
Geometric & Topological Representations of Maximum Classes with Applications to Sample Compression

Benjamin I. P. Rubinstein¹ and J. Hyam Rubinstein²

¹ Computer Science Division, University of California, Berkeley, U.S.A.

² Department of Mathematics & Statistics, the University of Melbourne, Australia

¹ benr@cs.berkeley.edu, ² rubin@ms.unimelb.edu.au

Abstract

We systematically investigate finite maximum classes, which play an important role in machine learning as concept classes meeting Sauer’s Lemma with equality. 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 conjecture of Kuzmin & Warmuth. A bijection between 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.

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 [Flo89, War03, KW07]. It is this structure that admits many elegant geometric and algebraic topological representations upon which this paper focuses.

Littlestone & Warmuth [LW86] 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 [LW86]. 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. Recently Kuzmin and Warmuth achieved compression of maximum classes without the use of labels [KW07]. They also conjectured that their elegant Min-

Peeling Algorithm constitutes such an unlabeled d -compression scheme for d -maximum classes.

As in our previous work [RBR08], 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 & Warmuth 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 [KW07, Conjecture 1]. 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 presented in [BDL98], Kuzmin & Warmuth ask whether unlabeled compression schemes for infinite classes such as positive half spaces can be constructed explicitly [KW07]. 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 and Hyperbolic 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 on viewing simple PL hyperplane arrangements as maximum classes [GW94]., The close relationship between such arrangements and their Hyperbolic versions suggests that they could be equivalent. Although PL sweeping does not immediately admit corner-peeling or compression, the PL representation result is used to prove the peeling conjecture [KW07, Conjecture 1] for VC dimension two.

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.

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 [KW07, Open Problem 3]. We present non-embeddability results following from our earlier counter-examples to Kuzmin & Warmuth's minimum degree conjecture [RBR08] and our new results on corner-peeling.

2 Background

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 $f : X \rightarrow Y$.

Definition 2 A homotopy is a continuous map $F : X \times [0, 1] \rightarrow Y$. 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 $f : X \rightarrow Y$ and $g : Y \rightarrow X$ such that $f \circ g$ and $g \circ f$ are homotopic to the identity maps on X and Y respectively. We say that X and Y have the same homotopy type if there is a homotopy equivalence between them.

Definition 3 A cubical complex is a union of solid cubes of the form $[a_1, b_1] \times \dots \times [a_m, b_m]$, 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.

2.2 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 [HLW94]; $\mathcal{G}(C)$ 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 [RBR08]. Each edge in $\mathcal{G}(C)$ is labeled by the component on which the two vertices differ.

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 [BEHW89]. One key to all such results is Sauer's Lemma.

Definition 5 The VC-dimension of $C \subseteq \{0, 1\}^X$ is defined as $\text{VC}(C) = \sup \left\{ n \mid \exists Y \in \binom{X}{n}, \Pi_Y(C) = \{0, 1\}^n \right\}$ 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 6 ([VC71, Sau72, She72]) $|C| \leq \sum_{i=1}^{\text{VC}(C)} \binom{n}{i}$ for all $C \subseteq \{0, 1\}^n$.

Motivated by maximizing concept class cardinality under a fixed VC-dimension, which is related to constructing general sample compression schemes (see Section 2.3), Welzl defined the following special classes in [Wel87].

Definition 7 Concept class $C \subseteq \{0, 1\}^X$ is called maximal if $\text{VC}(C \cup \{c\}) > \text{VC}(C)$ for all $c \in \{0, 1\}^X \setminus C$. 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.

The reduction of $C \subseteq \{0, 1\}^n$ with respect to $i \in [n] = \{1, \dots, n\}$ is class $C^i = \Pi_{[n] \setminus \{i\}}(\{c \in C \mid i \in I_{\mathcal{G}(C)}(c)\})$ where $I_{\mathcal{G}(C)}(c) \subseteq [n]$ denotes the labels of the edges incident to vertex c ; the tail is $\text{tail}_i(C) = \{c \in C \mid i \notin I_{\mathcal{G}(C)}(c)\}$. Welzl showed that if C is d -maximum, then $\Pi_{[n] \setminus \{i\}}(C)$ and C^i are maximum of VC-dimensions d and $d-1$ 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 showed in [Flo89] that maximum $C \subseteq \{0, 1\}^n$ of VC-dim d is the union of a maximally overlapping *d-complete collection of cubes* [RBR08]—defined as a collection of d -cubes which project onto all $\binom{n}{d}$ possible sets of d coordinate directions. (See also [Ney06] for a different proof of this.) It has long been known that VC-1 maximum classes have one-inclusion graphs that are trees [Dud85]; in [RBR08] we extended this result by showing that when viewed as complexes, d -maximum classes are contractible d -cubical complexes. 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 [Ede87], but not all finite maximum classes correspond to such Euclidean arrangements [Flo89].

2.3 Sample Compression Schemes

Littlestone and Warmuth showed that the existence of a compression scheme of finite size is sufficient for learnability of C , and conjectured the converse, that $\text{VC}(C) = d < \infty$ implies a compression scheme of size d [LW86]. Later

Warmuth weakened the conjectured size to $O(d)$ [War03]. To-date it is only known that maximum classes can be d -compressed [Flo89]. Unlabeled compression was first explored in [BDL98]; Kuzmin and Warmuth define such compression as follows, explicitly constructing schemes of size d for maximum classes [KW07].

Definition 8 Let C be a d -maximum class on a finite domain X . A representation mapping r of C satisfies:

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

As with all published labeled schemes, known unlabeled compression schemes for maximum classes exploit their special recursive projection-reduction structure and so it is doubtful whether such schemes will generalize. Kuzmin and Warmuth 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 [KW07, Conjecture 2]. The authors also conjecture 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 [KW07, Conjecture 1]. Possibly the most 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 [KW07, Open Problem 3].

3 Preliminaries

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 $[n] \setminus \{i\}$ and the corresponding $(d-1)$ -maximum reduction.

Definition 9 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, i.e. 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.

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'

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

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 10 $\text{MAXIMUMCLASSES}(n, d)$ (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 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 $\text{MAXIMUMCLASSES}(n, d)$ returns only d -maximum classes in the n -cube, and that all such classes are produced by the algorithm.

Let classes $C \in \text{MAXIMUMCLASSES}(n-1, d)$ and $C' \in \text{MAXIMUMCLASSES}(n-1, d-1)$ 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^* formed from C and C' by Algorithm 1. The algorithm partitions C into C' -connected components C_1, \dots, 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 $\{C_i \times \{1\}\}_{i \notin S_0}$. By definition C^* is a d -complete collection of cubes with cardinality equal to $\binom{n}{d}$ since $|C^*| = |C'| + |C|$, as in [KW07]. So by [RBR08, Theorem 34] C^* is d -maximum.

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

d -cubes in C^* which contain color i . It is thus clear that C^* 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

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

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 $u, v \in C_j \setminus C'$, $u, v \in C'$ or WLOG $u \in C_j \setminus C'$, $v \in C'$. 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. [KW07, Corollary 7.5]). ■

3.2 Corner-Peeling

Kuzmin and Warmuth conjectured in [KW07, Conjecture 2] 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 $C_{t+1} = C_t \setminus \{v_t\}$. 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 11 Let $C \subseteq \{0, 1\}^n$ be a class with $d = \text{VC}(C)$. We say that C can be corner-peeled if there exists an ordering $v_1, \dots, v_{|C|}$ of the vertices of C such that, for each $t \in [|C|]$ where $C_0 = C$,

1. $v_t \in C_{t-1}$ and $C_t = C_{t-1} \setminus \{v_t\}$;
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.

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 12 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 13 If $C \subseteq \{0, 1\}^n$ is shortest-path closed and $v \in C$ is a corner vertex of C , then $C \setminus \{v\}$ 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 $\{u, v\} \subseteq C \setminus \{c\}$. 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 $C \setminus \{c\}$. 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 $C \setminus \{c\}$ is shortest-path closed. ■

3.2.1 Corner-Peeling implies Compression

Theorem 14 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 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* [Flo89, BDL98, KW07].

Proposition 15 All corner-peeling unlabeled compression schemes are acyclic.

Proof: We follow the proof that the Min-Peeling Algorithm is acyclic [KW07]. Let $v_1, \dots, v_{|C|}$ 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.

Definition 16 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 17 *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 $C^* \subseteq \{0, 1\}^n$ be a 2-maximum class and consider color $i \in [n]$. Then the reduction C^{*i} is an unrooted tree with at least two leaves, each of which lifts to an i -colored edge in C^* . Since the leaves are of degree 1 in C^{*i} , the corresponding lifted edges belong to exactly one 2-cube in C^* and so lie in ∂C^* . Consider now a d -maximum class $C^* \subseteq \{0, 1\}^n$ for $d > 2$, and make the inductive assumption that the projection $C = \Pi_{[n-1]}(C^*)$ has two of each type of $(d - 1)$ -cube, and that the reduction $C' = C^{*n}$ 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 $I \setminus \{x_n\}$ in $\partial C'$. By the same argument as in the base case, these lift to two I -colored cubes in ∂C^* . 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 ∂C^* , otherwise its unique lift is contained in ∂C^* . Therefore ∂C^* contains at least two I -colored cubes. ■

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

Example 18 *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. It is not hard to verify that D is contractible, but has no (geometric) boundary.*

Although Theorem 17 will not be explicitly used in the sequel, we return to boundaries of maximum complexes later.

4 Euclidean Arrangements

Definition 19 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 one 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 $n - k$. 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 $[n] \setminus \{i\}$ corresponds to removing the i^{th} plane;
- The reduction C^i is the new arrangement given by the intersection of the 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 $(n - k + 1)$ -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 [Ede87].

Lemma 20 *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 i.e. 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 by [RBR08]. ■

Corollary 21 *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 20 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. 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. ■

Corollary 22 *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} .*

	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

Figure 1: A 2-maximum class in $\{0, 1\}^4$ having a simple linear line arrangement in \mathbb{R}^2 .

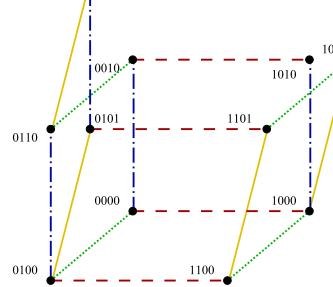


Figure 2: The 2-maximum class in the 4-cube, enumerated in Figure 1.

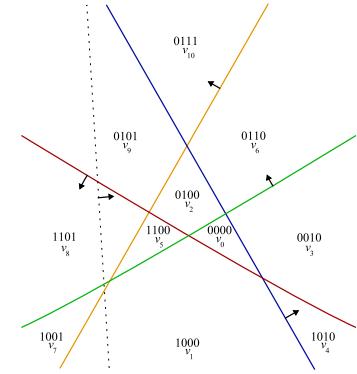


Figure 3: A simple linear line arrangement corresponding to the class in Figure 1, swept by the dashed line. Each cell has a unique vertex.

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. ■

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 23 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 in Kuzmin & Warmuth [KW07], 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 $|\mathcal{H}_+ \cap c| = 1$, i.e. 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 24 Figure 1 enumerates the 11 vertices of a 2-maximum class in the 4-cube. Figures 3 and 2 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 [KW07, Conjecture 1].

Theorem 25 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 subclass $C' \subseteq \partial C$ by Corollary 21. 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 -simplex¹ 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} . 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 14 completes the proof that this corner-peeling of C constitutes unlabeled compression. ■

Corollary 26 *The sequence of cubes $C'_0, \dots, C'_{|C|}$, 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.

Example 27 *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 corner-peeling must begin with the 2-cube.*

The next result follows from our counter-examples to Kuzmin & Warmuth's minimum degree conjecture [RBR08].

Corollary 28 *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

We briefly discuss the Klein model of hyperbolic geometry [Rat94, 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

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

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 29 *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 30 *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 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 plane.

Example 31 *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, i.e. a corner-vertex.*

Example 32 *Assume that we have a family of planes in the unit 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*

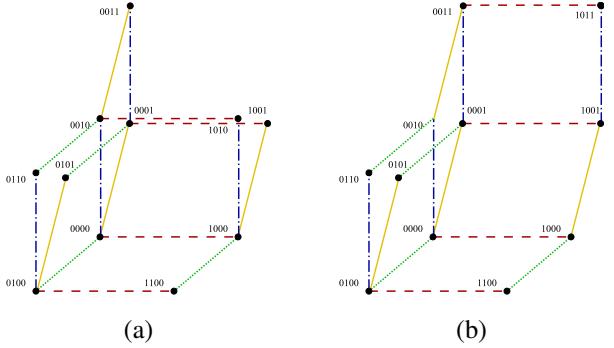


Figure 4: 2-maximum classes in $\{0, 1\}^4$ that can be represented as hyperbolic arrangements but not as Euclidean arrangements.

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 25. 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 26). Similar to Euclidean sweepouts, hyperbolic sweepouts correspond to corner-peeling and not min-peeling.

Theorem 33 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 25. 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 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 k -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 simplex as before

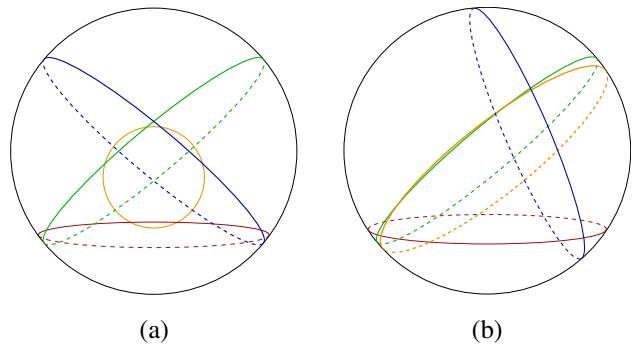


Figure 5: Hyperbolic hyperplane arrangements corresponding to the classes in Figure 4. In both cases the four hyperbolic planes meet in 6 straight line segments (not shown). The planes’ colors correspond to the edges’ colors in Figure 4.

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 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 34 Constructed with lifting, Figure 4 completes the enumeration, up to symmetry, of the 2-maximum classes in $\{0, 1\}^4$ begun with Example 24. These cases cannot be represented as simple Euclidean linear arrangements, since their boundaries do not satisfy the condition of Corollary 22 but can be represented as hyperbolic arrangements as in Figure 5. Figures 6 and 7 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.

Corollary 35 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 [RBR08].

Corollary 30 gives a proper superset of simple linear hyperplane arrangement-induced maximum classes as hyperbolic arrangements. We will prove in the next section 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.

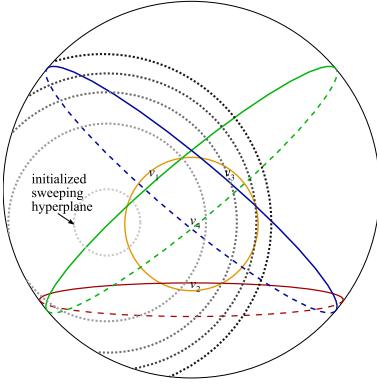


Figure 6: The simple hyperbolic arrangement of Figure 5.(a) with a generic sweeping hyperplane shown in several positions before and after it sweeps past four cells.

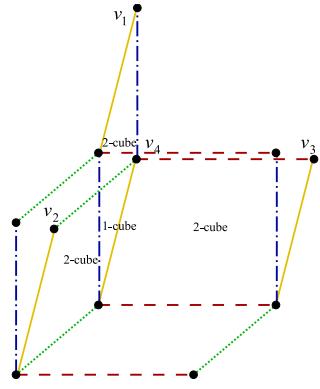


Figure 7: The 2-maximum class in $\{0, 1\}^4$ of Figure 4.(a), with the first four corner-vertices peeled by the hyperbolic arrangement sweeping of Figure 6. Notice that three 2-cubes are peeled, then a 1-cube (all shown) followed by 2-cubes.

Example 36 Let \mathcal{L} be the set of lines in the plane of the form $L_{2m} = \{(x, y) \mid x = 2m\}$ and $L_{2n+1} = \{(x, y) \mid y = 2n\}$ 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 sweep-out by the lines $\{(x, y) \mid x + (1 + \epsilon)y = t\}$ for $t \geq 0$ and any small fixed irrational $\epsilon > 0$. C is a 2-maximum class and the unlabeled compression scheme is also of size 2.

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} in order of increasing $m + n$. The lines $x + (1 + \epsilon)y = t$ 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 sweep-

out process, so that they have either singleton representatives or the empty set. In this way, as in [KW07], 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 37 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 and α, β are both close to 1. This example has similar properties to Example 36: the compression scheme is again given by decreasing t whereas corner-peeling corresponds to increasing t . Note that cells shrink to points, as $\mathbf{x} \rightarrow \mathbf{1}$ and the volume of cells converge to zero as $n \rightarrow \infty$, or equivalently any $x_i \rightarrow 1$.

Example 38 In the hyperbolic plane \mathbb{H}^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

A *PL hyperplane* is the image of a proper piecewise-linear homeomorphism from the $(k-1)$ -ball B^{k-1} into B^k , i.e. the inverse image of S^k is S^{k-1} [RS82]. 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.

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 39 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

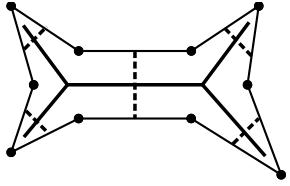


Figure 8: A 1-maximum class (thick solid lines) with its fattening (thin solid lines with points), bisecting sets (dashed lines) and induced complementary cells.

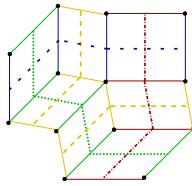


Figure 9: The top of Figure 4.(b) (i.e. the 2-cubes seen from above) gives part of the boundary of a regular neighborhood in \mathbb{R}^3 .

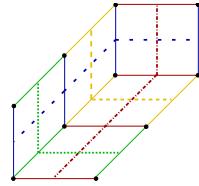


Figure 10: The bottom of Figure 4.(b) (i.e. the 2-cubes seen from below) gives the rest of the boundary of a regular neighborhood.

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 collapses onto C and so $\partial\mathcal{N}$ is a homology $(n - 1)$ -sphere (by a standard, well-known argument from topology [Maz61]). 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 generalised Poincaré conjecture [Sma61], 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 8.) 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 in a ball.

7.1.3 Shifting

To complete the induction step, we use the technique of shifting, [Alo83], [Fra83], [Hau95]. 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^* . We claim that the regular neighborhood of C is a ball if and only if the same is true for C^* . 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.

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, i.e. 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. 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 39, let $\partial\mathcal{N} = S^{n-1}$ denote the boundary of the n -ball \mathcal{N} constructed above (cf. Figures 9 and 10). 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 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 balls obtained in this way of dimension varying between $n - d$ and $n - 1$. Each of these faces has boundary a 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 , i.e. two copies of C glued together along their boundaries. The proof of Theorem 39 is now complete.

Example 40 Consider the bounded below 2-maximum class $\tilde{C} \subseteq \{0, 1\}^5$. We claim that \tilde{C} cannot be realized as an ar-

angement 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, i.e. 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 22: 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 41 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 22. 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' 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 and so collapsing shrinks the larger region to one of the same homeomorphism type). 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 22. 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

Theorem 42 Every 2-maximum class can be corner-peeled.

Proof: By Theorem 39, we can represent any 2-maximum class $C \subseteq \{0, 1\}^n$ by a simple family of 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 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, i.e. 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 Conclusions and Open Problems

We saw in Corollary 22 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 41, 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 43 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 44 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 [RS82] 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 21 and 22.

Question 45 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 46 Does a corner-peeling scheme exist with corner vertex sequence having minimum degree?

Theorem 39 suggests the following.

Question 47 Can any d -maximum class in $\{0, 1\}^n$ be represented by a simple arrangement of hyperplanes in \mathbb{H}^n ?

Question 48 Which compression schemes arise from sweeping across simple hyperbolic hyperplane arrangements?

Kuzmin & Warmuth note that there are unlabeled compression schemes that are cyclic [KW07]. In Proposition 15 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?

Acknowledgments: We thank Peter Bartlett and the first anonymous referee for their very helpful feedback.

References

- [Alo83] N. Alon. On the density of sets of vectors. *Discrete Math.*, 46(2):199–202, 1983.
- [BDL98] S. Ben-David and A. Litman. Combinatorial variability of Vapnik-Chervonenkis classes with applications to sample compression schemes. *Discrete Applied Math.*, 86(1):3–25, 1998.
- [BEHW89] 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.
- [Dud85] 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.
- [Ede87] H. Edelsbrunner. *Algorithms in Combinatorial Geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, 1987.
- [Flo89] S. Floyd. Space-bounded learning and the Vapnik-Chervonenkis dimension. Technical Report TR-89-061, ICSI, UC Berkeley, 1989.
- [Fra83] P. Frankl. On the trace of finite sets. *Journal of Comb. Theory (A)*, 34(1):41–45, 1983.
- [GW94] B. Gärtner and E. Welzl. Vapnik-Chervonenkis dimension and (pseudo-) hyperplane arrangements. *Discrete and Comp. Geometry*, 12:399–432, 1994.
- [Hau95] D. Haussler. Sphere packing numbers for subsets of the boolean n -cube with bounded Vapnik-Chervonenkis dimension. *Journal of Comb. Theory (A)*, 69:217–232, 1995.
- [HLW94] D. Haussler, N. Littlestone, and M.K. Warmuth. Predicting $\{0, 1\}$ functions on randomly drawn points. *Information and Computation*, 115(2):284–293, 1994.
- [KW07] D. Kuzmin and M.K. Warmuth. Unlabeled compression schemes for maximum classes. *Journal of Machine Learning Research*, 8(Sep):2047–2081, 2007.
- [LW86] N. Littlestone and M.K. Warmuth. Relating data compression and learnability. Unpublished manuscript, 1986.
- [Maz61] B. Mazur. A note on some contractible 4-manifolds. *Annals of Math.*, 73:221–228, 1961.
- [Ney06] T. Neylon. *Sparse Solutions for Linear Prediction Problems*. PhD thesis, NYU, 2006.
- [Rat94] J.G. Ratcliffe. *Foundations of Hyperbolic Manifolds*. Springer-Verlag, 1994.
- [RBR08] 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*, 2008. in press.
- [RS82] C. Rourke and B. Sanderson. *Introduction to Piecewise-Linear Topology*. Springer-Verlag, 1982.
- [Sau72] N. Sauer. On the density of families of sets. *Journal of Comb. Th. (A)*, 13:145–147, 1972.
- [She72] S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Math.*, 41(1):247–261, 1972.
- [Sma61] S. Smale. Generalized Poincaré conjecture in dimensions greater than four. *Annals of Math.*, 74:391–406, 1961.
- [VC71] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Prob. and its Applic.*, 16(2):264–280, 1971.
- [War03] M.K. Warmuth. Compressing to VC dimension many points. In *Proceedings of the 16th Annual Conference on Learning Theory*, 2003.
- [Wel87] E. Welzl. Complete range spaces. Unpublished notes, 1987.