Unlabeled Sample Compression Schemes and Corner Peelings for Ample and Maximum Classes

Jérémie Chalopin (ID)
CNRS, Aix-Marseille Université, Université de Toulon, LIS, Marseille, France
jeremie.chalopin@lis-lab.fr

Victor Chepoi (ID)
Aix-Marseille Université, CNRS, Université de Toulon, LIS, Marseille, France
victor.chepoi@lis-lab.fr

Shay Moran (ID)
Department of Computer Science, Princeton University, Princeton, USA
shaym@cs.princeton.edu

Manfred K. Warmuth
Computer Science Department, University of California, Santa Cruz, USA
manfred@ucsc.edu

──── **Abstract** ────

We examine connections between combinatorial notions that arise in machine learning and topological notions in cubical/simplicial geometry. These connections enable to export results from geometry to machine learning. Our first main result is based on a geometric construction by H. Tracy Hall (2004) of a partial shelling of the cross-polytope which can not be extended. We use it to derive a maximum class of VC dimension 3 that has no corners. This refutes several previous works in machine learning from the past 11 years. In particular, it implies that the previous constructions of optimal unlabeled compression schemes for maximum classes are erroneous.

On the positive side we present a new construction of an optimal unlabeled compression scheme for maximum classes. We leave as open whether our unlabeled compression scheme extends to ample (a.k.a. lopsided or extremal) classes, which represent a natural and far-reaching generalization of maximum classes. Towards resolving this question, we provide a geometric characterization in terms of unique sink orientations of the 1-skeletons of associated cubical complexes.

1 Introduction

The Sauer-Shelah-Perles Lemma [28, 30, 32] is arguably the most basic fact in VC theory; it asserts that any class $C \subseteq \{0,1\}^n$ satisfies $|C| \leq \binom{n}{\leq d}$, where $d = \text{VC-dim}(C)$. A beautiful generalization of Sauer-Shelah-Perles's inequality asserts that $|C| \leq |\overline{X}(C)|$, where $\overline{X}(C)$ is the family of subsets that are shattered by C.[1] The latter inequality is a part of the Sandwich Lemma [1, 4, 7, 23], which also provides a lower bound for $|C|$ (and thus "sandwiches" $|C|$) in terms of the number of its *strongly shattered subsets* (see Section 2). A class C is called *maximum/ample* if the Sauer-Shelah-Perles/Sandwich upper bounds are tight (respectively). Every maximum class is ample, but not vice versa.

Maximum classes were studied mostly in discrete geometry and machine learning, e.g. [34, 11, 9, 10, 14]. The history of ample classes is more interesting as they were discovered independently by several works in disparate contexts [1, 15, 4, 19, 3, 7, 35]. Consequently, they received different names such as lopsided classes [15], extremal classes [4, 19], and ample classes [3, 7]. Lawrence [15] was the first to define them for the investigation of the possible sign patterns realized by points of a convex set of $\mathbb{R}^d$. Interestingly, Lawrence's definition of these classes does not use the notion of shattering nor the Sandwich Lemma. In this context, these classes were discovered by Bollobás and Radcliffe [4] and Bandelt et al. [3], and the equivalence between the two definitions appears in [3]. Ample classes admit a multitude of combinatorial and geometric characterizations [3, 4, 15] and comprise many natural examples arising from discrete geometry, combinatorics, graph theory, and geometry of groups [3, 15].

Main Results

Corner Peelings. A *corner* in an ample class C is any concept $c \in C$ that belongs to a unique maximal cube of C (equivalently, c is a corner if $C \setminus \{c\}$ is also ample, see Claim 6). A sequence of corner removals leading to a single concept is called a *corner peeling*. Wiedemann [35] and independently Chepoi (unpublished, 1996) asked whether every ample class has a corner. The machine learning community studied this question independently in the context of *sample compression schemes* for maximum classes: Rubinstein and Rubinstein [25] showed that corner peelings lead to optimal *unlabeled sample compression schemes (USCS)*.

In Theorem 9 we refute this conjecture. The crux of the proof is an equivalence between corner peelings and partial shellings of the cross-polytope. This equivalence translates the question whether corners always exist to the question whether partial shellings can always be extended. The latter was an open question in Ziegler's book on polytopes [38], and was resolved in H. Tracy Hall's PhD thesis where she presented an interesting counterexample [13]. The ample class resulting from Hall's construction yields a maximum class without corners.

Sample Compression. Sample compression is a powerful technique to derive generalization bounds in statistical learning. Littlestone and Warmuth [16] introduced it and asked if every class of VC-dimension $d < \infty$ has a sample compression scheme of a finite size. This question was later precised by Floyd and Warmuth [10, 33] to *whether a sample compression scheme of size $O(d)$ exists*. The first question was recently resolved by [21] who exhibited an $\exp(d)$ sample compression. The second question however remains one of the oldest open problems in machine learning (for more background we refer the reader to [20] and the books [29, 36]).

---

[1] Note that this inequality indeed implies the Sauer-Shelah-Perles Lemma, since $|\overline{X}(C)| \leq \binom{n}{\leq d}$.

Rubinstein and Rubinstein [25, Theorem 16] showed that the existence of a corner peeling for a maximum class C implies a *representation map* for C (see Section 3 for a definition), which is known to yield an optimal unlabeled sample compression scheme of size VC-dim(C) [14].[2] They claim, using an interesting topological approach, that maximum classes admit corner peelings. Unfortunately, our Theorem 9 shows that this does not hold.

While our Theorem 9 rules out the program of deriving representation maps from corner peelings, in Theorem 10 we provide an alternative derivation of representation maps for maximum classes and therefore also of an unlabeled sample compression scheme for them.

Sample Compression and Unique Sink Orientations. We next turn to construction of representation maps for ample classes. In Theorem 15 we present geometric characterizations of such maps via *unique sink orientations*: an orientation of the edges of a cube B is a *unique sink orientation (USO)* if any subcube $B' \subseteq B$ has a unique sink. Szabó and Welzl [31] showed that any USO of B leads to a representation map for B. We extend this bijection to ample classes C by proving that representation maps are equivalent to orientations o of C such that (i) o is a USO on each subcube $B \subseteq C$, and (ii) for each $c \in C$ the edges outgoing from c belong to a subcube $B \subseteq C$. We further show that any ample class admits orientations satisfying each one of those conditions. However, *the question whether all ample classes admit representation maps remains open.*

Implications on Previous Works. Our Theorem 9 establishes the existence of maximum classes without any corners, thus countering several previous results in machine learning:

- Rubinstein and Rubinstein [25, Theorem 32] showed that any maximum class can be represented by a simple arrangement of piecewise-linear hyperplanes. In [25, Theorem 39], they claim that sweeping such an arrangement leads to a corner peeling of the corresponding maximum class. This is unfortunately false, as witnessed by Theorem 9.

- Kuzmin and Warmuth [14] constructed unlabeled sample compression schemes for maximum classes based on the presumed uniqueness of a certain matching (their Theorem 10). This theorem is wrong as it implies the existence of corners. However their conclusion is correct: in our Theorem 10 we show that such unlabeled compression schemes exist.

- Theorem 3 by Samei, Yang, and Zilles [27] is built on a generalization of Theorem 10 from [14] to the multiclass case which is also incorrect.

- Theorem 26 by Doliwa et al. [6] uses the result by [25] to show that the Recursive Teaching Dimension (RTD) of maximum classes equals to their VC dimension. However the VC dimension 3 maximum class from Theorem 9 has RTD at least 4. It remains open whether the RTD of every maximum class C is bounded by $O(\text{VC-dim}(C))$.

Organization. Section 2 presents the main definitions and notations. Section 3 reviews characterizations of ample/maximum classes and presents characteristic examples. Section 4 demonstrates the existence of the maximum class C_H without corners. Section 5 establishes the existence of representation maps for maximum classes. Section 6 establishes a bijection between representation maps and unique sink orientations for ample classes. Due to space limitations, some proofs are omitted and can be found in the full version of this paper [5].

---

[2] Pálvölgyi and Tardos [24] recently exhibited a (non-ample) class C with no USCS of size VC-dim(C).

2 Preliminaries

A *concept class* C is a set of subsets (concepts) of a finite ground set U which is called the *domain* of C and denoted $\text{dom}(C)$. We sometimes treat the concepts as characteristic functions rather than subsets. The *support* (or *dimension set*) $\text{supp}(C)$ of C is the set $\{x \in U : x \in c' \setminus c'' \text{ for some } c', c'' \in C\}$. $C^* := 2^U \setminus C$ is the *complement* of C. The *restriction* of C on $Y \subseteq U$ is the class $C|Y = \{c \cap Y : c \in C\}$ whose domain is Y. We use C_Y as shorthand for $C|(U \setminus Y)$; in particular, we write C_x for $C_{\{x\}}$, and c_x for $c|(U \setminus \{x\})$ for $c \in C$ (note that $c_x \in C_x$). A class $B \subseteq 2^U$ is a *cube* if there exists $Y \subseteq U$ such that $B|Y = 2^Y$ and B_Y contains a single concept (denoted by $\text{tag}(B)$). Note that $\text{supp}(B) = Y$ and therefore we say that B is a *Y-cube*; $|Y|$ is called the *dimension* $\dim(B)$ of B. Two cubes B, B' with the same support are called *parallel cubes*. A cube B is *maximal* if there is no cube B' such that $B \subsetneq B'$.

Let Q_n denote the *n-dimensional cube* where $n = |U|$; $c, c' \in Q_n$ are called adjacent if the symmetric difference $c \Delta c'$ is of size 1. The *1-inclusion graph* of C is the subgraph $G(C)$ of Q_n induced by the vertex-set C when the concepts of C are identified with the corresponding vertices of Q_n. Any cube $B \subseteq C$ is called *a cube of C*. The *cube complex* of C is the set $Q(C) = \{B : B \text{ is a cube of } C\}$. The *dimension* of $Q(C)$ is $\dim(Q(C)) := \max_{B \in Q(C)} \dim(B)$. A concept $c \in C$ is called a *corner* of C if c belongs to a unique maximal cube of C. The *reduction* C^Y of C to $Y \subseteq U$ is the class $C^Y := \{\text{tag}(B) : B \in Q(C) \text{ and } \text{supp}(B) = Y\}$ whose domain is $U \setminus Y$. When $x \in U$ we denote $C^{\{x\}}$ by C^x and call it *the x-hyperplane of C*. Note that a concept c belongs to C^x if and only if c and $c \cup \{x\}$ both belong to C. The union of all cubes of C having x in their support is called the *carrier* of C^x and is denoted by $N_x(C)$. If $c \in N_x(C)$, we also denote $c|U \setminus \{x\}$ by c^x (note that $c^x \in C^x$).

A class C is *connected* if the graph $G(C)$ is connected. Let $d_{G(C)}(c, c')$ denote the distance between c and c' in $G(C)$. Note that $d_{Q_n}(c, c') =: d(c, c')$ coincides with the Hamming distance $|c \Delta c'|$. Let $B(c, c') = \{t \subseteq U : d(c, t) + d(t, c') = d(c, c')\}$ be the *interval* between c and c' in Q_n. A class C is called *isometric* if $d(c, c') = d_{G(C)}(c, c')$ for any $c, c' \in C$ and *weakly isometric* if $d(c, c') = d_{G(C)}(c, c')$ if $d(c, c') \leq 2$. Any path connecting two concepts $\text{tag}(B)$ and $\text{tag}(B')$ of C^Y inside C^Y can be lifted to a path of Y-cubes connecting B and B' in C; such a path of cubes is called a *gallery*.

A class C *shatters* $Y \subseteq U$ if $C|Y = 2^Y$. C *strongly shatters* Y if C contains a Y-cube. Let $\overline{X}(C)$, $\underline{X}(C)$ denote the simplicial complexes $\overline{X}(C) = \{Y : C \text{ shatters } Y\}$, $\underline{X}(C) = \{Y : C \text{ strongly shatters } Y\}$. Note that $\underline{X}(C) \subseteq \overline{X}(C)$. The *VC-dimension* $\text{VC-dim}(C)$ of C is the size of a largest set C shatters. The *Sandwich Lemma* asserts that $|\underline{X}(C)| \leq |C| \leq |\overline{X}(C)|$.

A *labeled sample* is a set $s = \{(x_1, y_1), \ldots, (x_m, y_m)\}$, where $x_i \in U$ and $y_i \in \{0, 1\}$. An *unlabeled sample* is a set $\{x_1, \ldots, x_m\}$, where $x_i \in U$. Given a labeled sample $s = \{(x_1, y_1), \ldots, (x_m, y_m)\}$, the unlabeled sample $\{x_1, \ldots, x_m\}$ is the domain of s and is denoted by $\text{dom}(s)$. A sample s is *realizable* by a concept $c : U \to \{0, 1\}$ if $c(x_i) = y_i$ for every i, and s is realizable by a concept class C if it is realizable by some $c \in C$.

A *sample compression scheme* for a concept class C is best viewed as a protocol between a *compressor* and a *reconstructor*. The compressor gets a realizable sample s from which it picks a small subsample s'. The compressor sends s' to the reconstructor. Based on s', the reconstructor outputs a concept c that needs to be consistent with the entire input sample s. A sample compression scheme has size k if for every realizable input sample s the size of the compressed subsample s' is at most k. An *unlabeled (sample) compression scheme* (USCS) is a sample compression scheme in which the compressed subsample s' is unlabeled. So, the compressor removes the labels before sending the subsample to the reconstructor.

3 Ample and Maximum Classes

We briefly review the main characterizations and basic geometric examples of ample and maximum classes. The next theorem summarizes the main characterizations of ample classes:

▶ **Theorem 1** ([3, 4, 15]). *The following are equivalent for a class C: (1) C is ample; (2) C^* is ample; (3) $\underline{X}(C) = \overline{X}(C)$; (4) $|\underline{X}(C)| = |C|$; (5) $|\overline{X}(C)| = |C|$; (6) $C \cap B$ is ample for any cube B; (7) $(C^Y)_Z = (C_Z)^Y$ for all partitions $U = Y \uplus Z$; (8) for all partitions $U = Y \uplus Z$, either $Y \in \underline{X}(C)$ or $Z \in \underline{X}(C^*)$.*

Condition (3) leads to a simple definition of ampleness: C is ample if whenever $Y \subseteq U$ is shattered by C, then there is a Y-subcube of C. Thus, if C is ample we will write $X(C)$ instead of $\underline{X}(C) = \overline{X}(C)$. A *representation map* for an ample class C is a bijection $r : C \to X(C)$ satisfying the *non-clashing condition*: $c|(r(c) \cup r(c')) \neq c'|(r(c) \cup r(c'))$, for all $c, c' \in C, c \neq c'$. We continue with metric and recursive characterizations of ample classes:

▶ **Theorem 2** ([3]). *The following are equivalent for a class C: (1) C is ample; (2) C^Y is connected for all $Y \subseteq U$; (3) C^Y is isometric for all $Y \subseteq U$; (4) C is isometric, and both C_x and C^x are ample for all $x \in U$; (5) C is connected and all hyperplanes C^x are ample.*

▶ **Corollary 3.** *Two maximal cubes of an ample class C have different supports.*

Indeed, if B and B' are two d-cubes with the same support, by Theorem 2(2) B and B' can be connected in C by a gallery, and thus B is contained in a $d + 1$-cube. Therefore, B and B' cannot be maximal.

The Sandwich Lemma and Theorem 1(5) imply that maximum classes are ample. Basic examples of maximum classes are concept classes derived from hyperplane arrangements in $\mathbb{R}^n$, ball arrangements in $\mathbb{R}^n$, and unions of n intervals in $\mathbb{R}$. The following theorem summarizes some characterizations of maximum classes provided in [11, 9, 10, 34]:

▶ **Theorem 4.** *The following are equivalent for a class C: (1) C is maximum; (2) C_Y is maximum for all $Y \subseteq U$; (3) C_x and C^x are maximum for all $x \in U$; (4) C^* is maximum.*

We continue with some important geometric examples of ample classes.

1. Simplicial Complexes. Every simplicial complex S (viewed as a set system closed under taking subsets) is ample since $\overline{X}(S) = \underline{X}(S)$.

2. Realizable Ample Classes. Let $K \subseteq \mathbb{R}^n$ be a convex set. Let $C(K) := \{\mathrm{sign}(v) : v \in K, v_i \neq 0 \, \forall i \leq n\}$, where $\mathrm{sign}(v) \in \{\pm 1\}^n$ is the sign pattern of v. Lawrence [15] showed that $C(K)$ is ample, and called ample classes representable in this manner *realizable*.

3. Median Classes. A class C is called *median* if for every three concepts c_1, c_2, c_3 of C their *median $m(c_1, c_2, c_3) := (c_1 \cap c_2) \cup (c_1 \cap c_3) \cup (c_2 \cap c_3)$* also belongs to C. Median classes are ample by [3, Proposition 2]. Due to their relationships with other discrete structures, median classes are one of the most important examples of ample classes. Median classes are equivalent to finite median graphs (a well-studied class in metric graph theory, see [2]), to CAT(0) cube complexes, i.e., cube complexes of global nonpositive curvature (central objects in geometric group theory, see [12, 26]), and to the domains of event structures (a basic model in concurrency theory [22, 37]).

4. Convex Geometries and Conditional Antimatroids. Let C be a class such that (i) $\varnothing \in C$ and (ii) $c, c' \in C$ implies that $c \cap c' \in C$. Call $x \in c \in C$ *extremal* if $c \setminus \{x\} \in C$. We say that $c \in C$ is *generated* by $s \subseteq c$ if c is the smallest member of C containing s. A class C satisfying (i) and (ii) with the additional property that every member c of C is generated by its extremal points is called a *conditional antimatroid* [3, Section 3]. If $U \in C$, then we obtain the well-known structure of a *convex geometry* (called also an *antimatroid*) [8]. By [3, Proposition 1], conditional antimatroids C are ample since $\underline{X}(C)$ coincides with the sets of extremal points and $\overline{X}(C)$ coincides with the set of all minimal generating sets of sets from C. Convex geometries comprise many examples from geometry, ordered sets, and graphs; see the foundational paper [8]. For example, a *realizable convex geometry* is a convex geometry $C \subseteq U$ such that U can be realized as a set of $\mathbb{R}^n$ and $c \in C$ if and only if c is the intersection of a convex set of $\mathbb{R}^n$ with U.

4   Corner Peelings and Partial Shellings

In this section, we prove that corner peelings of ample classes are equivalent to isometric orderings of C as well as to partial shellings of the cross-polytope. This equivalence, combined with a result by Hall [13] yields a maximum class with VC dimension 3 without corners (Theorem 9 below). Let $C_< := (c_1, \ldots, c_m)$ be an ordering of the concepts in C. For any $1 \le i \le m$, let $C_i := \{c_1, \ldots, c_i\}$ denote the *i'th level set*. The ordering $C_<$ is called:

- an *ample* ordering if every level set C_i is ample;
- a *corner peeling* if every c_i is a *corner* of C_i;
- an *isometric* ordering if every level set C_i is isometric;
- a *weakly isometric* ordering if every level set C_i is weakly isometric.

▶ **Proposition 5.** *The following are equivalent for an ordering $C_<$ of an isometric class C: (1) $C_<$ is ample; (2) $C_<$ is a corner peeling; (3) $C_<$ is isometric; (4) $C_<$ is weakly isometric.*

Proof. Clearly, (3)⇒(4). Conversely, suppose $C_<$ is weakly isometric but one of its levels is not isometric. Hence, there exists $i < j$ such that any shortest (c_i, c_j)-path in C contains some c_k with $k > j$. Additionally, assume that c_i, c_j minimizes the distance $d(c_i, c_j)$ among all such pairs. Since C_j is weakly isometric, necessarily $d(c_i, c_j) \ge 3$. Let c_r be the first concept among $\{c_{j+1}, \ldots, c_m\}$ lying in $B(c_i, c_j) \cap C$. If $d(c_i, c_r) \ge 3$ or $d(c_r, c_j) \ge 3$ (say the first), then one can replace c_i, c_j by c_i, c_r, which contradicts the choice of c_i, c_j. Thus, $d(c_i, c_r), d(c_r, c_j) \le 2$, and at least one of them equals 2 (say $d(c_i, c_r) = 2$). Now, weak isometricity implies that c_i and c_r have a common neighbor c_ℓ with $\ell < \max\{i, r\} = r$. If $\ell < j$ then c_ℓ, c_j contradicts the minimality of c_i, c_j, and if $j < \ell < r$ then c_ℓ contradicts the minimality of c_r. This shows (4)⇒(3).

If c_i is a corner of C_i, then any two neighbors of c_i in C_i have a second common neighbor in C_i, and therefore $d_{G(C_{i-1})}$ is the restriction of $d_{G(C_i)}$ on C_{i-1}. Since $C_m = C$ is isometric, this proves (2)⇒(3). We now prove (3)⇒(1)⇒(2) using the next lemma. For $t \notin C$, let $F[t]$ be the smallest cube of Q_n containing t and all neighbors of t in $G(C)$. Note that the dimension of $F[t]$ is the number of neighbors of t in $G(C)$.

▷ **Claim 6.** Let C be ample. Then: *(i)* If $t \notin C$ then $F[t] \subseteq C \cup \{t\}$. *(ii)* If c is a corner of C then $C \setminus \{c\}$ is ample. *(iii)* If $t \notin C$ and $C' := C \cup \{t\}$ is isometric then C' is ample and t is a corner of C'.

Proof. Item (i): Suppose $F[t] \setminus C \neq \{t\}$. Pick $s \neq t$ that is closest to t in $F[t] \setminus C$ (with respect to the Hamming distance of Q_n). Then t and s are not adjacent (by the definition of $F[t]$). By the choice of s, $B(s,t) \setminus \{s,t\} \subseteq C$, i.e., $B(s,t) \cap C^* = \{t,s\}$, contrary to the ampleness of C^*.

Item (ii): If $c \in C$ is a corner then there is a unique maximal cube $F \subseteq C$ containing it. Combined with Corollary 3, this implies that $\underline{X}(C \setminus \{c\}) = \underline{X}(C) \setminus \{\text{supp}(F)\}$. Next, since $|C| = |\underline{X}(C)|$, we get that $|C \setminus \{c\}| = |\underline{X}(C \setminus \{c\})|$, and by Theorem 1 $C \setminus \{c\}$ is ample.

Item (iii): To prove that C' is ample, we use Theorem 2(2). First note that by item (i), $F[t] \subseteq C'$. Let $F' \neq F''$ be parallel cubes of C'. If $t \notin F' \cup F''$, then a gallery connecting F' and F'' in C is a gallery in C'. So, assume $t \in F'$. If F' is a proper face of $F[t]$, then F' is parallel to a face F of $F[t]$ not containing t. Since F' and F are connected in $F[t]$ by a gallery and F and F'' are connected in C by a gallery, we obtain a gallery between F' and F'' in C'. Finally, let $F' = F[t]$. In this case, we assert that F'' does not exist. Otherwise, let π be the parallelism map between F' and F'' (π maps each concept in F' to its unique closest concept in F''). Note that for any $r \in F'$: $d(t, \pi(t)) = d(r, \pi(r)) = d(F', F'')$. Since C' is isometric, t and $\pi(t)$ can be connected in C' by a path P of length $d(t, \pi(t))$. Let s be the neighbor of t in P. Since $s \in C$ it follows that $s \in F[t] = F'$. So, s is a concept in F' that is closer to $\pi(t)$ than t; this contradicts that $d(t, t'') = d(F', F'')$. ◁

To show (1)⇒(2), let $C_<$ be an ample order of C. We assert that each c_i is a corner of C_i. Indeed, since C_{i-1} is ample and $c_i \notin C_{i-1}$, by Item (i) in Claim 6 the cube $F[c_i]$, defined with respect to C_{i-1}, is included in C_i. Thus, c_i belongs to a unique maximal cube $F[c_i]$ of C_i, i.e., c_i is a corner of C_i. To prove (3)⇒(1), let $C_<$ be isometric. The ampleness of each C_i follows by induction from Item (iii) of Claim 6. ◀

A concept class C is *dismantlable* if it admits an ordering satisfying any of the equivalent conditions (1)-(4) in Proposition 5. Isometric orderings of Q_n are closely related to shellings of its dual, the *cross-polytope* O_n (which we define next). Define $\pm U := \{\pm x_1, \ldots, \pm x_n\}$; so, $|\pm U| = 2n$, and we call $-x_i, +x_i$ *antipodal*. The n-dimensional *cross-polytope* is the pure simplicial complex of dimension n whose facets are all $\sigma \subseteq \pm U$ that contain exactly one element in each antipodal pair. Thus, O_n has 2^n facets and each facet σ of O_n corresponds to a vertex c of Q_n ($+x_i \in \sigma$ if and only if $x_i \in c$). Observe that $x_i \in c' \Delta c''$ if and only if $\{+x_i, -x_i\} \subseteq \sigma' \Delta \sigma''$ where σ' correspond to c' and σ'' corresponds to c''.

Let X be a pure simplicial complex (PSC) of dimension d, i.e., a simplicial complex in which all facets have size d. Two facets σ, σ' are adjacent if $|\sigma \Delta \sigma'| = 2$. A *shelling* of X is an ordering $\sigma_1, \ldots, \sigma_p$ of all of its facets such that $2^{\sigma_j} \bigcap (\bigcup_{i<j} 2^{\sigma_i})$ is a PSC of dimension $d - 1$ for every $j \leq p$ [38, Lecture 8]. A *partial shelling* is an ordering of some facets that satisfies the above condition. Note that $\sigma_1, \ldots, \sigma_m$ is a partial shelling if and only if for every $i < j$ there exists $k < j$ such that $\sigma_i \cap \sigma_j \subseteq \sigma_k \cap \sigma_j$, and $\sigma_k \cap \sigma_j$ is a facet of both σ_j and σ_k. X is *extendably shellable* if every partial shelling can be extended to a shelling. We next establish a relationship between partial shellings and isometric orderings.

▶ **Proposition 7.** *Every partial shelling of the cross-polytope O_n defines an isometric ordering of the corresponding vertices of the cube Q_n. Conversely, if C is an isometric class of Q_n, then any isometric ordering of C defines a partial shelling of O_n.*

Proof. Let $\sigma_1, \ldots, \sigma_m$ be a partial shelling of O_n and $c_1, \ldots, c_m$ be the ordering of the corresponding vertices of Q_n. We need to prove that each level set $C_j = \{c_1, \ldots, c_j\}$ is isometric. It suffices to show that for every $i < j$ there is $k < j$ such that $d(c_k, c_j) = 1$ and $c_k \in B(c_i, c_j)$. Equivalently, that $|\sigma_k \Delta \sigma_j| = 2$ and $\sigma_i \cap \sigma_j \subseteq \sigma_k \subseteq \sigma_i \cup \sigma_j$: since

$\sigma_1, \ldots, \sigma_m$ is a partial shelling, there is a facet σ_k with $k < j$ such that $|\sigma_k \cap \sigma_j| = n - 1$ and $\sigma_i \cap \sigma_j \subseteq \sigma_k \cap \sigma_j$. We claim that σ_k is the desired facet. It remains to show that (i) $|\sigma_j \Delta \sigma_k| = 2$ and (ii) $\sigma_k \subseteq \sigma_i \cup \sigma_j$. Item (i) follows since $|\sigma_j| = |\sigma_k| = n$, and $|\sigma_k \cap \sigma_j| = n-1$. For Item (ii), let $\sigma_j \setminus \sigma_k = \{+x\}$ and $\sigma_k \setminus \sigma_j = \{-x\}$. We need to show that $-x \in \sigma_i$, or equivalently that $+x \notin \sigma_i$. The latter follows since $+x \in \sigma_j \setminus \sigma_k$ and $\sigma_j \cap \sigma_i \subseteq \sigma_k$.

Conversely, let $c_1, \ldots, c_m$ be an isometric ordering and $\sigma_1, \ldots, \sigma_m$ be the ordering of the corresponding facets of O_n. We assert that this is a partial shelling. Let $i < j$. It suffices to exhibit $k < j$ such that $|\sigma_k \cap \sigma_j| = n - 1$ and $\sigma_i \cap \sigma_j \subseteq \sigma_k \cap \sigma_j$. Since C_j is isometric, c_j has a neighbor $c_k \in B(c_i, c_j) \cap C_j$. Since $d(c_j, c_k) = 1$ it follows that $|\sigma_k \cap \sigma_j| = n - 1$. Since $c_k \in B(c_i, c_j)$ it follows that $\sigma_i \cap \sigma_j \subset \sigma_k \subset \sigma_i \cup \sigma_j$ and hence that $\sigma_i \cap \sigma_j \subseteq \sigma_k \cap \sigma_j$.   ◀

▶ **Corollary 8.** *If all ample classes are dismantlable, then O_n is extendably shellable.*

Proof. Let $\sigma_1, \ldots, \sigma_m$ be a partial shelling of O_n and let $C = \{c_1, \ldots, c_m\}$ be the corresponding vertices of Q_n. By Proposition 7, the level sets are isometric, thus C is ample by Proposition 5. The complement C^* is also ample, thus dismantlable. Thus C^* contains a concept t such that $C^* \setminus \{t\}$ is ample. Consequently, $C' := C \cup \{t\}$ is ample. Let τ be the facet of O_n corresponding to t. Since $c_1, \ldots, c_m, t$ is an isometric ordering of C', by Proposition 7, $\sigma_1, \ldots, \sigma_m, \tau$ is a partial shelling of O_n.   ◀

It was asked in [38] if any cross-polytope O_n is extendably shellable. In the PhD thesis of H. Tracy Hall from 2004, a nice counterexample to this question is given [13]. Hall's counterexample arises from the 299 regions of an arrangement of 12 pseudo-hyperplanes. These regions are encoded as facets of the cross-polytope O_{12} and it is shown in [13] that the subcomplex of O_{12} consisting of all other facets admits a shelling which cannot be extended. By the proof of Corollary 8, the ample concept class C_H defined by those 299 simplices does not have any corner.[3] A counting shows that C_H is a maximum class of VC-dimension 3. This completes the proof of our first main result:

▶ **Theorem 9.** *There exists a maximum class C_H of VC-dimension 3 without any corner.*

However, conditional antimatroids and 2-dimensional ample classes are dismantlable. The 2-dimensional case was proved in [25, Theorem 34] for maximum classes and in [18] for ample classes. The proof for conditional antimatroids appears in the full version of this paper [5], as well as a different proof for the 2-dimensional case that is based on a local characterization of convex sets of ample classes.

5   Representation Maps for Maximum Classes

In this section, we prove that maximum classes admit representation maps and therefore by [14, Lemma 1], they admit optimal unlabeled compression schemes.

▶ **Theorem 10.** *Any maximum class $C \subseteq 2^U$ of VC-dimension d admits a representation map, and consequently, an unlabeled compression scheme of size d.*

The crux of the proof of Theorem 10 is the following proposition. Let C be a d-dimensional maximum class and let $D \subseteq C$ be a $(d-1)$-dimensional maximum subclass. A *missed simplex* for the pair (C, D) is a simplex $\sigma \in X(C) \setminus X(D)$. Note that any missed simplex has size d.

---

[3] For the interested reader, a file containing the 299 concepts of C_H represented as elements of $\{0, 1\}^{12}$ is available at `https://arxiv.org/src/1812.02099/anc/CH.txt`.

An *incomplete cube* Q for (C, D) is a cube of C such that $\mathrm{supp}(Q)$ is a missed simplex. For any incomplete cube Q with $\sigma = \mathrm{supp}(Q)$, $C|\sigma$ and $D|\sigma$ are maximum classes of dimensions d and $d-1$, respectively. Since $|\sigma| = d$, we have $|C|\sigma| = \binom{d}{\leq d} = \binom{d}{\leq d-1} + 1 = |D|\sigma| + 1$. Since $Q|\sigma = C|\sigma$, there exists a unique concept $c \in Q$ such that $c|\sigma \notin D|\sigma$. We denote c by $s(Q)$, and call c the *source* of Q. In fact, the source map is a bijection between missed simplices for (C, D) and concepts of $C \setminus D$:

▶ **Proposition 11.** *Each $c \in C \setminus D$ is the source of a unique incomplete cube. Moreover, if $r' : D \to X(D)$ is a representation map for D and $r : C \to X(C)$ extends r' by setting $r(c) = \mathrm{supp}(s^{-1}(c))$ for each $c \in C \setminus D$, then r is a representation map for C.*
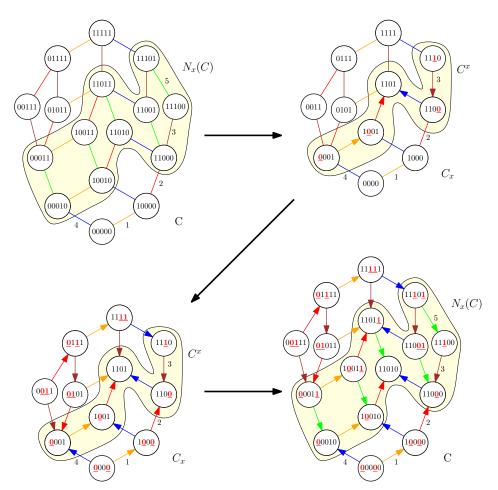

Figure 1 Illustrating the proof of Theorem 10 (when $x = 5$): to construct a representation map for C, we inductively construct a representation map r^x for C^x, extend it to a representation map r_x for C_x using Proposition 11 with $D = C^x$, and finally extend it to a representation map r for C. The representation maps r^x, r_x, and r are defined by the orientation as in Theorem 15 and by the coordinates of the underlined bits.

Proof of Theorem 10. Following the general idea of [14], we derive a representation map for C by induction on $|U|$. For the induction step (see Fig. 1), pick $x \in U$ and consider the maximum classes C_x and $C^x \subset C_x$ with domain $U \setminus \{x\}$. By induction, C^x has a representation map r^x. Use Proposition 11 to extend r^x to a representation map r_x of C_x. Define a map r on C as follows:

- $r(c) = r_x(c_x)$ if $c_x \notin C^x$ or $x \notin c$,
- $r(c) = r_x(c_x) \cup \{x\}$ if $c_x \in C^x$ and $x \in c$.

It is easy to verify that r is non-clashing: indeed, if $c' \neq c'' \in C$ satisfy $c'_x \neq c''_x$ then $c'_x | r_x(c'_x) \cup r_x(c''_x) \neq c''_x | r_x(c'_x) \cup r_x(c''_x)$. Since $r_x(c'_x) \subseteq r(c'), r_x(c''_x) \subseteq r(c'')$, it follows that also c', c'' disagree on $r(c') \cup r(c'')$. Else, $c'_x = c''_x \in C^x$ and $c'(x) \neq c''(x)$. In this case, $x \in r(c') \cup r(c'')$ and therefore c', c'' disagree on $r(c') \cup r(c'')$.

It remains to show that r is a bijection between C and $X(C) = \binom{U}{\leq d}$. It is easy to verify that r is injective. So, it remains to show that $|r(c)| \leq d$, for every $c \in C$. This is clear when $c_x \notin C^x$ or $x \notin c$. If $c_x \in C^x$ and $x \in c$, then $r(c) = r^x(c_x) \cup \{x\}$ and $|r^x(c_x)| \leq d - 1$ (since C^x is $(d-1)$-dimensional). Hence, $|r(c)| \leq d$ as required, concluding the proof. ◀

Proof of Proposition 11. Call a maximal cube of C a *chamber* and a facet of a chamber a *panel* (a σ'-*panel* if its support is σ'). Any σ'-panel in C satisfies $|\sigma'| = d - 1$ and $\sigma' \in X(D)$. Recall that a gallery between two parallel cubes Q', Q'' (say, two σ'-cubes) is any simple path of σ'-cubes $(Q_0 := Q', Q_1, \ldots, Q_k := Q'')$, where $Q_i \cup Q_{i+1}$ is a d-cube. By Theorem 2(3), any two parallel cubes of C are connected by a gallery in C. Since D is a maximum class, any panel of C is parallel to a panel that is a maximal cube of D. Also for any maximal simplex $\sigma' \in X(D)$, the class $C^{\sigma'}$ is a maximum class of dimension 1 and $D^{\sigma'}$ is a maximum class of dimension 0 (single concept). Thus $C^{\sigma'}$ is a tree (e.g. [11, Lemma 7]) which contains the unique concept $c \in D^{\sigma'}$. We call c *the root* of $C^{\sigma'}$ and denote the σ'-panel P such that $P^{\sigma'} = c$ by $P(\sigma')$.

▷ **Claim 12.** Let Q be an incomplete cube for (C, D) with source s and support σ, and let $x, y \in U$ such that $x \notin \sigma$ and $y \in \sigma$. Then, the following holds:
 (i) Q_x is an incomplete cube for (C_x, D_x) whose source is s_x.
 (ii) Q^y is an incomplete cube for (C^y, D^y) whose source is s^y.

Proof. Item (i): C_x and D_x are maximum classes on $U \setminus \{x\}$ of VC-dimensions d and $d - 1$, and $\text{supp}(Q_x) = \sigma$. Therefore, Q_x is an incomplete cube for (C^x, D^x). By definition, s is the unique concept $c \in Q$ such that $c|\sigma \notin D|\sigma$. Since $x \notin \sigma$, $D|\sigma = D_x|\sigma$ and s_x is the unique concept c of Q_x so that $c|\sigma \notin D_x|\sigma$, i.e., s_x is the source of Q_x.

Item (ii): C^y and D^y are maximum classes on $U \setminus \{y\}$ of VC-dimensions $d - 1$ and $d - 2$. Since $y \in \text{supp}(Q)$, $\dim(Q^y) = d - 1$ and Q^y is an incomplete cube for (C^y, D^y). Let $\sigma' = \sigma \setminus \{y\}$. It remains to show that $s^y|\sigma' \notin D^y|\sigma'$. Indeed, otherwise both extensions of s^y in σ, namely $s, s\Delta\{y\}$, are in $D|\sigma$ which contradicts that $s = s(Q)$. ◁

Next we prove that each concept of $C \setminus D$ is the source of a unique incomplete cube. Assume the contrary and let (C, D) be a counterexample minimizing the size of U. First, if a concept $c \in C \setminus D$ is the source of two incomplete cubes Q_1, Q_2, then $\text{dom}(C) = \text{supp}(Q_1) \cup \text{supp}(Q_2)$. Indeed, let $\sigma_1 = \text{supp}(Q_1)$ and $\sigma_2 = \text{supp}(Q_2)$. By Claim 12(i) and minimality of (C, D), $\text{dom}(C) = \sigma_1 \cup \sigma_2$. By Claim 12(ii) and minimality of (C, D), $\sigma_1 \cap \sigma_2 = \varnothing$. Indeed, if there exists x in $\sigma_1 \cap \sigma_2$, c^x is the source of the incomplete cubes Q_1^x and Q_2^x for (C^x, D^x), contrary to minimality of (C, D).

Next we assert that any $c \in C \setminus D$ is the source of at most 2 incomplete cubes. Indeed, let c be the source of incomplete cubes Q_1, Q_2, Q_3. Then $\text{dom}(C) = \text{supp}(Q_1) \cup \text{supp}(Q_2)$, i.e., $\text{supp}(Q_2) = \text{dom}(C) \setminus \text{supp}(Q_1)$. For similar reasons, $\text{supp}(Q_3) = \text{dom}(C) \setminus \text{supp}(Q_1) = \text{supp}(Q_2)$. Thus, by Corollary 3, $Q_2 = Q_3$.

▷ **Claim 13.** Let $c', c'' \in C \setminus D$ be neighbors and let $c'\Delta c'' = \{x\}$. Then, c' is the source of 2 incomplete cubes if and only if c'' is the source of 0 incomplete cubes. Consequently, every connected component in $G(C \setminus D)$ either contains only concepts c with $|s^{-1}(c)| \in \{0, 2\}$, or only concepts c with $|s^{-1}(c)| = 1$.

Proof. By minimality of (C, D), $(c')^x = (c'')^x$ is the source of a unique incomplete cube for (C^x, D^x) and $c'_x = c''_x$ is the source of a unique incomplete cube for (C_x, D_x). Let Q_1 be the incomplete cube for (C, D) such that $(c')^x$ is the source of Q_1^x. Let Q_2 be the incomplete cube for (C, D) such that c'_x is the source of $(Q_2)_x$. By Claim 12, items (i) and (ii), both $s(Q_1), s(Q_2)$ are in $\{c', c''\}$. Consequently, c' is the source of 2 incomplete cubes (Q_1 and Q_2) if and only if c'' is the source of 0 incomplete cubes.   ◁

Pick $c \in C \setminus D$ that is the source of two incomplete cubes for (C, D) and an incomplete cube Q such that $c = s(Q)$. Let $\sigma = \operatorname{supp}(Q)$, $x \in \sigma$, and $\sigma' = \sigma \setminus \{x\}$. The concept c belongs to a unique σ'-panel P. Let $L = (P_0 = P(\sigma'), P_1, \ldots, P_{m-1}, P_m = P)$ be the unique gallery between the root $P(\sigma')$ of the tree C^σ and P. For $i = 1, \ldots, m$, denote the chamber $P_{i-1} \cup P_i$ by Q_i. Since $P_i \cap D$ and $Q_i \cap D$ are ample for $i \geq 0$, and P_i is not contained in D for $i > 0$, it follows that the complements $P_i \setminus D$ and $Q_i \setminus D$ are nonempty ample classes. Hence $P_i \setminus D$ and $Q_i \setminus D$ induce nonempty connected subgraphs of $G(C \setminus D)$. Therefore, it follows that c and each concept $c' \in Q_i \setminus D$ are connected in $G(C \setminus D)$ by a path for $i > 0$, and by Claim 13 it follows that

> For every $i > 0$, each $c' \in Q_i \setminus D$ is the source of either 0 or 2 incomplete cubes.   (5.1)

Consider the chamber $Q_1 = P_0 \cup P_1$ and its source $s = s(Q_1)$. By the definition of the source, necessarily $s \in P_1$ and $s \notin D$. Therefore, Equation (5.1) implies that there must exist another cube Q' such that $s = s(Q')$. Let s' be the neighbor of s in $P_0 = P(\sigma')$; note that $s' \in D$. Since $\operatorname{supp}(Q_1) \cap \operatorname{supp}(Q') = \varnothing$, it follows that $s|\operatorname{supp}(Q') = s'|\operatorname{supp}(Q') \in D|\operatorname{supp}(Q')$, contradicting that $s = s(Q')$. This establishes the first assertion of Proposition 11.

We prove now that the map r defined in Proposition 11 is a representation map for C. It is easy to verify that it is a bijection between C and $X(C)$, so it remain to establish the non-clashing property: $c|(r(c) \cup r(c')) \neq c'|(r(c) \cup r(c'))$ for all distinct pairs $c, c' \in C$. This holds when $c, c' \in D$ because r' is a representation map. Next, if $c \in C \setminus D$ and $c' \in D$, this holds because $c|r(c) \notin D|r(c)$ by the properties of s. Thus, it remains to show that every distinct $c, c' \in C \setminus D$ satisfy $c|(\operatorname{supp}(Q) \cup \operatorname{supp}(Q')) \neq c'|(\operatorname{supp}(Q) \cup \operatorname{supp}(Q'))$, where $Q = s^{-1}(c), Q' = s^{-1}(c')$. Assume towards contradiction that this does not hold and consider a counterexample with minimal domain size $|U|$. By minimality, $\operatorname{supp}(Q') \cup \operatorname{supp}(Q) = U$ (or else (C_x, D_x), for some $x \notin \operatorname{supp}(Q') \cup \operatorname{supp}(Q)$ would be a smaller counterexample). Therefore, since c, c' are distinct, there must be $x \in U = \operatorname{supp}(Q') \cup \operatorname{supp}(Q)$ such that $c(x) \neq c'(x)$, which is a contradiction. This ends the proof of Proposition 11.   ◀

6   Representation Maps for Ample Classes

In this section, we provide combinatorial and geometric characterizations of representation maps of ample classes (which lead to optimal unlabeled compression schemes).

▶ **Theorem 14.** *For an ample class C and a map $r : C \to X(C)$, (i)-(iii) are equivalent:*
  (i) *r is a representation map;*
  (ii) *$c'|r(c') \Delta r(c'') \neq c''|r(c') \Delta r(c'')$ for all $c', c'' \in C, c' \neq c''$;*
  (iii) *r is a bijection and for every realizable sample s of C, there is a unique $c \in C$ that is consistent with s and $r(c) \subseteq \operatorname{dom}(s)$.*

This theorem implies that for any representation map $r : C \to X(C)$ and any x-edge cc', $r(c) \Delta r(c') = \{x\}$. Hence, r defines an orientation o_r of $G(C)$: an x-edge cc' is oriented from c to c' iff $x \in r(c) \setminus r(c')$. Moreover, as a corollary of Theorem 14, we can show that:
(C1) for any $c \in C$, all outgoing neighbors of c belong to a cube of C;
(C2) o_r is a USO on each cube of C.

An orientation o of the edges of $G(C)$ is a *unique sink orientation (USO)* if o satisfies (C1) and (C2). The *out-map* r_o of an orientation o associates to each $c \in C$ the coordinate set of the edges outgoing from c. We continue with a characterization of representation maps of ample classes as out-maps of USOs, extending a similar result of Szabó and Welzl [31] for cubes. This characterization is "local-to-global", since (C1) and (C2) are conditions on the *stars* $\mathrm{St}(c)$ of all concepts $c \in C$ ($\mathrm{St}(c)$ is the set of all faces of the cubes containing c).

▶ **Theorem 15.** *For an ample class C and a map $r : C \to 2^U$, (i)-(iii) are equivalent:*
  (i) *r is a representation map;*
  (ii) *r is the out-map of a USO;*
 (iii) *$r(c) \in X(C)$ for any $c \in C$ and o_r satisfies (C2).*

Proof. The implication (i)⇒(ii) is a consequence of Theorem 14. Now, we prove (ii)⇒(i). Clearly, property (C1) implies that $r(c) \in X(C)$ for any $c \in C$, whence r is a map from C to $X(C)$. Let C be an ample class of smallest size admitting a non-representation map $r : C \to X(C)$ satisfying (C1) and (C2). Hence there exist $u_0, v_0 \in C$ such that $u_0|(r(u_0)\Delta r(v_0)) = v_0|(r(u_0)\Delta r(v_0))$, i.e., $(u_0 \Delta v_0) \cap (r(u_0)\Delta r(v_0)) = \varnothing$; (u_0, v_0) is called a *clashing pair*.

▷ **Claim 16.** If (u_0, v_0) is a clashing pair, then $C = C \cap B(u_0, v_0)$ and $r(u_0) = r(v_0) = \varnothing$.

Proof. Since $C \cap B(u_0, v_0)$ is ample and $(u_0\Delta v_0) \cap (r(u_0)\Delta r(v_0)) = \varnothing$, (u_0, v_0) is a clashing pair for $C \cap B(u_0, v_0)$ and the restriction r_B of r to $\mathrm{supp}(B(u_0, v_0))$. Since r_B and $C \cap B(u_0, v_0)$ satisfy (C1) and (C2), by minimality of C, $C = C \cap B(u_0, v_0)$. Moreover, if $r(u_0) \neq r(v_0)$, then there is $x \in r(u_0)\Delta r(v_0)$ and $x \in u_0\Delta v_0$, contradicting that (u_0, v_0) is a clashing pair.
  Suppose $r(u_0) \neq \varnothing$ and pick $x \in r(u_0) = r(v_0)$. Consider the carrier $N_x(C)$ of C^x. Note that $r(u_0) \subseteq \mathrm{supp}(N_x(C))$. Indeed, let $y \in r(u_0)$. By (C1), u_0 belongs to an $\{x, y\}$-square of C, whence $y \in \mathrm{supp}(N_x(C))$. Analogously, $r(v_0) \subseteq \mathrm{supp}(N_x(C))$, thus (u_0, v_0) is a clashing pair for $N_x(C)$ and the restriction of r to $N_x(C)$. $N_x(C)$ is ample as the product of C^x by an x-edge. By minimality of C, $C = N_x(C)$. Define $r^x : C^x \mapsto X(C^x)$ by
  ■ $r^x(c) = r(c) \setminus \{x\}$ if $x \in r(c)$,
  ■ $r^x(c) = r(cx) \setminus \{x\}$ otherwise.
Consequently, for an x-edge of C between c and cx, $r^x(c)$ is the label of the origin of this edge minus x; we call r^x the *x-out-map* of r. We assert that r^x satisfies (C1) and (C2). Condition (C1) is trivial because it holds for cubes of C. To establish condition (C2), suppose that there exists a cube B' of C^x and $u', v' \in B'$ such that $r^x(u') \cap \mathrm{supp}(B') = r^x(v') \cap \mathrm{supp}(B')$. The cube $B := B' \times \{x\}$ is included in C since B' is a cube of C^x. Then among the four pairs $(u', v'), (u', v'x), (u'x, v'), (u'x, v'x)$ of B one can select a pair (u, v) such that $r(u) = r^x(u') \cup \{x\} = r^x(v') \cup \{x\} = r(v)$, a contradiction with condition (C2) for C and r. This shows that r^x satisfies (C1) and (C2). Recall that $x \in r(u_0) = r(v_0)$, suppose wlog that $x \in v_0 \setminus u_0$, and let $u'_0 = u_0$ and $v'_0 = v_0 \setminus \{x\}$. Then $r^x(u'_0) = r(u_0)\setminus\{x\} = r(v_0)\setminus\{x\} = r^x(v'_0)$, and consequently (u'_0, v'_0) is a clashing pair for the restriction of r on $C^x \cap B(u'_0, v'_0)$. Since $C^x \cap B(u'_0, v'_0)$ is ample and smaller than C, this contradicts the minimality of C. ◁

▷ **Claim 17.** C is a cube minus a vertex.

Proof. By (C2), C is not a cube. If C is not a cube minus a vertex, since the complement $C^* = 2^U \setminus C$ is also ample (thus $G(C^*)$ is connected), $G(C^*)$ contains an x-edge ww' with $x \notin w$ and $x \in w'$. Consider C_x and define the map $r_x : C_x \mapsto X(C_x)$ by
  ■ $r_x(c) = r(c)$ if $c \in C$ and $x \notin r(c)$,
  ■ $r_x(c) = r(cx)$ otherwise.

Hence $r_x(c \setminus \{x\}) = r(c)$ for each $c \in C$ with $x \notin r(c)$. We call r_x the *x-in-map* of r; r_x satisfies (C1), because r satisfies (C1). Suppose that r_x violates (C2). Then there exists a cube B' of C_x and $u', v' \in B'$ such that $(u' \Delta v') \cap (r_x(u') \Delta r_x(v')) = \varnothing$. Let $u \in \{u', u'x\}$ such that $r(u) = r(u')$ and let $v \in \{v', v'x\}$ such that $r(v) = r(v')$. The restriction C'_x of the ample class $C' := C \cap B(u, v)$ is the cube B'. Since $w, w' \notin C$ and ww' is an x-edge, $w \notin C'_x$. Thus there exists $y \in \mathrm{supp}(C)$ such that C' and the edge ww' of C^* belong to different y-half-spaces $C^- = \{c \in C : y \notin c\}$ and $C^+ = \{c \in C : y \in c\}$ of the cube 2^U. Since $y \in \mathrm{supp}(C)$, the half-space containing ww' also contains a concept of C. Hence, C' is a proper ample subset of C. Since $u \subseteq u' \cup \{x\}$, $v \subseteq v' \cup \{x\}$, $x \notin r(u) = r_x(u')$, $x \notin r(v) = r_x(v')$, we deduce that $u \cap (r(u) \Delta r(v)) = u' \cap (r_x(u') \Delta r_x(v'))$ and $v' \cap (r_x(u') \Delta r_x(v')) = v \cap (r(u) \Delta r(v))$. Since $(u' \Delta v') \cap (r_x(u') \Delta r_x(v')) = \varnothing$, (u, v) is a clashing pair for the restriction of r on C', contrary to the minimality of C. Hence by minimality of C, r_x is a representation map for C_x.

Consider a clashing pair (u_0, v_0) for C and r, and let $u'_0 = u_0 \setminus \{x\}$ and $v'_0 = v_0 \setminus \{x\}$. Observe that $r_x(u'_0) = r(u_0) = r(v_0) = r_x(v'_0) = \varnothing$. Since r_x is a representation map for C_x, necessarily $u'_0 = v'_0$. Consequently, $u_0 \Delta v_0 = \{x\}$, i.e., $u_0 v_0$ is an x-edge of $G(C)$. This is impossible since C satisfies (C2). Therefore, C is necessarily a cube minus a vertex.   ◁

Now, we complete the proof of the implication (ii)⇒(i). By Claim 16, $r(u_0) = r(v_0) = \varnothing$. By condition (C1), $r(c) \neq U$ for any $c \in C$. Thus there exists a set $s \in X(C) = 2^U \setminus \{U, \varnothing\}$ such that $s \neq r(c)$ for any $c \in C$. Every s-cube B of C contains a source $p(B)$ for o_{r_B} (i.e., $s \subseteq r(p(B))$). For each s-cube B of C, let $t(B) = r(p(B)) \setminus s$. Notice that $\varnothing \subsetneq t(B) \subsetneq U \setminus s$ since $s \subsetneq r(p(B)) \subsetneq U$. Consequently, there are $2^{|U|-|s|} - 2$ choices for $t(B)$ and since C is a cube minus one vertex by Claim 17, there are $2^{|U|-|s|} - 1$ s-cubes in C. Consequently, there exist two s-cubes B, B' such that $t(B) = t(B')$. Thus $\varnothing \subsetneq s \subsetneq r(p(B)) = r(p(B'))$ and $(p(B), p(B'))$ is a clashing pair for C and r, contradicting Claim 16.

The implication (ii)⇒(iii) is trivial. To prove (iii)⇒(ii), we show by induction on $|U|$ that a map $r : C \to X(C)$ satisfying (C2) also satisfies (C1). For any $x \in U$, let r^x denote the x-out-map defined in Claim 16. Recall that if cc' is an x-edge directed from c to c', then $x \in r(c)$ and r^x maps $c^x = (c')^x \in C^x$ to $r(c) \setminus \{x\} \in X(C^x)$. Thus r^x maps C^x to $X(C^x)$. Moreover, each cube B^x of C^x is contained in a unique cube B of C such that $\mathrm{supp}(B) = \mathrm{supp}(B^x) \cup \{x\}$. If there exist $c_1^x, c_2^x \in B^x$ such that $r^x(c_1^x) = r(c_2^x)$, then there exist $c_1, c_2 \in B$ such that $r(c_1) = r^x(c_1^x) \cup \{x\} = r^x(c_2^x) \cup \{x\} = r(c_2)$, contradicting (C2). Consequently, o_{r_x} satisfies (C2). By induction hypothesis, o_{r_x} satisfies (C1) for any $x \in U$.

For any concept $c \in C$, pick $x \in r(c)$. Since r^x satisfies (C1), c^x belongs to a σ'-cube in C^x with $\sigma' = r^x(c^x) = r(c) \setminus \{x\}$. This implies that c belongs to a σ-cube in C with $\sigma = \sigma' \cup \{x\} = r(c)$. Thus o_r satisfies (C1), concluding the proof of Theorem 15.   ◀

We conclude with some remarks regarding Theorems 14 and 15. First, corner peelings correspond *exactly* to acyclic USOs. Second, given a representation map for C one can derive representations maps for intersections of C with cubes, reductions C^Y, and restrictions C_Y. Third, there exist a bijection $r' : C \to X(C)$ satisfying (C1) and an injection $r'' : C \to 2^U$ satisfying (C2). Nevertheless, we were not able to find a map satisfying (C1) and (C2). It is surprising that, while each d-cube has at least $d^{\Omega(2^d)}$ USOs [17], it is so difficult to find a single USO for ample classes. One can try to find such maps by extending the approach for maximum classes: given ample classes C and D with $D \subset C$, a representation map r for C is called *D-entering* if all edges cd with $c \in C \setminus D$ and $d \in D$ are directed by o_r from c to d. The representation map defined in Proposition 11 is D-entering. Given $x \in \mathrm{dom}(C)$, suppose that r_x is a C^x-entering representation map for C_x. We can extend the orientation o_{r_x} to an orientation o of $G(C)$ as follows. Each x-edge cc' of $G(C)$ is directed arbitrarily, while each

other edge cc' is directed as the edge $c_x c'_x$ is directed by o_{r_x}. Since o_{r_x} satisfies (C1), (C2) and r_x is C^x-entering, o also satisfies (C1), (C2), thus the map r_o is a representation map for C. So, ample classes would admit representation maps, if *for any ample classes $D \subseteq C$, any representation map r' of D extends to a D-entering representation map r of C.*

References

1   Richard P. Anstee, Lajos Rónyai, and Attila Sali. Shattering News. *Graphs Combin.*, 18(1):59–73, 2002. `doi:10.1007/s003730200003`.

2   Hans-Jürgen Bandelt and Victor Chepoi. Metric Graph Theory and Geometry: a Survey. In J. E. Goodman, J. Pach, and R. Pollack, editors, *Surveys on Discrete and Computational Geometry: Twenty Years Later*, volume 453 of *Contemp. Math.*, pages 49–86. Amer. Math. Soc., Providence, RI, 2008. `doi:10.1090/conm/453/08795`.

3   Hans-Jürgen Bandelt, Victor Chepoi, Andreas W. M. Dress, and Jack H. Koolen. Combinatorics of Lopsided Sets. *European J. Combin.*, 27(5):669–689, 2006. `doi:10.1016/j.ejc.2005.03.001`.

4   Béla Bollobás and Andrew J. Radcliffe. Defect Sauer Results. *J. Combin. Theory Ser. A*, 72(2):189–208, 1995. `doi:10.1016/0097-3165(95)90060-8`.

5   Jérémie Chalopin, Victor Chepoi, Shay Moran, and Manfred K. Warmuth. Unlabeled Sample Compression Schemes and Corner Peelings for Ample and Maximum Classes. *arXiv preprint*, 1812.02099, 2018. `arXiv:1812.02099`.

6   Thorsten Doliwa, Gaojian Fan, Hans Ulrich Simon, and Sandra Zilles. Recursive Teaching Dimension, VC-dimension and Sample Compression. *J. Mach. Learn. Res.*, 15(1):3107–3131, 2014. URL: `http://jmlr.org/papers/v15/doliwa14a.html`.

7   Andreas W. M. Dress. Towards a Theory of Holistic Clustering. In *Mathematical Hierarchies and Biology*, volume 37 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 271–290. DIMACS, Amer. Math. Soc., 1996. `doi:10.1090/dimacs/037/19`.

8   Paul H. Edelman and Robert E. Jamison. The Theory of Convex Geometries. *Geom. Dedicata*, 19(3):247–270, 1985. `doi:10.1007/BF00149365`.

9   Sally Floyd. *On Space Bounded Learning and the Vapnik-Chervonenkis Dimension*. PhD thesis, International Computer Science Institut, Berkeley, CA, 1989. URL: `https://www.icsi.berkeley.edu/icsi/node/2289`.

10   Sally Floyd and Manfred K. Warmuth. Sample Compression, Learnability, and the Vapnik-Chervonenkis Dimension. *Mach. Learn.*, 21(3):269–304, 1995. `doi:10.1007/BF00993593`.

11   Bernd Gärtner and Emo Welzl. Vapnik-Chervonenkis dimension and (pseudo-)hyperplane arrangements. *Discrete Comput. Geom.*, 12(4):399–432, 1994. `doi:10.1007/BF02574389`.

12   Mikhaïl Gromov. Hyperbolic Groups. In S. M. Gersten, editor, *Essays in Group Theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987. `doi:10.1007/978-1-4613-9586-7_3`.

13   H. Tracy Hall. *Counterexamples in Discrete Geometry*. PhD thesis, University of California, 2004.

14   Dima Kuzmin and Manfred K. Warmuth. Unlabeled Compression Schemes for Maximum Classes. *J. Mach. Learn. Res.*, 8:2047–2081, 2007. URL: `http://www.jmlr.org/papers/v8/kuzmin07a.html`.

15   James F. Lawrence. Lopsided Sets and Orthant-intersection of Convex Sets. *Pacific J. Math.*, 104(1):155–173, 1983. `doi:10.2140/pjm.1983.104.155`.

16   Nick Littlestone and Manfred K. Warmuth. Relating Data Compression and Learnability. Technical report, Department of Computer and Information Sciences, Santa Cruz, CA, 1986.

17   Jiří Matoušek. The Number Of Unique-Sink Orientations of the Hypercube. *Combinatorica*, 26(1):91–99, 2006. `doi:10.1007/s00493-006-0007-0`.

18   Tamás Mészáros and Lajos Rónyai. Shattering-Extremal Set Systems of VC Dimension at most 2. *Electron. J. Combin.*, 21(4):P4.30, 2014. URL: `http://www.combinatorics.org/ojs/index.php/eljc/article/view/v21i4p30`.

19   Shay Moran. Shattering-extremal Systems. *arXiv preprint*, 1211.2980, 2012. `arXiv:1211.2980`.

20   Shay Moran and Manfred K. Warmuth. Labeled Compression Schemes for Extremal Classes. In *ALT 2016*, volume 9925 of *Lecture Notes in Comput. Sci.*, pages 34–49. Springer, 2016. `doi:10.1007/978-3-319-46379-7_3`.

21   Shay Moran and Amir Yehudayoff. Sample Compression Schemes for VC Classes. *J. ACM*, 63(3):21:1–21:10, 2016. `doi:10.1145/2890490`.

22   Mogens Nielsen, Gordon D. Plotkin, and Glynn Winskel. Petri nets, event structures and domains, part I. *Theoret. Comput. Sci.*, 13(1):85–108, 1981. `doi:10.1016/0304-3975(81)90112-2`.

23   Alain Pajor. *Sous-espaces ℓ_1^n des Espaces de Banach.* Travaux en Cours. Hermann, Paris, 1985.

24   Dömötör Pálvölgyi and Gábor Tardos. Unlabeled Compression Schemes Exceeding the VC-dimension. *arXiv preprint*, 1811.12471, 2018. `arXiv:1811.12471`.

25   Benjamin I. P. Rubinstein and J. Hyam Rubinstein. A Geometric Approach to Sample Compression. *J. Mach. Learn. Res.*, 13:1221–1261, 2012. URL: `http://www.jmlr.org/papers/v13/rubinstein12a.html`.

26   Michah Sageev. CAT(0) cube complexes and groups. In M. Bestvina, M. Sageev, and K. Vogtmann, editors, *Geometric Group Theory*, volume 21 of *IAS/Park City Math. Ser.*, pages 6–53. Amer. Math. Soc., Inst. Adv. Study, 2012. `doi:10.1090/pcms/021/02`.

27   Rahim Samei, Boting Yang, and Sandra Zilles. Generalizing Labeled and Unlabeled Sample Compression to Multi-label Concept Classes. In *ALT 2014*, volume 8776 of *Lecture Notes in Comput. Sci.*, pages 275–290. Springer, 2014. `doi:10.1007/978-3-319-11662-4_20`.

28   Norbert Sauer. On the Density of Families of Sets. *J. Combin. Theory Ser. A*, 13(1):145–147, 1972. `doi:10.1016/0097-3165(72)90019-2`.

29   Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms.* Cambridge Univ. Press, 2014. `doi:10.1017/CBO9781107298019`.

30   Saharon Shelah. A Combinatorial Problem, Stability and Order for Models and Theories in Infinitary Languages. *Pacific J. Math.*, 41(1):247–261, 1972. `doi:10.2140/pjm.1972.41.247`.

31   Tibor Szabó and Emo Welzl. Unique Sink Orientations of Cubes. In *FOCS 2001*, pages 547–555. IEEE Computer Society, 2001. `doi:10.1109/SFCS.2001.959931`.

32   Vladimir N. Vapnik and Alexey Y. Chervonenkis. On the Uniform Convergence of Relative Frequencies of Events to their Probabilities. *Theory Probab. Appl.*, 16(2):264–280, 1971. `doi:10.1137/1116025`.

33   Manfred K. Warmuth. Compressing to VC Dimension Many Points. In *COLT/Kernel 2003*, volume 2777 of *Lecture Notes in Comput. Sci.*, pages 743–744. Springer, 2003. `doi:10.1007/978-3-540-45167-9_60`.

34   Emo Welzl. Complete Range Spaces, 1987. Unpublished notes.

35   Douglas H. Wiedemann. *Hamming Geometry.* PhD thesis, University of Waterloo, 1986. re-typeset July, 2006.

36   Avi Wigderson. *Mathematics and Computation.* Princeton Univ. Press, 2019.

37   Glynn Winskel. *Events in Computation.* PhD thesis, Edinburgh University, 1980.

38   Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Grad. Texts in Math.* Springer–Verlag, New York, 1995. `doi:10.1007/978-1-4613-8431-1`.