and (b) Example 23.

cannot be contractible. Moreover any hyperplane arrangement represents a connected complex so there cannot be such an arrangement for this example. This example is the same class (up to flipping coordinate labels) as in Kuzmin and Warmuth (2007, Table 2) but there appear to be some errors there in describing the compression scheme.

Example 23 *Let C be the maximal class of VC-dimension 2 in the 4-cube with concepts and labels defined in Figure 18(b). The class is enlarged by adding an extra vertex 1111 x_4 to complete the labeling.*

This forms an unlabeled compression scheme and is the same as in Kuzmin and Warmuth (2007, Table 1). The one-inclusion graph is contractible, consisting of three 2-cubes with common vertex 0100 and three edges attached to these 2-cubes. It is easy to form a hyperbolic-line arrangement consisting of three lines meeting in three points forming a triangle and three further lines near the boundary of the hyperbolic plane which do not meet any other line.

It is easy to see that there is a corner-peeling sequence, but there is not such an obvious way of using this to form a compression scheme. The idea is that the label x_1x_4 comes from picking the origin at 0000 and considering the shortest path to the origin as giving the label. There are numerous ways of corner peeling this one-inclusion complex. The only other comment is that the final vertices 0111, 1011, 1101 and 1111 are labeled in a different manner. Namely putting the origin at 0000 means that 0111 has shortest path with label $x_2x_3x_4$. We replace this by the label x_1x_4 since clearly this satisfies the no-clashing condition. Then the final vertex 1111 has the remaining label x_4 to uniquely specify it.

13. Conclusions and Open Problems

We saw in Corollary 23 that d-maximum classes represented by simple linear-hyperplane arrangements in $\mathbb{R}^d$ have underlying cubical complexes that are homeomorphic to a d-ball. Hence the VC dimension and the dimension of the cubical complex are the same. Moreover in Theorem 33, we proved that d-maximum classes represented by PL-hyperplane arrangements in $\mathbb{R}^d$ are those whose underlying cubical complexes are manifolds or equivalently d-balls.

Question 47 *Does every simple PL-hyperplane arrangement in B^d, where every subcollection of d planes transversely meet in a point, represent the same concept class as some simple linear-hyperplane arrangement?*

Question 48 *What is the connection between the VC dimension of a maximum class induced by a simple hyperbolic-hyperplane arrangement and the smallest dimension of hyperbolic space containing such an arrangement? In particular, can the hyperbolic space dimension be chosen to only depend on the VC dimension and not the dimension of the binary cube containing the class?*

We gave an example of a 2-maximum class in the 5-cube that cannot be realized as a hyperbolic-hyperplane arrangement in $\mathcal{H}^3$. Note that the Whitney embedding theorem (Rourke and Sanderson, 1982) proves that any cubical complex of dimension d embeds in $\mathbb{R}^{2d}$. Can such an embedding be used to construct a hyperbolic arrangement in $\mathcal{H}^{2d}$ or a PL arrangement in $\mathbb{R}^{2d}$?

The structure of the boundary of a maximum class is strongly related to corner peeling. For Euclidean-hyperplane arrangements, the boundary of the corresponding maximum class is homeomorphic to a sphere by Corollaries 22 and 23.

Question 49 *Is there a characterization of the cubical complexes that can occur as the boundary of a maximum class? Characterize maximum classes with isomorphic boundaries.*

Question 50 *Does a corner-peeling scheme exist with corner vertex sequence having minimum degree?*

Theorem 32 suggests the following.

Question 51 *Can any d-maximum class in $\{0,1\}^n$ be represented by a simple arrangement of hyperplanes in $\mathbb{H}^n$?*

Question 52 *Which compression schemes arise from sweeping across simple hyperbolic-hyperplane arrangements?*

Kuzmin and Warmuth (2007) note that there are unlabeled compression schemes that are cyclic. In Proposition 17 we show that corner-peeling compression schemes (like min-peeling) are acyclic. So compression schemes arising from sweeping across simple arrangements of hyperplanes in Euclidean or hyperbolic space are also acyclic. Does acyclicity characterize such compression schemes?

We have established peeling of all finite maximum and a family of infinite maximum classes by representing them as PL-hyperplane arrangements and sweeping by multiple generic hyperplanes. A larger class of arrangements has these properties—namely those which are d-contractible—and we have shown that the corresponding one-inclusion graphs are precisely the strongly contractible ones. Finally we have established that there are d-maximal classes that cannot be embedded in any $(d+k)$-maximum classes for any constant k. Some important open problems along these lines are the following.

Question 53 *Prove peeling of maximum classes using purely combinatorial arguments*

Question 54 *Can all maximal classes be peeled by representing them by hyperplane arrangements and then using a sweeping technique (potentially solving the Sample Compressibility conjecture)? The obvious candidate for this approach is to use d-contractible PL-hyperplane arrangements.*

Question 55 *What about more general collections of infinite maximum classes, or infinite arrangements?*

Question 56 *Is it true that any d-contractible PL-hyperplane arrangement is equivalent to a hyperbolic-hyperplane arrangement?*

Question 57 *Is it true that all strongly contractible classes, with largest dimension d of cubes can be embedded in maximum classes of VC-dimension d?*

Acknowledgments

We thank Peter Bartlett for his very helpful feedback, and gratefully acknowledge the support of the NSF through grants DMS-0434383 and DMS-0707060, and the support of the Siebel Scholars Foundation.

References

N. Alon. On the density of sets of vectors. *Discrete Mathematics*, 46(2):199–202, 1983.

S. Ben-David and A. Litman. Combinatorial variability of Vapnik-Chervonenkis classes with applications to sample compression schemes. *Discrete Applied Mathematics*, 86(1):3–25, 1998.

A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *Journal of the ACM*, 36(4):929–965, 1989.

R. M. Dudley. The structure of some Vapnik-Chervonenkis classes. In L.M. Le Cam and R.A. Olshen, editors, *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman*, volume II, pages 495–507. Wadsworth, 1985.

H. Edelsbrunner. *Algorithms in Combinatorial Geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, 1987.

S. Floyd. Space-bounded learning and the Vapnik-Chervonenkis dimension. Technical Report TR-89-061, ICSI, UC Berkeley, 1989.

P. Frankl. On the trace of finite sets. *Journal of Combinatorial Theory (A)*, 34(1):41–45, 1983.

M. Freedman. The topology of four-dimensional manifolds. *Journal of Differential Geometry*, 17: 357–454, 1982.

B. Gärtner and E. Welzl. Vapnik-Chervonenkis dimension and (pseudo-) hyperplane arrangements. *Discrete and Computational Geometry*, 12:399–432, 1994.

D. Haussler. Sphere packing numbers for subsets of the boolean *n*-cube with bounded Vapnik-Chervonenkis dimension. *Journal of Combinatorial Theory (A)*, 69:217–232, 1995.

D. Haussler, N. Littlestone, and M. K. Warmuth. Predicting $\{0,1\}$ functions on randomly drawn points. *Information and Computation*, 115(2):284–293, 1994.

L. Hellerstein, 28 June 2006. pers. comm.

D. Helmbold, R. Sloan, and M. K. Warmuth. Learning integer lattices. *SIAM Journal on Computing*, 21(2):240–266, 1992.

D. Kuzmin and M. K. Warmuth. Unlabeled compression schemes for maximum classes. *Journal of Machine Learning Research*, 8(Sep):2047–2081, 2007.

N. Littlestone and M. K. Warmuth. Relating data compression and learnability. Unpublished manuscript `http://www.cse.ucsc.edu/~manfred/pubs/lrnk-olivier.pdf`, 1986.

M. Marchand and J. Shawe-Taylor. The decision list machine. In *Advances in Neural Information Processing Systems 15*, pages 921–928, 2003.

B. Mazur. A note on some contractible 4-manifolds. *Annals of Mathematics*, 73:221–228, 1961.

T. Neylon. *Sparse Solutions for Linear Prediction Problems*. PhD thesis, NYU, 2006.

U. Pachner. Konstruktionsmethoden und das kombinatorische Homöomorphieproblem für Triangulationen kompakter semilinearer Mannigfaltigkeiten. (german) [Construction methods and the combinatorial homeomorphism problem for triangulations of compact semilinear manifolds]. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 57:69–86, 1987.

G. Perelman. Finite extinction time for the solutions to the Ricci flow on certain 3-manifolds, 2002. `http://arXiv.org/math.DG/0211159v1`.

J. G. Ratcliffe. *Foundations of Hyperbolic Manifolds*. Springer-Verlag, 1994.

C. Rourke and B. Sanderson. *Introduction to Piecewise-Linear Topology*. Springer-Verlag, 1982.

B. I. P. Rubinstein and J. H. Rubinstein. Geometric & topological representations of maximum classes with applications to sample compression. In *Proceedings of the 21st Annual Conference on Learning Theory (COLT'08)*, pages 299–310, 2008.

B. I. P. Rubinstein, P. L. Bartlett, and J. H. Rubinstein. Shifting, one-inclusion mistake bounds and tight multiclass expected risk bounds. In *Advances in Neural Information Processing Systems 19*, pages 1193–1200, 2007.

B. I. P. Rubinstein, P. L. Bartlett, and J. H. Rubinstein. Shifting: one-inclusion mistake bounds and sample compression. *Journal of Computer and System Sciences: Special Issue on Learning Theory 2006*, 75(1):37–59, 2009.

N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory (A)*, 13:145–147, 1972.

S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41(1):247–261, 1972.

S. Smale. Generalized Poincaré conjecture in dimensions greater than four. *Annals of Mathematics*, 74:391–406, 1961.

V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

U. von Luxburg, O. Bousquet, and B. Schölkopf. A compression approach to support vector model selection. *Journal of Machine Learning Research*, 5:293–323, 2004.

M. K. Warmuth. Compressing to VC dimension many points. In *Proceedings of the 16th Annual Conference on Learning Theory*, 2003.

E. Welzl. Complete range spaces. Unpublished notes, 1987.