Labeled sample compression schemes for complexes of oriented matroids

Victor Chepoi[1], Kolja Knauer[1,2], and Manon Philibert[1]

[1]LIS, Aix-Marseille Université, CNRS, and Université de Toulon
Faculté des Sciences de Luminy, F-13288 Marseille Cedex 9, France
{victor.chepoi,kolja.knauer,manon.philibert}@lis-lab.fr

[2]Departament de Matemàtiques i Informàtica, Universitat de Barcelona (UB), Barcelona, Spain

Abstract We show that the topes of a complex of oriented matroids (abbreviated COM) of VC-dimension d admit a proper labeled sample compression scheme of size d. This considerably extends results of Moran and Warmuth on ample classes, of Ben-David and Litman on affine arrangements of hyperplanes, and of the authors on complexes of uniform oriented matroids, and is a step towards the sample compression conjecture – one of the oldest open problems in computational learning theory. On the one hand, our approach exploits the rich combinatorial cell structure of COMs via oriented matroid theory. On the other hand, viewing tope graphs of COMs as partial cubes creates a fruitful link to metric graph theory.

1. Introduction

1.1. **General setting.** Littlestone and Warmuth [51] introduced sample compression schemes as an abstraction of the underlying structure of learning algorithms. Roughly, the aim of a sample compression scheme is to compress samples of a *concept class* (i.e., of a set system) $\mathcal{C}$ as much as possible, such that data coherent with the original samples can be reconstructed from the compressed data. There are two types of sample compression schemes: labeled, see [35, 51] and unlabeled, see [7, 34, 49]. A labeled compression scheme of size k compresses every sample of $\mathcal{C}$ to a labeled subsample of size at most k and an unlabeled compression scheme of size k compresses every sample of $\mathcal{C}$ to a subset of size at most k of the domain of the sample (see the end of the introduction for precise definitions). The Vapnik-Chervonenkis dimension (*VC-dimension*) of a set system, was introduced by [69] as a complexity measure of set systems. VC-dimension is central in PAC-learning and plays an important role in combinatorics, algorithmics, discrete geometry, and combinatorial optimization. In particular, it coincides with the rank in the theory of (complexes of) oriented matroids. Furthermore, within machine learning and closely tied to the topic of this paper, the *sample compression conjecture* of [35] and [51] states that *any set system of VC-dimension d has a labeled sample compression scheme of size $O(d)$.* This question remains one of the oldest open problems in computational learning theory.

1.2. **Related work.** The best-known general upper bound is due to Moran and Yehudayoff [58] and shows that there exist labeled compression schemes of size $O(2^d)$ for any set system of VC-dimension d. The labeled compression scheme of [58] is not proper (i.e., does not necessarily return a set from the input set system) and it is even open if there exist proper labeled sample compression schemes which compress samples with support larger than d to subsamples with strictly smaller support [56]. From below, Floyd and Warmuth [35] showed that there are classes of VC-dimension d admitting no labeled compression scheme of size less than d and that no concept class of VC-dimension d admits a labeled compression scheme of size at most $\frac{d}{5}$. Pálvölgyi and Tardos [64] exhibited a concept class of VC-dimension 2 with no unlabeled compression scheme of size 2. However, no similar results are known for labeled sample compression schemes. Prior to [64], it was shown in [61] that the concept class of positive halfspaces in $\mathbb{R}^2$ (which has VC-dimension 2) does not admit proper unlabeled sample compression schemes of size 2.

For more structured concept classes better upper bounds are known. Ben-David and Litman [7] proved a compactness lemma, which reduces the existence of labeled or unlabeled compression schemes for arbitrary concept classes to finite concept classes. They also obtained unlabeled compression schemes for regions in arrangements of affine hyperplanes (which correspond to realizable affine oriented matroids in our language). Finally, they obtained sample compression schemes for concept classes by embedding them into concept classes for which such schemes were known. Helmbold, Sloan, and Warmuth [43] constructed unlabeled compression schemes of size d for intersection-closed concept classes of VC-dimension d. They compress each sample to a minimal generating set and show that the size of this set is upper bounded by the VC-dimension. An important class for which positive results are available is given by ample set systems [3, 27] (originally introduced as lopsided sets by Lawrence [50]). They capture an important variety of combinatorial objects, e.g., (conditional) antimatroids, see [29], diagrams of (upper locally) distributive lattices, median graphs or CAT(0) cube complexes, see [3] and were rediscovered in various disguises, e.g. by [10] as *extremal for (reverse) Sauer* and by [59] as *shattering-extremal* [59]. Moran and Warmuth [57] provide labeled sample compression schemes of size d for ample set systems of VC-dimension d. For maximum concept classes (a subclass of ample set systems) unlabeled sample compression schemes of size d have been designed by Chalopin et al. [11]. They also characterized unlabeled compression schemes for ample classes via the existence of *unique sink orientations* of their graphs. However, the existence of such orientations remains open.

1.3. **OMs and COMs.** A structure somewhat opposed to ample classes are Oriented Matroids (OMs), see the book of Björner et al. [8]. Co-invented by Bland and Las Vergnas [9] and Folkman and Lawrence [36], and further investigated by Edmonds and Mandel [30] and many other authors, oriented matroids represent a unified combinatorial theory of orientations of ordinary matroids, which simultaneously captures the basic properties of sign vectors representing the regions in a hyperplane arrangement in $\mathbb{R}^d$ and of sign vectors of the circuits in a directed graph. OMs provide a framework for the analysis of combinatorial properties of geometric configurations occurring in discrete geometry and in machine learning. Point and vector configurations, order types, hyperplane and pseudo-line arrangements, convex polytopes, directed graphs, and linear programming find a common generalization in this language. The Topological Representation Theorem of [36] connects the theory of OMs on a deep level to arrangements of pseudohyperplanes and distinguishes it from the theory of ordinary matroids.

Complexes of Oriented Matroids (COMs) were introduced by Bandelt, Chepoi, and Knauer [4] as a natural common generalization of ample classes and OMs. Ample classes are exactly the COMs with cubical cells, while OMs are the COMs with a single cell. In general COMs, the cells are OMs and the resulting cell complex is contractible. In the realizable setting, a COM corresponds to the intersection pattern of a hyperplane arrangement with an open convex set, see Figure 1. Examples of COMs neither contained in the class of OMs nor in ample classes include linear extensions of a poset or acyclic orientations of mixed graphs, see [4], CAT(0) Coxeter complexes of [40], hypercellular and Pasch graphs, see [17], and Affine Oriented Matroids through [6] and [23]. Note that none of the listed examples is contained in the classes of OMs or ample classes. Apart from the above, COMs already lead to new results and questions in various areas such as combinatorial semigroup theory by [54], algebraic combinatorics in relation to the Varchenko determinant by [44, 45], neural codes [46], poset cones, see [26], as well as sweeping sequences, see [63]. In particular, relations to COMs have already been established within sample compression, by [18, 19, 53] and [11]. A central feature of COMs is that they can be studied via their tope graphs, see Figure 1. Indeed, the characterization of their tope graphs by [47] establishes an embedding of the theory of COMs into metric graph theory, with theoretical and algorithmic implications. Namely, tope graphs of COMs form a subclass of the ubiquitous metric graph class of partial cubes, i.e., isometric subgraphs of hypercubes, with applications ranging from interconnection networks [38] and media theory [33], to chemical graph theory [32]. On the other hand, tope graphs of COMs can be recognized in polynomial time [31, 47]. The graph theoretic view has been used in several recent publications, see [16, 48, 52] and is essential to our work.

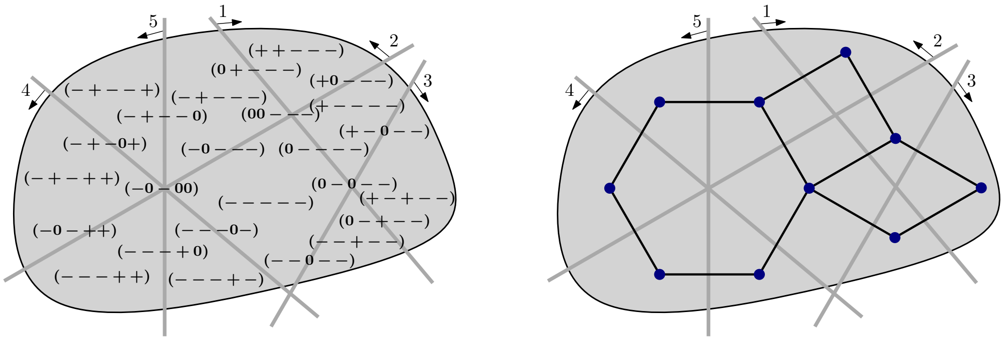

FIGURE 1. A realizable COM and its tope graph.

1.4. **Labeled sample compression schemes.** As we explain later, COMs can be defined as sets of sign vectors, which is another unifying feature for OMs and ample classes. This turns out to be beneficial for the present paper, since the language of sign vectors is perfectly suited for defining sample compression schemes formally. The following formulation is due to [12], for classical formulations, see [51, 57, 58]. Let U be a finite set, called the *universe* and $\mathcal{C}$ be a family of subsets of U, called a *concept class* and whose elements are called *concepts*. We view $\mathcal{C}$ as a set of $\{-1,+1\}$-vectors, i.e., $\mathcal{C} \subseteq \{-1,+1\}^U$. We also consider sets of $\{-1,0,+1\}$-vectors, i.e., subsets of $\{\pm 1, 0\}^U$ endowed with the product order $\leq$ between sign vectors relative to the ordering $0 \leq -1, +1$. The sign vectors of the set $\mathbf{Samp}(\mathcal{C}) = \bigcup_{C \in \mathcal{C}} \{S \in \{\pm 1, 0\}^U : S \leq C\}$ are *realizable samples* for $\mathcal{C}$.

Definition 1 (Labeled sample compression schemes)**.** A *labeled sample compression scheme* of size k for a concept class $\mathcal{C} \subseteq \{-1,+1\}^U$ is a pair (α, β) of mappings, where $\alpha : \mathbf{Samp}(\mathcal{C}) \to \{\pm 1, 0\}^U$ is called the *compression function* and $\beta : \mathrm{Im}(\alpha) \to \{-1,+1\}^U$ the *reconstruction function* such that for any realizable sample $S \in \mathbf{Samp}(\mathcal{C})$, it holds $\alpha(S) \leq S \leq \beta(\alpha(S))$ and $|\underline{\alpha}(S)| \leq k$, where $\underline{\alpha}(S)$ is the support of the sign vector $\alpha(S)$, i.e., the non-zero entries of $\alpha(S)$. A labeled sample compression scheme is *proper* if $\beta(\alpha(S)) \in \mathcal{C}$ for all $S \in \mathbf{Samp}(\mathcal{C})$.

The condition $S \leq \beta(\alpha(S))$ means that the restriction of $\beta(\alpha(S))$ on the support of S coincides with the input sample S. In particular, if S is a concept of $\mathcal{C}$, then $\beta(\alpha(S)) = S$, i.e., the reconstructor must reconstruct the input concept. Notice that the labeled compression schemes of size $O(2^d)$ of [58] are not proper (i.e., $\beta(\alpha(S))$ is not necessarily a concept of $\mathcal{C}$) and they use additional information. The compression schemes developed in [12] for balls in graphs are proper but also use additional information. The *unlabeled sample compression schemes* [49] (which are not the subject of this paper) are defined analogously, with the difference that in the unlabeled case $\alpha(S)$ is a subset of size at most k of the support of S.

The definition of labeled compression scheme implies that if $\mathcal{C}' \subseteq \mathcal{C}$ and (α, β) is a labeled sample compression scheme for $\mathcal{C}$, then (α, β) is a labeled sample compression scheme for $\mathcal{C}'$. However, (α, β) is in general not proper for $\mathcal{C}'$. Still, this yields an approach (suggested in [67] and implicit in [35]) to obtain improper schemes. For instance, using the result of [57] that ample classes of VC-dimension d admit labeled sample compression schemes of size d, one can try to extend a given set system to an ample class without increasing the VC-dimension too much and then apply their result. In [18] it is shown that partial cubes of VC-dimension 2 can be extended to ample classes of VC-dimension 2. Furthermore, in [19] it is shown that OMs and complexes of uniform oriented matroids (CUOMs) can be extended to ample classes without increasing the VC-dimension.

Thus, in these classes there exist improper labeled sample compression schemes whose size is the VC-dimension. On the other hand, there exist partial cubes of VC-dimension 3 that cannot be extended to ample classes of VC-dimension 3, see [19], as well as set systems of VC-dimension 2, that cannot be extended to partial cubes of VC-dimension 2, see [18]. In [19] it is conjectured that every COM of VC-dimension d can be extended to an ample class of VC-dimension d. This would yield improper labeled sample compression schemes for COMs of size d.

1.5. **Our result.** In this paper, we follow a different strategy to give (stronger) proper labeled sample compression schemes of size d for general COMs of VC-dimension d, see Theorem 3. More precisely, we show that the set systems defined by the topes of COMs satisfy the strong form of the sample compression conjecture, i.e., COMs of VC-dimension d admit *proper labeled sample compression schemes of size d.*

Our work substantially extends the result of [57] for ample concept classes, the result of [7] for concept classes arising from arrangements of affine hyperplanes (i.e., realizable Affine Oriented Matroids), and our results [19] for OMs and CUOMs. Many classes of COMs are only covered by this new result. For example, the classes of COMs mentioned in Subsection 1.3 are neither ample, nor affine, nor uniform. Some of these examples are realizable and can be embedded into realizable Affine Oriented Matroid to which one can apply the result of [7]. However, this will lead only improper compression schemes. One important class of COMs, which is neither realizable, nor ample, nor affine, nor uniform, is the class of non-realizable OMs. By the *Topological Representation Theorem of Oriented Matroids* of Folkman and Lawrence [36], the topes of OMs can be characterized as the inclusion maximal cells of an arrangement of pseudohyperplanes. An OM is non-realizable if it it represented by a non-stretchable arrangement, i.e., an arrangement whose pseudohyperplanes cannot be replaced by linear hyperplanes.

To illustrate the representation by pseudohyperplanes, in Figure 2 we give an example of an arrangement U of pseudolines in $\mathbb{R}^2$ and its graph of regions, i.e., the tope graph of the resulting COM. While this example is stretchable, there are many non-stretchable arrangements. Indeed, most OMs are non-realizable [8, Theorems 7.4.2 and 8.7.5]. Deciding stretchability of a pseudoline arrangement and more generally realizability of an OM is a complete problem of the existential theory of the reals, hence in particular NP-hard, see [68]. By a result of Edmonds and Mandel [30], all arrangements of pseudohyperplanes can be considered piecewise-linear.

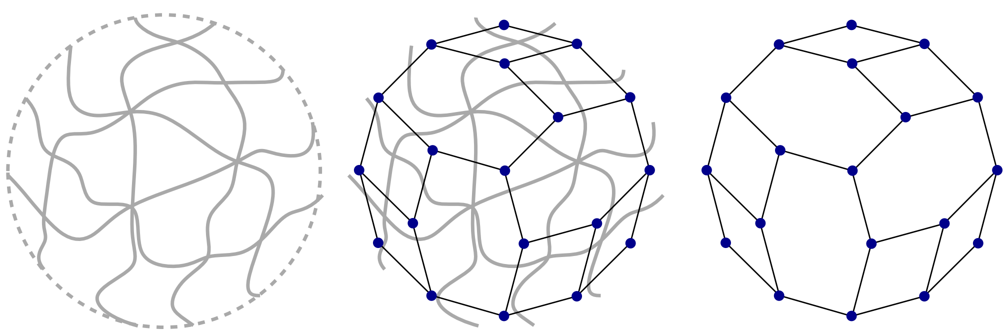

FIGURE 2. A pseudoline arrangement U and its region graph.

1.6. **Pseudohyperplane arrangements and Machine Learning.** Pseudohyperplane arrangements have already arisen in the context of sample compression schemes and VC-dimension in [37, 55, 65, 66] in the treatment of maximum and ample classes. More recently, particular piecewise-linear pseudohyperplane arrangements and their regions occurred in the study of deep feedforward neural networks with ReLU activations [24, 39, 41, 42, 60]. In this theory they appear under the names "arrangements of bent hyperplanes" and "activation regions", respectively. Recall that a (trained) feedforward neural network used to answer Yes/No (i.e., $\{-1,+1\}$) classification problems is a particular type of function $F:\mathbb{R}^d\to\mathbb{R}$. The inputs to F are data feature vectors and the outputs are used to answer the binary classification problem by partitioning the input space $\mathbb{R}^d$ into activation regions.

Next, we closely follow [39] and [42]. A ReLU function $\mathrm{ReLU}:\mathbb{R}\to\mathbb{R}$ is defined by $\mathrm{ReLU}(x)=\max\{0,x\}$. ReLU is among the most popular activation functions for deep neural networks. Let $\sigma:\mathbb{R}^d\to\mathbb{R}^d$ denote the function that applies ReLU to each coordinate. Let $n_0,\dots,n_k,n_{k+1}=1$ be a sequence of natural numbers and let $A_i:\mathbb{R}^{n_{i-1}}\to\mathbb{R}^{n_i}, i=1,\dots,k+1$ be (parametrized) affine maps. A ReLU (Rectified Linear Unit) network $\mathcal{N}$ of architecture $(n_0,\dots,n_k)$, depth $k+1$, and $n:=\sum_{i=0}^m n_i$ neurons is a neural network in which the map F is defined as the composition of the layer maps $F_1=\sigma\circ A_1,\dots,F_k=\sigma\circ A_k, F_{k+1}=A_{k+1}$. An activation pattern for $\mathcal{N}$ is an assignment of a $\{-1,+1\}$-sign to each neuron. Given a vector θ of trainable parameters, the activation pattern of the neurons defines a partition of the input space $\mathbb{R}^d$ into activation regions. The activation regions can be viewed as the regions defined by the arrangement of bent hyperplanes associated to layers; for the precise definition see [39, Section 6] and [42]. Activation regions are convex polyhedra [42] and one of important questions in the complexity analysis of deep ReLU networks is counting the number of such activation regions [41, 42, 60]. Notice that the arrangements of bent hyperplanes may not be arrangements of pseudohyperplanes in the classical sense [8] because two bent hyperplanes may not intersect transversally. Transversality of arrangements of bent hyperplanes was investigated in depth in the recent paper [39]. It will be interesting to further investigate how sample compression schemes can be useful in the setting of deep ReLU networks.

2. Preliminaries

2.1. **OMs and COMs.** We recall the basic theory OMs and COMs from [8] and [4], respectively. Let U be a set of size m and let $\mathcal{L}$ be a *system of sign vectors*, i.e., maps from U to $\{-1,0,+1\}$. The elements of $\mathcal{L}$ are referred to as *covectors* and denoted by capital letters X,Y,Z. For $X\in\mathcal{L}$, the subset $\underline{X}=\{e\in U: X_e\neq 0\}$ is the *support* of X and its complement $X^0=U\setminus\underline{X}=\{e\in U: X_e=0\}$ is the *zero set* of X. For $X,Y\in\mathcal{L}$, $\mathrm{Sep}(X,Y)=\{e\in U: X_eY_e=-1\}$ is the *separator* of X and Y. The *composition* of X and Y is the sign vector $X\circ Y$, where for all $e\in U$, $(X\circ Y)_e=X_e$ if $X_e\neq 0$ and $(X\circ Y)_e=Y_e$ if $X_e=0$. Let $\leq$ be the product ordering on $\{\pm1,0\}^U$ relative to the ordering $0\leq -1,+1$. A system of sign vectors $(U,\mathcal{L})$ is *simple* if for each $e\in U$, $\{X_e: X\in\mathcal{L}\}=\{-1,0,+1\}$ and for all $e\neq f$ there exist $X,Y\in\mathcal{L}$ with $\{X_eX_f,Y_eY_f\}=\{+1,-1\}$. In this paper, we consider only simple systems of sign vectors.

Definition 2 (OMs)**.** An *oriented matroid* (OM) is a system of sign vectors $\mathcal{M}=(U,\mathcal{L})$ satisfying

- **(Z)** the zero sign vector $\mathbf{0}$ belongs to $\mathcal{L}$.
- **(C)** (Composition) $X\circ Y\in\mathcal{L}$ for all $X,Y\in\mathcal{L}$.
- **(SE)** (Strong elimination) for each pair $X,Y\in\mathcal{L}$ and for each $e\in\mathrm{Sep}(X,Y)$, there exists $Z\in\mathcal{L}$ such that $Z_e=0$ and $Z_f=(X\circ Y)_f$ for all $f\in U\setminus\mathrm{Sep}(X,Y)$.
- **(Sym)** (Symmetry) $-\mathcal{L}=\{-X: X\in\mathcal{L}\}=\mathcal{L}$, that is, $\mathcal{L}$ is closed under sign reversal.

Notice that the axiom (Z) is implied by the three other axioms. The poset $(\mathcal{L},\leq)$ of an OM $\mathcal{M}$ with an artificial global maximum $\hat{1}$ forms the (graded) *big face lattice* $\mathcal{F}_{\mathrm{big}}(\mathcal{M})$. The length of maximal chains of $\mathcal{F}_{\mathrm{big}}(\mathcal{M})$ minus 1 is the *rank* of $\mathcal{L}$ and denoted $\mathrm{rank}(\mathcal{M})$. The rank of the underlying matroid $\underline{\mathcal{M}}$ equals $\mathrm{rank}(\mathcal{M})$ [8, Thm 4.1.14]. The *topes* $\mathcal{T}$ of $\mathcal{M}$ are the co-atoms of $\mathcal{F}_{\mathrm{big}}(\mathcal{M})$. By simplicity the topes are $\{-1,+1\}$-vectors and $\mathcal{T}$ can be seen as a family of subsets of U. For each $T\in\mathcal{T}$, an element $e\in U$ belongs to the corresponding set if and only if $T_e=+1$. The *tope graph* $G(\mathcal{M})$ of an OM $\mathcal{M}$ is the 1-inclusion graph of the set $\mathcal{T}$ of topes of $\mathcal{L}$, i.e., the subgraph of the hypercube $\mathrm{Q}(U)$ induced by the vertices corresponding to $\mathcal{T}$, see Figure 1.

In *realizable OMs* (i.e., OMs arising from central hyperplane arrangements of $\mathbb{R}^d$), $X\leq Y$ for two covectors X,Y if and only if the (open) cell corresponding to X is contained in the cell corresponding to Y. Consequently, the topes of realizable OMs are the covectors of the inclusion maximal (open) cells (which all have dimension d), called *regions*. Therefore, the tope graph of a realizable OM can be viewed as the adjacency graph of regions: the vertices of this graph are the regions of a hyperplane arrangement and two regions are adjacent in this graph if they are separated by a unique hyperplane of the arrangement. The *Topological Representation Theorem of Oriented Matroids* of [36], generalizes this correspondence to all OMs: tope graphs of OMs can be characterized as the adjacency graphs of maximal (open) cells of pseudohyperplane arrangements in $\mathbb{R}^d$ [8], where d is the rank of the OM. More precisely, two topes are adjacent if and only if the corresponding regions are separated by a unique pseudohyperplane, see Figure 1. It is also well-known (see for example [8]) that $\mathcal{L}$ can be recovered from its tope graph $G(\mathcal{L})$ (up to isomorphism). Therefore, *we can define all terms in the language of tope graphs.*

Another important axiomatization of OMs is in terms of *cocircuits* of $\mathcal{L}$. These are the atoms of $\mathcal{F}_{\text{big}}(\mathcal{L})$. Their collection is denoted by $\mathcal{C}^*$ and axiomatized as follows: a system of sign vectors $(U, \mathcal{C}^*)$ is an *oriented matroid* (OM) if $\mathcal{C}^*$ satisfies (Sym) and the two axioms:

- **(Inc)** (Incomparability) $\underline{X} \subseteq \underline{Y}$ implies $X = \pm Y$ for all $X, Y \in \mathcal{C}^*$.
- **(E)** (Elimination) for each pair $X, Y \in \mathcal{C}^*$ with $X \neq -Y$ and for each $e \in \text{Sep}(X, Y)$, there exists $Z \in \mathcal{C}^*$ such that $Z_e = 0$ and $Z_f \in \{0, X_f, Y_f\}$ for all $f \in U$.

The set $\mathcal{L}$ of covectors can be derived from $\mathcal{C}^*$ by taking the closure of $\mathcal{C}^*$ under composition.

COMs are defined by replacing the global axiom (Sym) with a weaker local axiom:

Definition 3 (COMs)**.** A *complex of oriented matroids* (COM) is a system of sign vectors $\mathcal{M} = (U, \mathcal{L})$ satisfying (SE) and the following axiom:

- **(FS)** (Face symmetry) $X \circ -Y \in \mathcal{L}$ for all $X, Y \in \mathcal{L}$.

One can see that OMs are exactly the COMs containing the zero vector $\mathbf{0}$ (axiom (Z)), see [4]. The twist between (Sym) and (FS) allows to keep on using the same concepts, such as topes, tope graphs, the sign-order and the big face (semi)lattice in a completely analogous way. On the other hand, it leads to a combinatorial and geometric structure that is built from OMs as cells but is much richer than OMs. Let $\mathcal{M} = (U, \mathcal{L})$ be a COM and $X \in \mathcal{L}$ a covector. The *face* of X is $\text{F}(X) := \{Y \in \mathcal{L} : X \leq Y\}$ (see [4,8]) and $\text{Q}(X)$ denotes the smallest cube of $\{-1, +1\}^U$ containing the topes of $\text{F}(X)$. A *facet* of $\mathcal{M}$ is an inclusion maximal proper face. From the definition, any face $\text{F}(X)$ consists of the sign vectors of all faces of the subcube of $[-1, +1]^U$ with barycenter X. By [4, Lemma 4], each face $\text{F}(X)$ of a COM $\mathcal{M}$ is isomorphic to an OM, which however is not simple, because all $Y \in \text{F}(X)$ coincide on $\underline{X}$. Thus, we consider its *simplification* $\mathcal{M}(X)$ obtained by deleting all the elements of $\underline{X}$. Deletion again gives an OM as is explained in Section 2.3. *Ample classes* (called also lopsided [3,50] or extremal [10,57]) are exactly the COMs, in which all faces are cubes. Since OMs are COMs, each face of an OM is an OM and the facets correspond to cocircuits. Furthermore, by [4, Section 11] replacing each combinatorial face $\text{F}(X)$ of $\mathcal{M}$ by a PL-ball, we obtain a contractible cell complex associated to each COM. The *topes* $\mathcal{T}$ and the *tope graph* $G(\mathcal{M})$ of a COM $\mathcal{M}$ are defined as for OMs. Again, the COM $\mathcal{M}$ can be recovered from $G(\mathcal{M})$, see [4,47]. For $X \in \mathcal{L}$, the topes in $\text{F}(X)$ induce a subgraph of $G(\mathcal{M})$, which we denote by $[X]$. We show that $[X]$ is isomorphic to the tope graph $G(\mathcal{M}(X))$ of $\mathcal{M}(X)$ and it is crucial for this paper.

2.2. **Realizable COMs.** In this subsection, we recall the geometric illustration of the axioms in the case of realizable COMs given in the paper [4]. Let U be an affine arrangement of hyperplanes of $\mathbb{R}^d$ and C an open convex set. Restrict the arrangement pattern to C, that is, remove all sign vectors which represent the open regions disjoint from C. Denote the resulting set of sign vectors by $\mathcal{L}(U, C)$ and call it a *realizable COM*. If U is a central arrangement with C being any open convex set containing the origin, then $\mathcal{L}(U, C)$ coincides with the realizable oriented matroid of U. If the arrangement U is affine and C is the entire space, then $\mathcal{L}(U, C)$ coincides with the realizable affine oriented matroid of U. The realizable ample sets arise by taking the central arrangement U of all coordinate hyperplanes restricted to an arbitrary open convex set C of $\mathbb{R}^d$ (this model was first considered in [50]).

We argue, why a realizable COM satisfies the axioms from Definition 3. Let X and Y be sign vectors belonging to $\mathcal{L}(U,C)$ and designating two open regions of C defined by U. Let x,y be two points in these regions. Connect x,y by a line segment and choose $\epsilon > 0$ so that the open ball of radius ϵ around x is contained in C and intersects only those hyperplanes from U containing x. Pick any point w from the intersection of this ball with the open line segment between x and y. The corresponding sign vector W is the composition $X \circ Y$, establishing (C). If we select a point v on the ray from y through x within the ϵ-ball but beyond x, then the corresponding sign vector V has the opposite signs as W at the coordinates corresponding to the hyperplanes from U containing x and not including the ray from y through x. Hence, $V = X \circ -Y$, yielding (FS). Now, assume that the hyperplane e from U separates x and y, that is, the line segment between x and y crosses e at some point z. The corresponding sign vector Z is then zero at e and equals the composition $X \circ Y$ at all coordinates where X and Y are sign-consistent, establishing (SE). If the hyperplanes of U have a non-empty intersection in C, then any point o from this intersection corresponds to the zero sign vector, showing that central hyperplane arrangements define OMs. In this case, $\mathcal{L}(U,C)$ coincides with $\mathcal{L}(U,\mathbb{R}^d)$ as well as with $\mathcal{L}(U,C_\epsilon)$, where C_ϵ is any open ball centered at o. The face $\mathrm{F}(X)$ of a covector $X \in \mathcal{L}(U,C)$ is obtained by taking any point $x \in C$ corresponding to X and a small ϵ-ball C_ϵ centered at x. Then $\mathrm{F}(X)$ coincides with the OM $\mathcal{L}(U,C_\epsilon)$. Finally, notice that the topes of $\mathcal{L}(U,C)$ correspond to the connected components of C minus the hyperplanes of U. Two such topes are adjacent in the tope graph if and only if the corresponding regions are separated by a single hyperplane. Furthermore, the distance between any two topes in the tope graph of $\mathcal{L}(U,C)$ is equal to the number of hyperplanes separating the two regions corresponding to these topes (for $C = \mathbb{R}^d$ this was proved by Deligne [21, Proposition 1.3]).

2.3. **Deletions and duality.** We continue with deletions in OMs and COMs. Let $\mathcal{M} = (U,\mathcal{L})$ be a COM and $A \subseteq U$. Given a sign vector $X \in \{\pm1,0\}^U$, by $X \setminus A$ (or by $X_{|U\setminus A}$) we refer to the *restriction* of X to $U \setminus A$, that is $X \setminus A \in \{\pm1,0\}^{U\setminus A}$ with $(X \setminus A)_e = X_e$ for all $e \in U \setminus A$. The *deletion* of A is defined as $\mathcal{M}\setminus A = (U\setminus A, \mathcal{L}\setminus A)$, where $\mathcal{L}\setminus A := \{X\setminus A : X \in \mathcal{L}\}$. We often consider the following type of deletion. For a covector $X \in \mathcal{L}$, we denote by $\mathcal{M}(X) = (U \setminus \underline{X}, \mathrm{F}(X) \setminus \underline{X})$ the simple OM defined by the face $\mathrm{F}(X)$. Note that $\mathcal{M}(X) = \mathcal{M} \setminus \underline{X}$, since for every $Y \in \mathcal{L}$ we have that $Y \setminus \underline{X} = (X \circ Y) \setminus \underline{X}$ and $X \circ Y \in \mathrm{F}(X)$. The classes of COMs and OMs are closed under deletion, see [4, Lemma 1]. The cocircuits and the covectors of deletions of OMs are described in the following way:

Lemma 1. *[8] Let $\mathcal{M} = (U,\mathcal{L})$ be an* OM *with the set of cocircuits $\mathcal{C}^*$ and $A \subseteq U$. Then the cocircuits of $\mathcal{M} \setminus A$ are the minimal elements of $\mathcal{C}^* \setminus A$ and the covectors of $\mathcal{M} \setminus A$ are $\mathcal{L} \setminus A$.*

We briefly recall the duality of OMs, see [8, Section 3.4]. The duality is defined via orthogonality of circuits and cocircuits, which can be viewed as a synthetic version of classical orthogonality of vectors. Two sign-vectors $X,Y \in \{\pm1,0\}^U$ are *orthogonal*, denoted $X \perp Y$, if either $\underline{X} \cap \underline{Y} = \varnothing$ or there are $e,f \in \underline{X} \cap \underline{Y}$ such that $X_eY_e = -X_fY_f$. Oriented matroids can be defined in terms of their *vectors* $\mathcal{V}$ and *circuits* $\mathcal{C}$, which can be derived from the cocircuits $\mathcal{C}^*$ using the following result:

Theorem 1. *[8, Theorem 3.4.3 and Proposition 3.7.12] Let $\mathcal{M}$ be an OM. The set $\mathcal{V}$ consists of all $Y \in \{\pm1,0\}^U$ such that $Y \perp X$ for any $X \in \mathcal{C}^*$ and $\mathcal{C}$ consists of the minimal members of $\mathcal{V} \setminus \{\boldsymbol{0}\}$.*

We will also make use of the version of Lemma 1 for circuits:

Lemma 2. *[8] Let $\mathcal{M} = (U,\mathcal{L})$ be an* OM *with the set of circuits $\mathcal{C}$ and $A \subseteq U$. Then the circuits of $\mathcal{M} \setminus A$ are $X \in \mathcal{C}$ such that $\underline{X} \cap A = \varnothing$.*

Remark 1. Throughout the paper we will use letters like S, S' for samples, T, T' for topes, and X, Y, Z for cocircuits, covectors, and circuits.

2.4. **Partial cubes and pc-minors.** It is well-known, see for example [4, 8], that tope graphs of OMs and COMs are partial cubes, which we introduce now. Let $G = (V, E)$ be a finite, connected, simple graph. The *distance* $d(u, v) := d_G(u, v)$ between vertices u and v is the length of a shortest (u, v)-path, and the *interval* $I(u, v) := \{x \in V : d(u, x) + d(x, v) = d(u, v)\}$ consists of all vertices on shortest (u, v)-paths. A subgraph H is *convex* if $I(u, v) \subseteq H$ for any $u, v \in H$ and *gated* [28] if for every vertex $x \notin H$ there exists a vertex x' (the *gate* of x) in H such that $x' \in I(x, y)$ for each vertex y of H. It is easy to see that gates are unique and that gated sets are convex. An induced subgraph H of G is *isometric* if the distance between vertices in H is the same as that in G. A graph $G = (V, E)$ is *isometrically embeddable* into a graph $H = (W, F)$ if there exists $\varphi : V \to W$ such that $d_H(\varphi(u), \varphi(v)) = d_G(u, v)$ for all $u, v \in V$. A graph G is a *partial cube* if it admits an isometric embedding into a hypercube $\mathrm{Q}_m = \mathrm{Q}(U)$. For an edge uv of G, let $W(u, v) = \{x \in V : d(x, u) < d(x, v)\}$. For an edge uv, the sets $W(u, v)$ and $W(v, u)$ are called *complementary halfspaces* of G.

Theorem 2. *[25] A graph G is a partial cube if and only if G is bipartite and for any edge uv the sets $W(u, v)$ and $W(v, u)$ are convex.*

Djoković [25] introduced the following binary relation Θ on the edges of G: for two edges $e = uv$ and $e' = u'v'$, we set $e\Theta e'$ if $u' \in W(u, v)$ and $v' \in W(v, u)$. If G is a partial cube, then Θ is an equivalence relation. Each Θ-class E_e corresponds to a coordinate $e \in U$ of the hypercube $\mathrm{Q}(U)$ into which G is isometrically embedded. Let $\{G_e^-, G_e^+\}$ be the complementary halfspaces of G defined by setting $G_e^- := G(W(u, v))$ and $G_e^+ := G(W(v, u))$ for an arbitrary edge $uv \in E_e$ (for $S \subseteq V(G)$ we denote by $G(S)$ the subgraph of G induced by S). An *elementary pc-restriction* consists of taking one of the halfspaces G_e^- and G_e^+. A *pc-restriction* is a convex subgraph of G induced by the intersection of a set of halfspaces of G. Since any convex subgraph of a partial cube G is the intersection of halfspaces [1, 2, 13], the pc-restrictions of G coincide with the convex subgraphs of G. Denote by $\pi_e(G)$ an *elementary pc-contraction*, i.e., the graph obtained from G by contracting the edges in E_e. For a vertex v of G, let $\pi_e(v)$ be the image of v under the contraction. We apply π_e to subsets $S \subseteq V$, by setting $\pi_e(S) := \{\pi_e(v) : v \in S\}$. By [14, Theorem 3], the class of partial cubes is closed under pc-contractions. Since pc-contractions commute, for a set A of Θ-classes, we denote by $\pi_A(G)$ the isometric subgraph of $\mathrm{Q}(U \setminus A)$ obtained from G by contracting the equivalence classes of edges from A. pc-Contractions and pc-restrictions also commute in partial cubes. A *pc-minor* of G is a partial cube obtained from G by pc-restrictions and pc-contractions. A deletion $\mathcal{M} \setminus A$ in a COM $\mathcal{M}$ translates to the contraction of the Θ-classes E_e with $e \in A$ in its tope graph $G(\mathcal{M})$. Since tope graphs of COMs and OMs are partial cubes, we can describe pc-restrictions and pc-contractions on sign-vectors in terms of partial cubes. First recall the following fundamental lemma from [4] and [47]:

Lemma 3. *For each covector X of a COM $\mathcal{M}$, $[X]$ is a gated subgraph of the tope graph $G(\mathcal{M})$ of $\mathcal{M}$. Moreover, for any tope T of $\mathcal{M}$, $X \circ T$ is the gate of T in $[X]$ and in $\mathrm{Q}(X)$.*

Let G be an isometric subgraph of the hypercube $Q(U)$ and H be an isometric subgraph of the hypercube $Q(U \setminus A)$ for some $A \subseteq U$. We say that G and H are *U-isomorphic* if there exists an isomorphism between G and H which maps each edge of a Θ-class E_e of G to an edge of E_e of H.

Lemma 4. *Let $\mathcal{M} = (U, \mathcal{L})$ be a COM and $A \subseteq U$. Then $\pi_A(G(\mathcal{M}))$ is the tope graph of $\mathcal{M} \setminus A$. If $X \in \mathcal{L}$, then the tope graph $[X]$ of $(U, F(X))$ is U-isomorphic to the tope graph $G(\mathcal{M}(X)) = \pi_{\underline{X}}(G(\mathcal{M}))$ of $\mathcal{M}(X) = \mathcal{M} \setminus \underline{X}$.*

Proof. That $G(\mathcal{M} \setminus A) = \pi_A(G(\mathcal{M}))$ follows from the equivalence between deletion in COMs and pc-contraction in their tope graphs. To prove that $[X]$ is U-isomorphic to $G(\mathcal{M}(X))$, note that

$[X]$ is obtained from $G(\mathcal{M})$ by a pc-restriction: $[X]$ is the intersection of the halfspaces defined by the Θ-classes E_e with $e \in \underline{X}$ and containing $[X]$. We assert that the pc-restrictions and the pc-contractions over $\underline{X}$ give the same result, i.e., that $\pi_{\underline{X}}(G(\mathcal{M}))$ is U-isomorphic to $[X]$. Indeed, by Lemma 3, $[X]$ is a gated subgraph of $G(\mathcal{M})$. Pick any $e \in \underline{X}$ and consider the elementary pc-contraction of the class E_e. By Lemma 3, the gate of any tope T of $\mathcal{M}$ in $[X]$ and in the cube $\mathrm{Q}(X)$ is $X \circ T$. Therefore, if $T, T' \in \{-1,+1\}^U$ such that $\mathrm{Sep}(T,T') = e$, T is a vertex of $G(\mathcal{M})$ not belonging to $[X]$, and T' belongs to $\mathrm{Q}(X)$, then necessarily $T' = X \circ T$ and thus T' must be a vertex of $[X]$. This implies that the intersection of the cube $\mathrm{Q}(X)$ with $\pi_e(G(\mathcal{M}))$ (which is the tope graph of the face of X in $\mathcal{M} \setminus e$) coincides with $[X]$. Consequently, $[X]$ coincides with $\pi_e(G(\mathcal{M}))$. Performing elementary pc-contractions for all elements of $\underline{X}$ we conclude that $[X]$ is U-isomorphic to $\pi_{\underline{X}}(G(\mathcal{M})) = G(\mathcal{M}(X))$. □

2.5. **VC-dimension.** Let $\mathcal{S}$ be a family of subsets of an m-element set U. A subset X of U is *shattered* by $\mathcal{S}$ if for all $Y \subseteq X$ there exists $S \in \mathcal{S}$ such that $S \cap X = Y$. The *Vapnik-Chervonenkis dimension* (VC-dimension) [69] $\text{VC-dim}(\mathcal{S})$ of $\mathcal{S}$ is the cardinality of the largest subset of U shattered by $\mathcal{S}$. Any set system $\mathcal{S} \subseteq 2^U$ can be viewed as a subset of vertices of the m-dimensional hypercube $\mathrm{Q}_m = \mathrm{Q}(U)$. Denote by $G(\mathcal{S})$ the *1-inclusion graph* of $\mathcal{S}$, i.e., the subgraph of $\mathrm{Q}(U)$ induced by the vertices of $\mathrm{Q}(U)$ corresponding to $\mathcal{S}$. A subgraph G of $\mathrm{Q}(U)$ has VC-dimension d if G is the 1-inclusion graph of a set system of VC-dimension d. For partial cubes, the notions of shattering and VC-dimension can be formulated in terms of pc-minors. First, note that if G' is a pc-minor of a partial cube G and G' shatters a subset X of U, then G also shatters X. Thus a partial cube G has VC-dimension $\leq d$ if and only if G does not have the hypercube Q_{d+1} as a pc-minor. More precisely a subset $D \subseteq U$ of the Θ-classes of G shatters G if $\pi_{U \setminus D}(G)$ is isomorphic to a hypercube. This is well-defined, since the embeddings of partial cubes are unique up to isomorphism, see e.g. [62, Chapter 5].

The *VC-dimension* $\text{VC-dim}(\mathcal{M})$ of a COM $\mathcal{M} = (U, \mathcal{L})$ is the VC-dimension of its tope graph $G(\mathcal{M})$ and we say that $D \subseteq U$ is shattered by $\mathcal{M}$ if D is shattered by $G(\mathcal{M})$. The *VC-dimension* $\text{VC-dim}(X)$ of a covector $X \in \mathcal{L}$ of $\mathcal{M}$ is the VC-dimension of the OM $\mathcal{M}(X)$, i.e., by Lemma 4, it is the VC-dimension of the graph $[X]$. The VC-dimension of OMs, COMs, and their covectors can be expressed in the following way:

Lemma 5. *[19, Lemma 13] For a COM $\mathcal{M}$, $\text{VC-dim}(\mathcal{M}) = \max\{\text{VC-dim}(\mathcal{M}(X)) : X \in \mathcal{L}\}$. If $\mathcal{M}$ is an OM and X a cocircuit of $\mathcal{M}$, then $\text{VC-dim}(X) + 1 = \text{VC-dim}(\mathcal{M}) = \text{rank}(\mathcal{M})$.*

That $\text{VC-dim}(X) = \text{VC-dim}(\mathcal{M}) - 1$ for cocircuits X of an OM $\mathcal{M}$ follows from the fact that the cocircuits are atoms of the big face lattice $\mathcal{F}_{\text{big}}(\mathcal{M})$ and this lattice is graded.

3. Auxiliary results

We establish and recall several auxiliary results about OMs and COMs. We also develop a correspondence between realizable samples and convex subgraphs of partial cubes. Finally, we define upper and lower covectors for a given sample, which are crucial notions for the main result.

3.1. **More about shattering in OMs and COMs.** We continue with new results about shattering in OMs and COMs. Let G be a partial cube, H a convex subgraph, and E_e a Θ-class of G. We say that E_e *crosses* H if H contains an edge of E_e. If E_e does not cross H and there exists an edge uv of E_e with $u \in H$ and $v \notin H$, then E_e and H *osculate*. Otherwise, E_e is *disjoint* from H. Denote by $\mathrm{osc}(H)$ the set of all e such that E_e osculates with H and by $\mathrm{cross}(H)$ the set of all e such that E_e crosses H.

Lemma 6. *Let G be a partial cube, H a convex subgraph of G, and $e \notin \mathrm{osc}(H)$. Then $\pi_e(H)$ is convex in $\pi_e(G)$ and $\mathrm{osc}(\pi_e(H)) = \mathrm{osc}(H)$, where $\mathrm{osc}(H)$ and $\mathrm{osc}(\pi_e(H))$ are considered in G and $\pi_e(G)$, respectively.*

Proof. Let $H' = \pi_e(H)$. First, since $e \notin \mathrm{osc}(H)$, the fact that H' is a convex subgraph of $\pi_e(G)$ comes from [17, Lemma 5]. Then, the inclusion $\mathrm{osc}(H) \subseteq \mathrm{osc}(H')$ is obvious. If there exists $e' \in \mathrm{osc}(H') \setminus \mathrm{osc}(H)$, then there exists an edge $\pi_e(u)\pi_e(v)$ in $\pi_e(E_{e'})$ with $\pi_e(u) \in V(H')$ and $\pi_e(v) \notin V(H')$. Then $\pi_e(u)\pi_e(v)$ comes from an edge uv of G belonging to $E_{e'}$. Since $e' \notin \mathrm{osc}(H)$ and the vertex v does not belong to H, the vertex u also does not belong to H. This implies that there exists an edge uw of E_e with $w \in V(H)$. If E_e and H contain an edge $u'w'$ and say $d(u, u') < d(w, u')$, then $u \in I(w, u')$, which contradicts the convexity of H. Thus E_e and H osculate, a contradiction. This establishes the equality $\mathrm{osc}(\pi_e(H)) = \mathrm{osc}(H)$. □

Lemma 7. *Let G be a partial cube and H a gated subgraph of G. If $D \subseteq \mathrm{cross}(H)$ is shattered by G, then D is shattered by H.*

Proof. Pick any Θ-class E_e with $e \in D$ and let v be any vertex of G. If v belongs to the halfspace G_e^- of G, then the gate v' of v in H also belongs to G_e^-. Indeed, since E_e crosses H, there exists a vertex $w \in G_e^- \cap H$. Then $v' \in I(v, w) \subset G_e^-$ by convexity of G_e^- and because v' is the gate of v in H. Analogously, if $v \in G_e^+$, then $v' \in G_e^+$.

Since G shatters D, for any sign vector $X \in \{-, +\}^D = \{-1, +1\}^D$, there exists a vertex v_X of G, whose restriction to D coincides with X. This means that for any $e \in D$, the vertex v_X belongs to the halfspace $G_e^{X_e}$. Since the gate v'_X of v_X in H also belongs to $G_e^{X_e}$, the restriction of v'_X to D also coincides with X. This implies that H also shatters D. □

The next lemma shows that the sets shattered by an OM $\mathcal{M}$ are exactly the *independent sets* of the underlying matroid $\underline{\mathcal{M}}$, i.e., the sets not containing supports of circuits of $\mathcal{M}$.

Lemma 8. *Let $\mathcal{M} = (U, \mathcal{L})$ be an* OM *and D be a subset of U. Then D is shattered by $\mathcal{M}$ if and only if D is independent in the underlying matroid $\underline{\mathcal{M}}$.*

Proof. By definition D is shattered by $\mathcal{M}$ if and only if D is shattered by $G(\mathcal{M})$. This is equivalent to $\pi_{U \setminus D}(G(\mathcal{M})) = Q_{U \setminus D}$. But since we have $\pi_{U \setminus D}(G(\mathcal{M})) = G(\mathcal{M}_{|D})$ this means $\mathcal{L}(\mathcal{M}_{|D}) = \{\pm 1, 0\}^D$. By Theorem 1 this is equivalent to $\mathcal{V}(\mathcal{M}_{|D}) = \{\mathbf{0}\}$ and $\mathcal{C}(\mathcal{M}_{|D}) = \varnothing$. Applying Lemma 2 this just means that the support of no circuit of $\mathcal{M}$ is contained in D. By definition this means that D is independent in $\underline{\mathcal{M}}$. □

An *antipode* of a vertex v in a partial cube G is a (necessarily unique) vertex $-v$ such that $G = I(v, -v)$. A partial cube G is *antipodal* if all its vertices have antipodes. By (Sym), a tope graph of a COM is the tope graph of an OM if and only if it is antipodal, see [47].

The next lemma can be seen as dual analogue of Lemma 8. It shows that the VC-dimension of OMs is defined locally at each tope T, by shattering subsets of $\mathrm{osc}(T)$.

Lemma 9. *Let $\mathcal{M} = (U, \mathcal{L})$ be an* OM *of rank d with tope graph $G(\mathcal{M})$. For any tope T of $\mathcal{M}$, $\mathrm{osc}(T)$ contains a subset D of size d shattered by $\mathcal{M}$.*

Proof. We proceed by induction on the size of U. If $\mathrm{osc}(T) = U$, then we are obviously done. Thus suppose that there exists $e \notin \mathrm{osc}(T)$. Consider the tope graph $G' = \pi_e(G)$ of the oriented matroid $\mathcal{M}' = \mathcal{M} \setminus e$. Let $T' = \pi_e(T)$. Then $\mathrm{osc}(T') = \mathrm{osc}(T)$ by Lemma 6. If $\mathrm{rank}(\mathcal{M}') = d$, by the induction hypothesis the set $\mathrm{osc}(T')$ contains a subset D of size d shattered by G'. Since G' is a pc-minor of G, the set $D \subset \mathrm{osc}(T') = \mathrm{osc}(T)$ is also shattered by G and we are done.

Thus, let $\mathrm{rank}(\mathcal{M}') < \mathrm{rank}(\mathcal{M})$. If the Θ-class E_e of G crosses the faces $\mathrm{F}(X)$ of all cocircuits $X \in \mathcal{L}$, then $\mathcal{L}$ is not simple. Therefore, there exists a cocircuit $X \in \mathcal{L}$ whose face $\mathrm{F}(X)$ is not crossed by E_e. However, since when we contract E_e the rank decreases by 1, the resulting OM $\mathcal{M}'$ coincides with $\mathrm{F}(X)$. Indeed, after contraction the rank of $\mathrm{F}(X)$ remains the same. Hence, if X would remain a cocircuit, then the global rank would not decrease. Hence, G' is the tope graph of $\mathcal{M}(X)$. Since G is an antipodal partial cube and $G_e^+ = \mathrm{F}(X)$, we have $G_e^- \cong G_e^+$. This shows that $G \cong G_e^+ \square K_2 \cong G' \square K_2$. This implies that E_e osculate with $\{T\}$ in G, contrary to the assumption $e \notin \mathrm{osc}(T)$. □

Next we give a shattering property of COMs. The *distance* $d(A,B)$ between sets A, B of vertices of G is $\min\{d(a,b) : a \in A, b \in B\}$. The set $\mathrm{pr}_B(A) = \{a \in A : d(a,B) = d(A,B)\}$ is the *metric projection* of B on A. For two covectors $X, Y \in \mathcal{L}$ of a COM $\mathcal{M}$, we denote by $\mathrm{pr}_{[X]}([Y])$ the metric projection of $[X]$ on $[Y]$ in $G(\mathcal{M})$. Since $[X]$ and $[Y]$ are gated by Lemma 3, $\mathrm{pr}_{[X]}([Y])$ consists of the gates of vertices of $[X]$ in $[Y]$, see [28]. Two faces $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ of $\mathcal{M}$ are *parallel* if $\mathrm{pr}_{[X]}([Y]) = [Y]$ and $\mathrm{pr}_{[Y]}([X]) = [X]$. A *gallery* between two parallel faces $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ of $\mathcal{M}$ is a sequence of faces $(\mathrm{F}(X) = \mathrm{F}(X_0), \mathrm{F}(X_1), \ldots, \mathrm{F}(X_{k-1}), \mathrm{F}(X_k) = \mathrm{F}(Y))$ such that either $k = 0$ (i.e., $\mathrm{F}(X) = \mathrm{F}(Y)$) or any two faces of this sequence are parallel and any two consecutive faces $\mathrm{F}(X_{i-1}), \mathrm{F}(X_i)$ are facets of a common face of $\mathcal{L}$. A *geodesic gallery* between $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ is a gallery of length $|\mathrm{Sep}(X,Y)|$. Two parallel faces $\mathrm{F}(X), \mathrm{F}(Y)$ are *adjacent* if $|\mathrm{Sep}(X,Y)| = 1$, i.e., $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ are opposite facets of a face of $\mathcal{L}$. See Figure 3 and recall the following result:

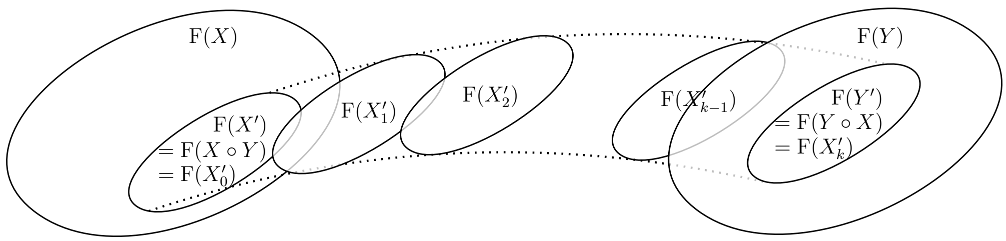

FIGURE 3. Illustration of Lemmas 10 and 11.

Lemma 10. *[19, Proposition 8] Let $\mathcal{M} = (U, \mathcal{L})$ be a COM and $X, Y \in \mathcal{L}$ (not necessarily distinct). Then:*

(i) *$d([X],[Y]) = |\mathrm{Sep}(X,Y)|$ and the gates of $[Y]$ in $[X]$ are the vertices of $[X \circ Y] \subseteq [X]$;*
(ii) *$\mathrm{F}(X)$ and $\mathrm{F}(Y)$ are parallel if and only if $\underline{X} = \underline{Y}$. If $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ are parallel, then they are connected by a geodesic gallery;*
(iii) *$\mathrm{pr}_{[Y]}([X]) = [X \circ Y]$, $\mathrm{pr}_{[X]}([Y]) = [Y \circ X]$, and $\mathrm{F}(X \circ Y)$ and $\mathrm{F}(Y \circ X)$ are parallel.*

A covector $X \in \mathcal{L}$ of a COM $\mathcal{M} = (U, \mathcal{L})$ *maximally shatters* a set $D \subseteq U$ if $[X]$ shatters D but $[X]$ does not shatter any superset of D. We also say that $X \in \mathcal{L}$ *locally maximally shatters* a set D if $[X]$ shatters D but D is not shattered by $[X']$ for any covector $X' > X$.

Lemma 11. *Let $\mathcal{M} = (U, \mathcal{L})$ be a COM and $X, Y \in \mathcal{L}$ (not necessarily distinct). Then:*

(i) *if $[X]$ and $[Y]$ shatter D, then the projections $[X \circ Y]$ and $[Y \circ X]$ also shatter D;*
(ii) *if $X \neq \pm Y$ and $[X]$ maximally shatters D and $[Y]$ shatters D, then $[X \circ Y] = [X]$ and $\mathrm{F}(X)$ is not a facet of $\mathcal{M}$;*
(iii) *if both $[X]$ and $[Y]$ shatter D, then there exist covectors $X' \geq X, Y' \geq Y$ such that $[X']$ and $[Y']$ both maximally shatter D, and $\mathrm{F}(X')$ and $\mathrm{F}(Y')$ are parallel. In particular, if $[X]$ shatters D, then there exists a covector $X' \geq X$ such that $[X']$ maximally shatters D.*

Proof. **Property (i):** Since $[X]$ and $[Y]$ shatter D, for any sign vector $Z \in \{\pm 1\}^D$ we can find two topes $T' \in [X]$ and $T'' \in [Y]$, such that $T'_{|D} = Z = T''_{|D}$. Since $X \leq T'$ and $Y \leq T''$, from $T'_{|D} = Z = T''_{|D}$ we conclude that $(X \circ Y)_{|D} < Z$ and in $[X \circ Y]$ we can find a tope T whose restriction to D coincides with Z. This proves that $[X \circ Y]$ shatters D, establishing (i).

Property (ii): If $[X]$ maximally shatters D, then $\text{VC-dim}(X) = |D| =: d$. By property (i), $[X \circ Y]$ also shatters D. If $\mathrm{F}(X \circ Y)$ is a proper face of $\mathrm{F}(X)$, then we obtain a contradiction with Lemma 5 applied to the OM $\mathcal{M}(X)$. Thus $\mathrm{F}(X \circ Y) = \mathrm{F}(X)$, showing that $X = X \circ Y$. This establishes the first assertion. By Lemma 10, the faces $\mathrm{F}(X)$ and $\mathrm{F}(Y \circ X)$ are parallel and therefore are connected by a geodesic gallery $(\mathrm{F}(X) = \mathrm{F}(X_0), \mathrm{F}(X_1), \ldots, \mathrm{F}(X_k) = \mathrm{F}(Y \circ X))$. Then either $k = 0$ and $\mathrm{F}(X) = \mathrm{F}(Y \circ X)$ holds or $\mathrm{F}(X)$ and $\mathrm{F}(X_1)$ are facets of a common face of $\mathcal{L}$. In the first case, since $X \neq \pm Y$, we conclude that $\mathrm{F}(X)$ is a proper face of $\mathrm{F}(Y)$, and thus is not a facet of $\mathcal{M}$. In the second case, $\mathrm{F}(X)$ is not a facet of $\mathcal{M}$ either. This proves (ii).

Property (iii): Let $d = |D|$. We can suppose that both X and Y locally maximally shatter the set D. Indeed, if D is shattered by a proper face $\mathrm{F}(X')$ of $\mathrm{F}(X)$, then we can replace the pair X, Y by the pair X', Y so that $[X']$ and $[Y]$ still shatter D. Thus D is not shattered by any proper faces of $\mathrm{F}(X)$ and $\mathrm{F}(Y)$. Since by (i), D is shattered by $[X \circ Y]$ and $[Y \circ X]$, we conclude that $X = X \circ Y$ and $Y = Y \circ X$ and thus the faces $\mathrm{F}(X)$ and $\mathrm{F}(Y)$ are parallel.

It remains to show that $[X]$ and $[Y]$ maximally shatter D. Suppose by way of contradiction that $[X]$ shatters a larger set $D' := D \cup \{e\}$. Consider the OM $\mathcal{M}' = \mathcal{M}(X) \setminus (U \setminus D')$. Note that $\mathcal{M}'$ maximally shatters D', i.e., VC-dim$(\mathcal{M}') = d + 1$. Since $[X]$ shatters D', the covectors of $\mathcal{M}'$ are $\{\pm 1, 0\}^{D'}$. Let X'' be a cocircuit of $\mathcal{M}'$ with $\underline{X'} = \{e\}$. By Lemma 5 applied to $\mathcal{M}'$, we conclude that X'' has VC-dimension d. Hence, X'' must shatter the set D. By Lemma 1, there is a cocircuit X' of $\mathrm{F}(X)$ such that $X'' = X' \setminus (U \setminus D')$. Since X'' shatters D, X' also shatters D. Since $X < X'$, this contradicts our assumption that X locally maximally shatters D. The second assertion follows by applying the first assertion with $Y = X$. This establishes (iii). $\square$

3.2. **Realizable and full samples as convex subgraphs.** Let $\mathcal{M} = (U, \mathcal{L})$, where $\mathcal{L} \subset \{\pm 1, 0\}^U$ is a system of sign vectors whose topes $\mathcal{T}$ induce an isometric subgraph G of $\mathrm{Q}(U)$. We denote by $\mathbf{Samp}(\mathcal{M}) = \mathbf{Samp}(\mathcal{L}) = \bigcup_{X \in \mathcal{L}} \{S \in \{\pm 1, 0\}^U : S \leq X\}$ the *set of realizable samples* for $\mathcal{M}$ (this is called the *polar complex* in neural codes [46]). Since for any $X \in \mathcal{L}$ there exists $T \in \mathcal{T}$ such that $X \leq T$, we have $\mathbf{Samp}(\mathcal{M}) = \mathbf{Samp}(\mathcal{T})$, see Figure 4.

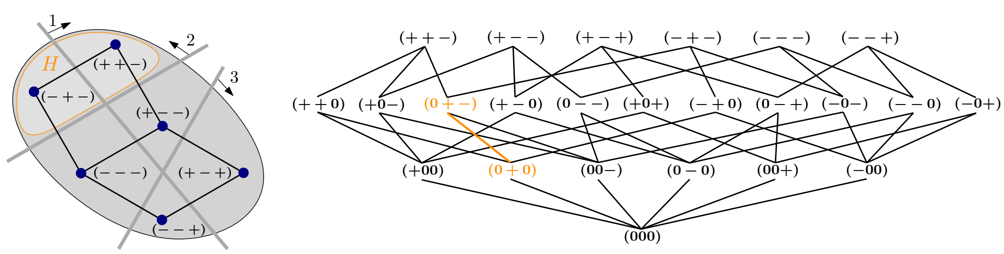

FIGURE 4. Left: the tope graph G of the pc-restriction $\mathcal{M}$ of the COM from Figure 1 to $\{1, 2, 3\}$ and a convex subgraph H of G. Right: the realizable samples of $\mathcal{M}$ and the interval $I(H)$ (in orange).

Extending the notation for covectors and their faces, for a sample $S \in \mathbf{Samp}(\mathcal{M})$ we set $\mathrm{F}(S) = \{X \in \mathcal{L} : S \leq X\}$ and let $[S]$ be the subgraph of G induced by all topes $T \in \mathcal{L}$ from $\mathrm{F}(S)$. For OMs, the set $\mathrm{F}(S)$ is called a *supertope* in [45]. For COMs, $\mathrm{F}(S)$ is called the *fiber* of S and it is known that they are COMs [4]. Since for any $S \in \mathbf{Samp}(\mathcal{M})$ there exists $T \in \mathcal{T}$ such that $S \leq T$, $[S] \neq \varnothing$. Moreover, $[S]$ is the intersection of halfspaces of G of the form G_e^+ if $S_e = +1$ and G_e^- if $S_e = -1$. Hence, $[S]$ is a nonempty convex subgraph of G for all $S \in \mathbf{Samp}(\mathcal{M})$.

Any convex subgraph H of a partial cube G is the intersection of all halfspaces of G containing H. Similarly to the fact that any polytope P in Euclidean space is the intersection of the halfspaces defined by its facet-defining hyperplanes, any convex set H in a partial cube is the intersection of the halfspaces defined by the Θ-classes in $\mathrm{osc}(H)$. Both for P and for H, this is a minimal representation as the intersection of halfspaces. However, H can be represented in different ways as the intersection of halfspaces. Indeed, any representation of H as an intersection of halfspaces of G yields a realizable sample S, where $S_e = \pm 1$ if $G_e^{\pm}$ participates in the representation and $S_e = 0$ otherwise. Notice that the Θ-classes osculating with H have to be part of every representation of H and the Θ-classes crossing H take part in no representation of H. This leads to two canonical representations of H, one using only the halfspaces whose Θ-class osculates with H and one using all halfspaces containing H:

$$(S_\perp)_e = \begin{cases} -1 & \text{if } e \in \operatorname{osc}(H) \text{ and } H \subseteq G_e^-, \\ +1 & \text{if } e \in \operatorname{osc}(H) \text{ and } H \subseteq G_e^+, \\ 0 & \text{otherwise.} \end{cases} \quad \text{and} \quad (S^\top)_e = \begin{cases} -1 & \text{if } H \subseteq G_e^-, \\ +1 & \text{if } H \subseteq G_e^+, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $(S^\top)^0 = \operatorname{cross}(H)$ and $(S_\perp)^0 = U \setminus \operatorname{osc}(H)$, i.e., $(S_\perp)^0$ consists of all e such that E_e crosses or is disjoint from H. If S is a sample arising from the representation of H as the intersection of halfspaces, then $S_\perp \leq S \leq S^\top$. Moreover, any sample S from the order interval $I(H) := [S_\perp, S^\top]$ arises from a representation of H, i.e., $[S] = [S_\perp] = [S^\top] = H$. Thus, for any convex subgraph H of G the set of all $S \in \mathbf{Samp}(\mathcal{M})$ such that $[S] = H$ is an interval $I(H) = [S_\perp, S^\top]$ of $(\mathbf{Samp}(\mathcal{M}), \leq)$. Note that the intervals $I(H)$ partition $\mathbf{Samp}(\mathcal{M})$. See Figure 4 for an illustration of the above. Moreover:

Lemma 12. *If $S, S' \in \mathbf{Samp}(\mathcal{L})$ and $S \leq S'$, then $[S'] \subseteq [S]$.*

Definition 4 (Full samples)**.** We say that a realizable sample $S \in \mathbf{Samp}(\mathcal{M})$ is *full* if the pc-minor $G' = \pi_{S^0}(G(\mathcal{M}))$ obtained from $G(\mathcal{M})$ by contracting the Θ-classes of S^0 has VC-dimension $d = \text{VC-dim}(\mathcal{M})$. Let $\mathbf{Samp}_f(\mathcal{M})$ denote the set of all full samples of $\mathcal{M}$.

Note that all topes of $\mathcal{M}$ are full samples since their zero set is empty. A convex subgraph H of G is *full* if the sample $S_\perp$ is full, where recall $I(H) = [S_\perp, S^\top]$. The image of H in G' (obtained from G by contracting the Θ-classes of $(S_\perp)^0 = U \setminus \operatorname{osc}(H)$) is a single vertex v_H and its degree is $|\operatorname{osc}(H)|$. If $D \subset \operatorname{osc}(v_H) = \operatorname{osc}(H)$ of size d is shattered by G', since G' is a pc-minor of G, D is also shattered by G. Hence, a convex set H of G is full if and only if G shatters a subset D of $\operatorname{osc}(H)$ of size $d = \text{VC-dim}(G)$. Since for any $S \in I(H)$ we have $(S^\top)^0 \subseteq S^0 \subseteq (S_\perp)^0$, if H is a full convex subgraph of a COM, then all samples in $I(H)$ are full. However, if S is a full sample and $H = [S]$, then not necessarily all samples from $I(H)$ are full:

Example 1. Let $\mathcal{M} = (U, \mathcal{L})$ be the COM with $\text{VC-dim}(\mathcal{M}) = 2$ defined on $U = \{1,2,3,4,5\}$ and whose tope graph consists of one edge on each of whose ends there is a pending 4-cycle. Formally, $\mathcal{M}$ has $(-,-,-,-,-), (+,-,-,-,-), (-,+,-,-,-), (+,+,-,-,-), (+,+,+,-,-), (+,+,+,+,-), (+,+,+,-,+), (+,+,+,+,+)$ as topes. The two 4-cycles are the convex sets $H_1 = [S_1]$ and $H_2 = [S_2]$ defined by the samples $S_1 = (0,0,-,-,-)$ and $S_2 = (+,+,+,0,0)$, while the middle edge is the convex set $H_3 = [S_3]$ where $S_3 = (+,+,0,-,-)$. Consider the samples $S_\perp$ and $S^\top$ for the convex set H_1: $S_\perp = (0,0,-,0,0)$ and $S^\top = (0,0,-,-,-) = S_1$. Notice that the sample $S_1 = S^\top$ is full since contracting $S_1^0 = \{1,2\}$ does not affect the other 4-cycle H_2. However, the convex set $H_1 = [S_1]$ is not full because the sample $S_\perp$ is not full: contracting $(S_\perp)^0 = \{1,2,4,5\}$, both 4-cycles will be contracted, thus the VC-dimension will decrease. Morally, being full is a local property in a COM.

We show next, that this problem does not arise in OMs.

Lemma 13. *Let $\mathcal{M} = (U, \mathcal{L})$ be an OM of rank d and let $G = G(\mathcal{M})$ be its tope graph. A sample $S \in \mathbf{Samp}(\mathcal{M})$ is full if and only if the convex subgraph $[S]$ is full.*

Proof. First notice that since in OMs the rank and the VC-dimension are equal, a sample S is full if and only if $\operatorname{rank}(\mathcal{M} \setminus S^0) = d$. Let S be a full sample, $H = [S]$, and recall that $(S_\perp)^0$ equals $\operatorname{cross}(H)$ plus the Θ-classes not osculating with H. We have to show that $\mathcal{M} \setminus (S_\perp)^0$ has rank d. First, let $\mathcal{M}' = \mathcal{M} \setminus \operatorname{cross}(H)$ and let $G' = \pi_{\operatorname{cross}(H)}(G)$ be its tope graph. Since $\operatorname{cross}(H) \subseteq S^0$ and S is full, $\operatorname{rank}(\mathcal{M}') = d$ and hence $\text{VC-dim}(G') = d$. The image of H in G' is a single vertex v_H. By Lemma 6, $\operatorname{osc}(v_H) = \operatorname{osc}(H)$. By Lemma 9, $\operatorname{osc}(v_H)$ contains a subset of size d shattered by $\mathcal{M}'$. Since $\operatorname{osc}(v_H) \cap (S_\perp)^0 = \varnothing$ we conclude that H is full. Conversely, if $H = [S]$ is a convex subgraph of G that is full, then from the discussion preceding Example 1 we deduce that all samples from $I(H)$ (and in particular, S) are full. $\square$

3.3. **The samples $\widehat{S}$ and $\widehat{\widehat{S}}$.** For a covector $X \in \mathcal{L}$ of a COM $\mathcal{M} = (U, \mathcal{L})$, let $\mathbf{Samp}(\mathrm{F}(X))$ denote the samples of OM $(U, \mathrm{F}(X))$, i.e., $\mathbf{Samp}(\mathrm{F}(X)) = \bigcup_{Y \in \mathrm{F}(X)} \{S \in \{\pm 1, 0\}^U : S \leq Y\}$. Clearly $\mathbf{Samp}(\mathrm{F}(X)) \subseteq \mathbf{Samp}(\mathcal{L})$. We also denote by $\mathbf{Samp}(\mathcal{M}(X))$ the samples of the simple OM $\mathcal{M}(X) = (U \setminus \underline{X}, F(X) \setminus \underline{X})$, i.e., $\mathbf{Samp}(\mathcal{M}(X)) = \bigcup_{Y \in \mathrm{F}(X) \setminus \underline{X}} \{S \in \{\pm 1, 0\}^{U \setminus \underline{X}} : S \leq Y\}$. Finally, denote by $\mathbf{Samp}_f(\mathrm{F}(X))$ the set of full samples from $\mathbf{Samp}(\mathrm{F}(X))$ and by $\mathbf{Samp}_f(\mathcal{M}(X))$ the set of full samples from $\mathbf{Samp}(\mathcal{M}(X))$.

As we noticed already, $F(X)$ is an OM, however it is not simple. Since all covectors from $F(X)$ have the values as X on the coordinates of $\underline{X}$, from the definition of $\mathcal{M}(X)$ we conclude that $F(X) = (F(X) \setminus \underline{X}) \times X_{|\underline{X}}$. This establishes a one-to-one correspondence φ_X between the covectors of $F(X)$ and the covectors of $\mathcal{M}(X)$ and between the topes of $\mathrm{F}(X)$ and the topes of $\mathcal{M}(X)$. Recall that by Lemma 4 the tope graph $[X]$ of $F(X)$ is isomorphic to the tope graph $G(\mathcal{M}(X))$ of $\mathcal{M}(X)$. Therefore, $[X]$ and $G(\mathcal{M}(X))$ have the same sets of convex subgraphs. This one-to-one correspondence φ_X also shows that any $S \in \mathbf{Samp}(\mathrm{F}(X))$ has the form $S = S' \times S''$, where $S' \in \mathbf{Samp}(\mathcal{M}(X))$ and $S'' \in \{\pm 1, 0\}^{\underline{X}}$ such that $S'' \leq X_{|\underline{X}}$.

The following samples $\widehat{S}$ and $\widehat{\widehat{S}}$ will be important in what follows:

Definition 5 ($\widehat{S}$ and $\widehat{\widehat{S}}$)**.** For a sample $S \in \mathbf{Samp}(\mathcal{L})$ and a covector $X \in \mathcal{L}$, set $\widehat{S} := X \circ S$ and $\widehat{\widehat{S}} := \widehat{S} \setminus \underline{X} = S \setminus \underline{X}$, where it will usually be clear which covector $X \in \mathcal{L}$ we are referring to.

From the definition it immediately follows that $\widehat{S}$ and $\widehat{\widehat{S}}$ have the same zero sets: $\widehat{S}^0 = \widehat{\widehat{S}}^0$. We continue with the following properties of $\widehat{S}$ and $\widehat{\widehat{S}}$:

Lemma 14. *Let $X \in \mathcal{L}$, $S \in \mathbf{Samp}(\mathcal{M})$, and $\mathrm{Sep}(X, S) = \varnothing$. Then $\widehat{S} \in \mathbf{Samp}(\mathrm{F}(X)), \widehat{\widehat{S}} \in \mathbf{Samp}(\mathcal{M}(X))$, the convex subgraphs $[\widehat{S}]$ of $[X]$ and $[\widehat{\widehat{S}}]$ of $G(\mathcal{M}(X))$ are U-isomorphic, and $[\widehat{S}] = [X] \cap [S] \neq \varnothing$.*

Proof. Since $S \in \mathbf{Samp}(\mathcal{M})$, there exists $Y \in \mathcal{L}$ such that $S \leq Y$. Then $\widehat{S} = X \circ S \leq X \circ Y$. Since $X \circ Y \in \mathrm{F}(X)$, we get $X \circ S \in \mathbf{Samp}(\mathrm{F}(X))$. Since $(X \circ Y) \setminus \underline{X} \in \mathcal{M}(X)$ and $\widehat{\widehat{S}} = \widehat{S} \setminus \underline{X} \leq (X \circ Y) \setminus \underline{X}$, we also deduce that $\widehat{\widehat{S}} \in \mathbf{Samp}(\mathcal{M}(X))$. From the definition of the convex subgraphs $[\widehat{S}]$ and $[\widehat{\widehat{S}}]$ and the way how the bijection φ_X between the topes of $[X]$ and the topes of $G(\mathcal{M}(X))$ is defined, we conclude that the convex subgraphs $[\widehat{\widehat{S}}]$ and $[\widehat{S}]$ are U-isomorphic.

Now, we prove that $[\widehat{S}] = [X] \cap [S] \neq \varnothing$. Since $\widehat{S} = X \circ S$, we have $X \leq \widehat{S}$ and by Lemma 12 we have $[\widehat{S}] \subseteq [X]$. Now we prove that $[\widehat{S}] \subseteq [S]$. Indeed, otherwise there exists a tope T of $\mathcal{L}$ such that $T \in [\widehat{S}] \setminus [S]$. This implies that $\widehat{S} \leq T$ and there exists an element $e \in U$ such that $T_e \neq S_e \neq 0$. If $X_e = 0$, then $\widehat{S}_e = (X \circ S)_e = S_e \neq 0$. Since $\widehat{S} \leq T$ this implies that $S_e = \widehat{S}_e = T_e$, a contradiction with the choice of e. Otherwise, if $X_e \neq 0$, then $\widehat{S}_e = (X \circ S)_e = X_e$. This implies that $T_e = X_e$, which is impossible because $T_e \neq S_e \neq 0$ and we have $\mathrm{Sep}(X, S) = \varnothing$. This proves that $[\widehat{S}] \subseteq [X] \cap [S]$. Consequently, $[X] \cap [S] \neq \varnothing$. To prove the converse inclusion $[X] \cap [S] \subseteq [\widehat{S}]$ pick any tope T of $\mathcal{L}$ belonging to $[X] \cap [S]$. Then $X \leq T$ and $S \leq T$ and thus $\widehat{S} = X \circ S \leq T$. This implies that $T \in [\widehat{S}]$ and we are done. □

Lemma 15. *Let $X \in \mathcal{L}, S \in \mathbf{Samp}(\mathcal{M})$, and $\mathrm{Sep}(X, S) = \varnothing$. Then:*

- (i) $\mathrm{osc}([\widehat{\widehat{S}}]) = \mathrm{osc}([\widehat{S}]) = \mathrm{osc}([S]) \cap X^0$, *where* $\mathrm{osc}([\widehat{S}])$ *is considered in* $[X]$ *and* $\mathrm{osc}([\widehat{\widehat{S}}])$ *in* $G(\mathcal{M}(X))$;
- (ii) $\widehat{\widehat{S}}^0 = \widehat{S}^0 = S^0 \cap X^0$.

Proof. To prove (i), first notice that since $[\widehat{\widehat{S}}]$ and $[\widehat{S}]$ are U-isomorphic by Lemma 14 and $[X]$ and $G(\mathcal{M}(X))$ are U-isomorphic by Lemma 4, we obtain that $\mathrm{osc}([\widehat{\widehat{S}}]) = \mathrm{osc}([\widehat{S}])$.

Now we show that $\mathrm{osc}([\widehat{S}]) \subseteq \mathrm{osc}([S]) \cap X^0$. Pick any $e \in \mathrm{osc}([\widehat{S}])$. Since $[\widehat{S}] \subseteq [X]$ and $[X]$ is isomorphic to $G(\mathcal{M}(X))$ by Lemma 4, the Θ-class E_e necessarily crosses $[X]$, whence $e \in X^0$.

Since $e \in \text{osc}([\widehat{S}])$, either $e \in \text{osc}([S])$ and we are done, or $e \in \text{cross}([S])$. Suppose by way of contradiction that $e \in \text{cross}([S])$. Then there exist two edges T_1T_2 and $T_1'T_2'$ of E_e such that $T_1 \in [\widehat{S}]$ and $T_2 \in [X] \setminus [\widehat{S}]$ and $T_1', T_2' \in [S]$. But then T_2 belongs to the interval either between T_1 and T_2' or between T_1 and T_1', contradicting the convexity of $[S]$. This proves that $e \in \text{osc}([S])$, establishing the inclusion $\text{osc}([\widehat{S}]) \subseteq \text{osc}([S]) \cap X^0$.

Conversely, pick any $e \in \text{osc}([S]) \cap X^0$. Then there exist two edges T_1T_2 and $T_1'T_2'$ of E_e, such that $T_1, T_2 \in [X]$ and $T_1' \in [S]$, $T_2' \notin [S]$. Since $X \in \mathcal{L}$, $[X]$ is gated. Denote by T_1'' and T_2'' the gates of respectively T_1' and T_2' in $[X]$: $T_1'' = X \circ T_1'$ and $T_2'' = X \circ T_2'$. Since T_1' and T_2' are adjacent, the topes T_1'' and T_2'' are either adjacent or coincide. Furthermore, since T_1'' belongs to the interval $I(T_1', T)$ between T_1' and any $T \in [S] \cap [X] \neq \varnothing$, the convexity of $[\widehat{S}] = [S] \cap [X]$ implies that $T_1' \in [\widehat{S}]$. Now, if $T_2'' = T_1''$, since $T_1' \in [S]$ and $T_2' \notin [S]$, the convexity of $[S]$ implies that T_1' is in the interval $I(T_2', T_2'') = I(T_2', T_1'')$. Since T_2'' is in the intervals $I(T_2', T_1)$ and $I(T_2', T_2)$, we conclude that T_1' also belongs to the intervals $I(T_2', T_1)$ and $I(T_2', T_2)$. But this is impossible because the edges $T_1'T_2'$ and T_1T_2 belong to the same Θ-class E_e. This proves that T_1'' and T_2'' are different and adjacent. Moreover, $T_1'' \in I(T_1', T_2'')$ and $T_2'' \in I(T_2', T_1'')$, proving that the edge $T_1''T_2''$ also belongs to the Θ-class E_e. Then we also have $T_1' \in I(T_2', T_1'')$ and $T_2' \in I(T_1', T_2'')$. Since $T_1' \in [S]$ and $T_2' \notin [S]$, the convexity of $[S]$ implies that $T_2'' \notin [S]$. Consequently, $T_1'' \in [S] \cap [X] = [\widehat{S}]$ and $T_2'' \in [X] \setminus [S] = [X] \setminus [\widehat{S}]$, establishing that $e \in \text{osc}([\widehat{S}])$. This proves the inclusion $\text{osc}([S]) \cap X^0 \subseteq \text{osc}([\widehat{S}])$ and concludes the proof of (i).

To prove (ii), first notice that $\widehat{\widehat{S}}^0 = \widehat{S}^0$ and that $\widehat{S}^0 \subseteq S^0 \cap X^0$. To prove the converse inclusion, pick any $e \in S^0 \cap X^0$. Then there exist two edges T_1T_2 and $T_1'T_2'$ of E_e, such that $T_1, T_2 \in [X]$ and $T_1', T_2' \in [S]$. As in previous proof, let T_1'' and T_2'' be the gates of T_1' and T_2' in $[X]$. Then as above we deduce that $T_1''T_2''$ is an edge of E_e belonging to $[X]$. If T is a tope of $[\widehat{S}] = [S] \cap [X]$ (such a tope exists by Lemma 14), then $T_1'' \in I(T_1', T)$ and $T_2'' \in I(T_2', T)$. Since $[S]$ is convex and $T_1', T_2' \in [S]$, we conclude that $T_1'', T_2'' \in [S]$. Consequently, $T_1''T_2''$ is an edge of $[\widehat{S}]$, hence $e \in \widehat{S}^0$, establishing (ii). □

3.4. **Lower and upper covectors.** Let $\mathcal{M} = (U, \mathcal{L})$ be a COM. We define lower and upper covectors for samples of $\mathcal{M}$. For a sample $S \in \mathbf{Samp}(\mathcal{M})$ of $\mathcal{M}$ consider the tope $T' = S \setminus S^0$ of $\mathcal{M}' := \mathcal{M} \setminus S^0$. Any minimal non-zero covector X' of $\mathcal{M}'$ such that $T' \geq X'$ is called a *lower covector* for S. Since $\mathcal{M}'$ is a COM and T' is a tope of $\mathcal{M}'$, lower covectors X' for S exist. Any covector of $\mathcal{M}$ such that $X \setminus S^0 = X'$ is called an *upper covector* for S. Again, upper covectors for S exist because X' is the restriction of some covector X of $\mathcal{M}$. Note that if $\mathcal{M}$ is an OM, then the lower and upper covectors are always cocircuits, which we will sometimes call *lower and upper cocircuits* for S. For lower covectors this follows by minimality, but for upper covectors this follows from S being full and is part of Lemma 19.

Recall that we denote by $\mathcal{M}'(X') = \mathcal{M}' \setminus \underline{X'}$ the simple OM defined by the face $\text{F}(X')$ of $\mathcal{M}'$ and by $\mathcal{M}(X) = \mathcal{M} \setminus \underline{X}$ the simple OM defined by the face $\text{F}(X)$ of $\mathcal{M}$.

Lemma 16. *If $S \in \mathbf{Samp}(\mathcal{M})$, X' is a lower covector for S, and $X \in \mathcal{L}$ is an upper covector for S such that $X \setminus S^0 = X'$, then $\text{Sep}(S, X) = \varnothing$ and $\text{VC-dim}(X) \geq \text{VC-dim}(X')$. Furthermore, if $\text{VC-dim}(X) = \text{VC-dim}(X')$, then $\widehat{\widehat{S}}$ is a full sample of $\mathcal{M}(X)$.*

Proof. Let X' be a lower covector for S. Since $X' = X \setminus S^0$ and $X' \leq T' = S \setminus S^0$, for any tope T of $\mathcal{M}$ such that $T' = T \setminus S^0$ (such tope T exists since $\mathcal{M}$ is simple), we have $S \leq T$ and $X \leq T$, yielding $T \in [S] \cap [X]$. Thus $\text{Sep}(S, X) = \varnothing$.

Now we prove that $\text{VC-dim}(X) \geq \text{VC-dim}(X')$. Let $\text{VC-dim}(X') = d$. By Lemma 9, there exists a set $D \subseteq \text{osc}([T']) \cap X'^0$ of size d shattered by $\mathcal{M}'(X')$. Since the tope graph of $\mathcal{M}'(X')$ is a pc-minor of $G(\mathcal{M})$, D is shattered by $\mathcal{M}$. Since $D \subseteq X'^0 \subseteq X^0$ and $[X]$ is a gated subgraph of the tope graph of $\mathcal{M}$, by Lemma 7, D is shattered by $\mathcal{M}(X)$. This shows that $\text{VC-dim}(X) \geq d = \text{VC-dim}(X')$.

Now suppose that VC-dim$(X) = d$ and we assert that $\widehat{\widehat{S}}$ is a full sample of $\mathcal{M}(X)$. By Lemma 13 applied to OM $\mathcal{M}(X)$, the sample $\widehat{\widehat{S}}$ is full if and only if the convex set $[\widehat{\widehat{S}}\,]$ is full. Since $\mathrm{Sep}(S, X) = \varnothing$, by Lemma 15(i), $\mathrm{osc}([\widehat{\widehat{S}}\,]) = \mathrm{osc}([S]) \cap X^0$. Since $D \subseteq \mathrm{osc}([T']) \cap X'^0$ and $\mathrm{osc}([T']) = \mathrm{osc}([S])$ (by Lemma 6), $X'^0 \subseteq X^0$, we deduce that $D \subseteq \mathrm{osc}([\widehat{\widehat{S}}\,])$. Consequently, $\mathcal{M}(X)$ shatters a set $D \subseteq \mathrm{osc}([\widehat{\widehat{S}}\,])$ of size d, establishing that the convex set $[\widehat{\widehat{S}}\,]$ is full in $\mathcal{M}(X)$. □

Lemma 17. *Let $S \in \mathbf{Samp}(\mathcal{M})$, X' be a lower covector for S, and $X \in \mathcal{L}$ be an upper covector for S such that $X' = X \setminus S^0$. Then $\mathcal{M}'(X') = \mathcal{M}(X) \setminus \widehat{\widehat{S}}^{\,0} = \mathcal{M}(X) \setminus \widehat{S}^0$. Consequently,* VC-dim$(X) \geq$ VC-dim(X').

Proof. First we prove the following claim:

Claim 1. $\underline{X'} \cup S^0 = \underline{X} \cup \widehat{S}^0$.

Proof. To prove the inclusion $\underline{X'} \cup S^0 \subseteq \underline{X} \cup \widehat{S}^0$ notice that $\underline{X'} \subseteq \underline{X}$ by the definition of X. If $e \in S^0 \setminus \underline{X}$, then $e \in X^0$. By Lemma 15(ii), $e \in S^0 \cap X^0 = \widehat{S}^0$, establishing that $\underline{X'} \cup S^0 \subseteq \underline{X} \cup \widehat{S}^0$. To prove the converse inclusion $\underline{X} \cup \widehat{S}^0 \subseteq \underline{X'} \cup S^0$ note that $\widehat{S}^0 \subseteq S^0$. If $e \in \underline{X} \setminus S^0$, then $e \in \underline{X'}$ because $X' = X \setminus S^0$, and we are done. □

Denote by $G(\mathcal{M}), G(\mathcal{M}')$, and $G(\mathcal{M}(X))$ the tope graphs of $\mathcal{M}, \mathcal{M}' = \mathcal{M} \setminus S^0$, and $\mathcal{M}(X)$, respectively. Denote also by G' the tope graph of $\mathcal{M}'(X')$ and by G'' the tope graph of $\mathcal{M}(X) \setminus \widehat{S}^0$. To prove that $\mathcal{M}'(X') = \mathcal{M}(X) \setminus \widehat{S}^0$ it suffices to establish that the tope graphs G' and G'' coincide. By Lemma 4, $[X]$ is isomorphic to $G(\mathcal{M}(X)) = \pi_{\underline{X}}(G(\mathcal{M}))$. Furthermore, by the same lemma, $G'' = G(\mathcal{M}(X) \setminus \widehat{S}^0) = \pi_{\widehat{S}^0}(G(\mathcal{M}(X))$. Consequently, $G'' = \pi_{\underline{X} \cup \widehat{S}^0} G(\mathcal{M})$. Analogously, by Lemma 4, $G' = G(\mathcal{M}'(X')) = \pi_{\underline{X'}}(G(\mathcal{M}'))$ and is isomorphic to $[X']$. Since $G(\mathcal{M}') = \pi_{S^0}(G(\mathcal{M}))$, we conclude that $G' = \pi_{S^0 \cup \underline{X'}}(G(\mathcal{M}))$. By Claim 1, $\underline{X'} \cup S^0 = \underline{X} \cup \widehat{S}^0$. Since the pc-contractions commute, we obtain that

$$G' = \pi_{S^0 \cup \underline{X'}}(G(\mathcal{M})) = \pi_{\underline{X} \cup \widehat{S}^0}(G(\mathcal{M})) = G'',$$

whence $\mathcal{M}'(X') = \mathcal{M}(X) \setminus \widehat{S}^0$. Since $\widehat{\widehat{S}}^{\,0} = \widehat{S}^0$, we obtain the equality $\mathcal{M}'(X') = \mathcal{M}(X) \setminus \widehat{S}^0 = \mathcal{M}(X) \setminus \widehat{\widehat{S}}^{\,0}$. Since $G' = G''$ is a pc-minor of $G(\mathcal{M}(X))$, also VC-dim$(X) \geq$ VC-dim(X') holds. □

In the following two results we suppose that $\mathcal{M} = (U, \mathcal{L})$ is an OM of VC-dimension d.

Lemma 18. *For any tope T of $\mathcal{M}$ and $e \in \mathrm{osc}([T])$, there exists a cocircuit X of $\mathcal{M}$ such that $e \in \underline{X}$, $X \leq T$, and $\mathcal{M}(X)$ has VC-dimension $d - 1$.*

Proof. Since T is a tope and $e \in \mathrm{osc}([T])$, T is incident to an edge of E_e, i.e., there is a tope T' of $\mathcal{M}$ such that $\mathrm{Sep}(T, T') = \{e\}$. Let X be a cocircuit of $\mathcal{M}$ such that its face $\mathrm{F}(X)$ contains T but not T'. This cocircuit X exists, otherwise all cocircuits Y of $\mathcal{M}$ would have $Y_e = 0$, contradicting the assumption that $\mathcal{M}$ is simple. Now, since $\mathcal{M}$ has VC-dimension d, $\mathcal{M}(X)$ has VC-dim $d-1$ by Lemma 5. Furthermore, as $T \in [X]$ and $T' \notin [X]$, we immediately get that $X \leq T$ and $e \in \underline{X}$. □

Lemma 19. *For any full sample S of $\mathcal{M}$ and $e \in \mathrm{osc}([S])$, there exists a lower cocircuit X' for S such that $e \in \underline{X'}$. For any such X', there exists an upper cocircuit X for S. Any such cocircuit X satisfies that* VC-dim$(X) = d - 1$, *$e \in \underline{X}$, $\mathrm{Sep}(S, X) = \varnothing$, and $\widehat{\widehat{S}}$ is a full sample of $\mathcal{M}(X)$.*

Proof. Since S is a full sample, $\mathcal{M}' = \mathcal{M} \setminus S^0$ has rank d. Moreover, $S \setminus S^0$ is a tope T' of $\mathcal{M}'$. By Lemma 6, $e \in \mathrm{osc}([S]) = \mathrm{osc}([T'])$ and by Lemma 18 there exists a cocircuit X' of $\mathcal{M}'$ such that $e \in \underline{X'}$, $X' \leq T'$, and $\mathcal{M}(X')$ has VC-dim $d - 1$. Thus X' is a lower cocircuit for S and hence there exists an upper covector X of $\mathcal{M}$ such that $X' = X \setminus S^0$. By Lemma 17, VC-dim$(X) \geq$ VC-dim$(X') = d-1$. If X was not a cocircuit, then $\mathrm{F}(X)$ is a proper face of $\mathrm{F}(Y)$ for some cocircuit Y of $\mathcal{M}$. Since in an OM the VC-dimension of any proper face is strictly smaller than the VC-dimension of the face itself and since $\mathcal{M}$ has VC-dimension d, we obtain a contradiction.

Thus X is a cocircuit of $\mathcal{M}$ (and an upper cocircuit for S) and VC-dim(X) = VC-dim$(X') = d-1$. In particular, $e \in \underline{X}$. By Lemma 16, $\mathrm{Sep}(S, X) = \varnothing$ and $\widehat{\widehat{S}}$ is a full sample of $\mathcal{M}(X)$. □

4. The main result

The goal of this section is to prove the following theorem:

Theorem 3. *The set $\mathcal{T}$ of topes of a complex of oriented matroids $\mathcal{M} = (U, \mathcal{L})$ of VC-dimension d admits a proper labeled sample compression scheme of size d.*

4.1. **The main idea.** Our labelled sample compression scheme takes any realizable sample S of a COM $\mathcal{M}$ and removes the zero set of S. Consequently, S becomes the tope $S \setminus S^0$ of the COM $\mathcal{M} \setminus S^0 =: \mathcal{M}'$. Then we consider a face $\mathrm{F}(X')$ of $\mathcal{M}'$ defined by a minimal covector X' of $\mathcal{M}'$ such that $S \setminus S^0 \geq X'$ (i.e., by a lower covector for S). This face defines the simple OM $\mathcal{M}'(X') = \mathcal{M}' \setminus \underline{X'}$. The compressor $\alpha(S)$ is then defined by applying to $\mathcal{M}'(X')$ and its tope $S \setminus (S^0 \cup \underline{X'})$ the *distinguishing lemma*, which allows to distinguish the full samples of an OM $\mathcal{M}$ of rank d by considering their restriction to subsets of size d. It constructs a function $f_{\mathcal{M}}$ that assigns such a subset to each full sample and is used by both compressor and reconstructor. The *localization lemma* is used by the reconstructor and designates the set of all potential covectors whose faces may contain topes T compatible with the initial sample S. These two lemmas are proved in next two subsections. Compressor and reconstructor are given in the last subsection and are illustrated by Example 2. The compressor generalizes the compressor for ample classes of Moran and Warmuth [57]. However, the reconstructor is more involved than that for ample classes.

4.2. **The distinguishing lemma.** In this subsection, $\mathcal{M} = (U, \mathcal{L})$ is an OM of VC-dimension/rank d. We continue with the definition of the function $f_{\mathcal{M}}$ defined on the set $\mathbf{Samp}_f(\mathcal{M})$ of full samples of $\mathcal{M}$. Fix a linear order on the ground set $U = \{1, \ldots, m\}$ of $\mathcal{M}$. For any subset $U' = U \setminus A$ of U we will consider the restriction of this linear order to U'. Suppose recursively that we have already defined the functions $f_{\mathcal{M}'}$ on the set $\mathbf{Samp}_f(\mathcal{M}')$ of full samples of all proper (i.e., $A \neq \varnothing$) deletions $\mathcal{M}' = (U \setminus A, \mathcal{L} \setminus A)$ of $\mathcal{M}$. Let $S \in \mathbf{Samp}_f(\mathcal{M})$ be a full sample of $\mathcal{M}$. If S is not a tope of $\mathcal{M}$, then we set $f_{\mathcal{M}}(S) = f_{\mathcal{M} \setminus S^0}(S \setminus S^0)$. Otherwise, if S is a tope of $\mathcal{M}$, then we set $f_{\mathcal{M}}(S) = \{e_S, f_{\mathcal{M}(X')}(S \setminus e_S)\}$, where:

- e_S is the smallest element of $\mathrm{osc}([S])$;
- X' is the lexicographically minimal lower cocircuit for S in $\mathcal{M}$ such that $e_S \in \underline{X'}$ and $X' \leq S$.

Equivalently, $f_{\mathcal{M}}(S)$ can be defined recursively by setting $f_{\mathcal{M}}(S) = \{e_S, f_{\mathcal{M}'(X')}(S \setminus (S^0 \cup \underline{X'}))\}$, where $\mathcal{M}' = \mathcal{M} \setminus S^0$ and:

- e_S is the smallest element of $\mathrm{osc}([S \setminus S^0]) = \mathrm{osc}([S])$;
- X' is the lexicographically minimal lower cocircuit for S in $\mathcal{M}'$ such that $e_S \in \underline{X'}$ and $X' \leq S \setminus S^0$.

Remark 2. Here we order sign vectors lexicographically by setting $0 < + < -$. This choice is needed in order to avoid freedom in the definition, but is arbitrary. Indeed, we will prove that taking any lower cocircuit X' for S in $\mathcal{M}'$ such that $e_S \in \underline{X'}$ and $X' \leq S$ will work.

The equality $\mathrm{osc}([S]) = \mathrm{osc}([S \setminus S^0])$ holds by Lemma 6. The cocircuit X' exists by Lemma 18. Since $S \setminus (S^0 \cup \underline{X'})$ is a tope (and thus a full sample) of $\mathcal{M}'(X')$ and since $\mathcal{M}'(X')$ has VC-dimension $d-1$ by Lemma 18, by induction hypothesis $f_{\mathcal{M}'(X')}(S \setminus (S^0 \cup \underline{X'}))$ is well-defined. Furthermore, $f_{\mathcal{M}}(S)$ has size d, thus $f_{\mathcal{M}}$ is a map from $\mathbf{Samp}_f(\mathcal{M})$ to $\binom{U}{d}$.

Now, we define an equivalence relation $\sim$ on the set $\mathbf{Samp}_f(\mathcal{M})$ of all full samples of $\mathcal{M}$:

Definition 6 (Equivalence classes of full samples)**.** Two full samples $S, S' \in \mathbf{Samp}_f(\mathcal{M})$ are *equivalent* (notation $S \sim S'$) if $f_{\mathcal{M}}(S) = f_{\mathcal{M}}(S')$ and $S_{|f_{\mathcal{M}}(S)} = S'_{|f_{\mathcal{M}}(S')}$ hold. Clearly, $\sim$ is an equivalence relation on $\mathbf{Samp}_f(\mathcal{M})$. Denote by $\Omega_1, \ldots, \Omega_k$ the equivalence classes of $\mathbf{Samp}_f(\mathcal{M})$.

The partition of $\mathbf{Samp}_f(\mathcal{M})$ into equivalence classes can be also viewed in the following way. For any set $D \subseteq U$ of size d and any $C \in \{\pm 1, 0\}^U$ with $\underline{C} = D$, we denote by $\Omega(C, D)$ the set of all $S \in \mathbf{Samp}_f(\mathcal{M})$ such that $f_{\mathcal{M}}(S) = D$ and $S_{|f_{\mathcal{M}}(S)} = C$. Then $\Omega(C, D)$ is either empty or is an equivalence class of $(\mathbf{Samp}_f(\mathcal{M}), \sim)$.

We continue with the distinguishing lemma, which shows that $f_{\mathcal{M}}$ distinguishes samples from different equivalence classes of $\sim$ and defines for all samples from the same equivalence class a nonempty convex set, which later in Definition 7 will be called the realizer and will be used by the reconstructor.

Lemma 20. *Let $\mathcal{M} = (U, \mathcal{L})$ be an* OM *of VC-dimension d. The function $f_{\mathcal{M}} : \mathbf{Samp}_f(\mathcal{M}) \to \binom{U}{d}$ has the following properties for all $S \in \mathbf{Samp}_f(\mathcal{M})$:*

- (i) $f_{\mathcal{M}}(S) \subseteq \operatorname{osc}([S])$,
- (ii) $f_{\mathcal{M}}(S)$ *is shattered by* $\mathcal{M}$,
- (iii) *for any equivalence class $\Omega_i, i = 1, \ldots, k$ of $(\mathbf{Samp}_f(\mathcal{M}), \sim)$, $\bigcap_{S \in \Omega_i} [S] \neq \varnothing$.*

Proof. Let $G := G(\mathcal{M})$ be the tope graph of $\mathcal{M}$. We proceed by induction on d. If $d = 1$, then $U = \{e\}$ and G is an edge between the topes $T_1 = (-1)$ and $T_2 = (+1)$, which are the only full samples of $\mathcal{M}$. Then $f_{\mathcal{M}}(T_1) = f_{\mathcal{M}}(T_2) = \{e\}$ and we obtain a function satisfying the conditions (i)-(iii). Thus, let $d \geq 2$.

Condition (i): By definition of $f_{\mathcal{M}}(S)$, the element e_S is chosen from $\operatorname{osc}([S \setminus S^0]) = \operatorname{osc}([S])$. Let $T' = S \setminus S^0$. By induction hypothesis, the remaining elements of $f_{\mathcal{M}}(S)$ will be chosen from $\operatorname{osc}([T' \setminus \underline{X'}])$. Note that $T'' = T' \setminus \underline{X'} = S \setminus (S^0 \cup \underline{X'})$ is a tope of $\mathcal{M}'(X')$ and $\operatorname{osc}([T''])$ is defined by the edges of the tope graph of $\mathcal{M}'(X')$ incident to T''. Since this is a subset of edges incident to T' in the tope graph of $\mathcal{M}'$, we conclude that $\operatorname{osc}([T'']) \subseteq \operatorname{osc}([T']) = \operatorname{osc}([S])$. This proves that $f_{\mathcal{M}}(S) \subseteq \operatorname{osc}([S])$.

Condition (ii): Suppose that $f_{\mathcal{M}}(S)$ is not shattered by $\mathcal{M}$. Define $D' = f_{\mathcal{M}'(X')}(T' \setminus \underline{X'})$, where $\mathcal{M}' = \mathcal{M} \setminus S^0$, $T' = S \setminus S^0$, and X' is any cocircuit of $\mathcal{M}'$ such that $e_S \in \underline{X'}$ and $X' \leq T'$, which exists by Lemma 19. By the induction hypothesis, D' is shattered by $\mathcal{M}'(X')$. By Lemma 19, there exists a cocircuit X of $\mathcal{M}$ such that $X \setminus S^0 = X'$ and $e_S \in \underline{X'}$. Since D' is shattered by $\mathcal{M}'(X')$, we get $D' \subseteq X'^0 \subseteq X^0$. Since $f_{\mathcal{M}}(S) = D' \cup \{e_S\}$ is not shattered by $\mathcal{M}$, by Lemma 8 there is a circuit Y of $\mathcal{M}$ such that $\underline{Y} \subseteq \{e_S\} \cup D'$ and $e_S \in \underline{Y}$. On the other hand, $D' \subseteq X^0$ and $e_S \in \underline{X}$, thus $|\underline{Y} \cap \underline{X}| = 1$. Since X is a cocircuit and Y is a circuit, this contradicts orthogonality of circuits and cocircuits in OMs, see Theorem 1.

Condition (iii): The case $d = 1$ was considered above, so let $d \geq 2$. Suppose that $\Omega_i = \Omega(C, D)$ for some $C \in \{\pm 1, 0\}^U$ and $D = \underline{C}$. Let Q, R be any two full samples of $\Omega(C, D)$ and denote $\mathcal{M}' = \mathcal{M} \setminus Q^0$ and $\mathcal{M}'' = \mathcal{M} \setminus R^0$. Thus $f_{\mathcal{M}}(Q) = f_{\mathcal{M}}(R) = D$ and $Q_{|f_{\mathcal{M}}(Q)} = R_{|f_{\mathcal{M}}(R)} = C$. By definition, $f_{\mathcal{M}}(Q) = \{e_Q, f_{\mathcal{M}'(X'_Q)}(Q \setminus (Q^0 \cup \underline{X'_Q}))\}$, where e_Q is the smallest element of $\operatorname{osc}([Q \setminus Q^0]) = \operatorname{osc}([Q])$ and X'_Q is a lower cocircuit for Q such that $e_Q \in \underline{X'_Q}$ and $X'_Q \leq Q \setminus Q^0$. Analogously, $f_{\mathcal{M}}(R) = \{e_R, f_{\mathcal{M}''(X'_R)}(R \setminus (R^0 \cup \underline{X'_R}))\}$, where e_R is the smallest element of $\operatorname{osc}([R \setminus R^0]) = \operatorname{osc}([R])$ and X'_R is a lower cocircuit for R such that $e_R \in \underline{X'_R}$ and $X'_R \leq R \setminus R^0$. Since $f_{\mathcal{M}}(Q) = f_{\mathcal{M}}(R)$, by the minimality in the choice of the elements e_Q and e_R, both are the smallest elements of the respective sets $f_{\mathcal{M}}(Q)$ and $f_{\mathcal{M}}(R)$. Consequently, $e_Q = e_R =: e$ and $D = \{e\} \cup D'$, where $f_{\mathcal{M}'(X'_Q)}(Q \setminus (Q^0 \cup \underline{X'_Q})) = f_{\mathcal{M}''(X'_R)}(R \setminus (R^0 \cup \underline{X'_R})) =: D'$.

By Lemma 19, there exists an upper cocircuit X_Q of $\mathcal{M}$ such that $X_Q \setminus Q^0 = X'_Q$, $e \in \underline{X_Q}$, and VC-dim$(X_Q) = d-1$. Analogously, there exists an upper cocircuit X_R of $\mathcal{M}$ such that $X_R \setminus R^0 = X'_R$, $e \in \underline{X_R}$, and VC-dim$(X_R) = d-1$. Furthermore, by the same Lemma 19 and by Lemma 14, we have $[Q] \cap [X_Q] \neq \varnothing$ and $[R] \cap [X_R] \neq \varnothing$. Since both faces $\mathrm{F}(X_Q) \cong \mathcal{M}(X_Q)$ and $\mathrm{F}(X_R) \cong \mathcal{M}(X_R)$ of $\mathcal{M}$ shatter the same set $D' \subseteq U$, Lemma 11 implies that $X_Q = X_R$ or $X_Q = -X_R$. Indeed, let $X_Q \neq \pm X_R$. Since X_Q, X_R maximally shatter D', by Lemma 11(ii) $X_Q = X_Q \circ X_R$ and $X_R = X_R \circ X_Q$. By Lemma 11(iii) there exists a geodesic gallery between $\mathrm{F}(X_Q)$ and $\mathrm{F}(X_R)$. Since X_Q and X_R are cocircuits of $\mathcal{M}$, $\mathrm{F}(X_Q)$ and $\mathrm{F}(X_R)$ are facets of $\mathcal{M}$. Therefore $\mathrm{F}(X_Q)$ and $\mathrm{F}(X_R)$ must be consecutive in the gallery and the face containing them as facets must coincide with $\mathcal{M}$. Thus, $X_Q = \pm X_R$.

But if $X_Q = -X_R$ holds, since $e \in \underline{X_Q} \cap \underline{X_R}$, we have $e \in \mathrm{Sep}(X_Q, X_R)$. Since $[Q] \cap [X_Q] \neq \varnothing$ and $[R] \cap [X_R] \neq \varnothing$, for any two topes $T' \in [Q] \cap [X_Q]$ and $T'' \in [R] \cap [X_R]$ we will have $T'_e = -T''_e$. Since $e \in \mathrm{osc}([Q]) \cap \mathrm{osc}([R])$, we get $Q_e = -R_e$, which contradicts the assumption $Q_{|f_{\mathcal{M}}(Q)} = R_{|f_{\mathcal{M}}(R)}$. Hence, $X_Q = X_R$. Since the equality $X_Q = X_R$ holds for any $Q, R \in \Omega(C, D)$, there exists a cocircuit X of $\mathcal{M}$ such that for any $S \in \Omega(C, D)$, we have $X \setminus S^0 = X'_S$, $e \in \underline{X'_S}$, VC-dim$(X) = d-1$, and $[S] \cap [X] \neq \varnothing$.

By the induction hypothesis, the function $f_{\mathcal{M}(X)}$ defined on the set $\mathbf{Samp}_f(\mathcal{M}(X))$ of full samples of $\mathcal{M}(X)$ satisfies the properties (i)-(iii) of the lemma. Let C' denote the restriction of C to D'. Denote by $\Omega'(C', D')$ the set of all $Q' \in \mathbf{Samp}_f(\mathcal{M}(X))$ such that $f_{\mathcal{M}(X)}(Q') = D'$ and $Q'_{|D'} = C'$. For any $Q \in \Omega(C, D)$, we have $[Q] \cap [X] \neq \varnothing$, thus $\mathrm{Sep}(X, Q) = \varnothing$. By Lemma 14, $[\widehat{Q}] = [Q] \cap [X] \neq \varnothing$. By the same lemma, $\widehat{\widehat{Q}} \in \mathbf{Samp}(\mathcal{M}(X))$ and $[\widehat{\widehat{Q}}\,]$ is U-isomorphic to $[\widehat{Q}]$. By Lemma 19, $\widehat{\widehat{Q}}$ is a full sample of $\mathcal{M}(X)$, i.e., $\widehat{\widehat{Q}} \in \mathbf{Samp}_f(\mathcal{M}(X))$. We assert that $\widehat{\widehat{Q}} \in \Omega'(C', D')$. Recall that X is an upper cocircuit for Q and X'_Q is a lower cocircuit for Q such that $X \setminus Q^0 = X'_Q$. By Lemma 17, $\mathcal{M}'(X'_Q) = \mathcal{M}(X) \setminus \widehat{Q}^0 = \mathcal{M}(X) \setminus \widehat{\widehat{Q}}\,^0$. This implies that $f_{\mathcal{M}(X)}(\widehat{\widehat{Q}}\,) = f_{\mathcal{M}'(X'_Q)}(Q \setminus (Q^0 \cup \underline{X'_Q})) = D'$ and consequently that $\widehat{\widehat{Q}}_{|D'} = Q_{|D'} = C'$. This establishes the inclusion $\{\widehat{\widehat{Q}} : Q \in \Omega(C, D)\} \subseteq \Omega'(C', D')$. Since $\Omega(C, D) = \Omega_i \neq \varnothing$, the set $\Omega'(C', D')$ is nonempty and thus is an equivalence class of $(\mathbf{Samp}_f(\mathcal{M}(X)), \sim)$. By the induction hypothesis, in $G(\mathcal{M}(X))$ we have $\bigcap_{Q' \in \Omega'(C', D')} [Q'] \neq \varnothing$. Denote this intersection by $\mathcal{R}'(C', R')$.

Let $\mathcal{R}(C, R)$ denotes the (nonempty) set of topes T of $G(\mathcal{M})$ of the form $T = T' \times X_{|\underline{X}}$ for some tope T' of $\mathcal{M}(X)$ belonging to the set $\mathcal{R}'(C', R')$. By the one-to-one correspondence φ_X between the topes of $[X]$ and the topes of $G(\mathcal{M}(X))$ we conclude that $\mathcal{R}(C, R) \subseteq [X]$. Pick any sample $Q \in \Omega(C, D)$. Since $\widehat{\widehat{Q}} \in \Omega'(C', D')$, we get $T' \in \mathcal{R}'(C', R') \subseteq [\widehat{\widehat{Q}}\,]$ (recall that $[\widehat{\widehat{Q}}\,]$ is considered in $G(\mathcal{M}(X))$). By the U-isomorphism between the convex subgraphs $[\widehat{Q}]$ and $[\widehat{\widehat{Q}}\,]$ (Lemma 14), we deduce that the tope $T = T' \times X_{|\underline{X}}$ belongs in $G(\mathcal{M})$ to $[\widehat{Q}]$. Consequently, the inclusion $\mathcal{R}(C, R) \subseteq [\widehat{Q}] \cap [X]$ holds for any $Q \in \Omega(C, D)$. Since for any $Q \in \Omega(C, D)$, $[\widehat{Q}] = [Q] \cap [X]$ by Lemma 14, we conclude that $(\bigcap_{Q \in \Omega(C,D)} [Q]) \cap [X] = \bigcap_{Q \in \Omega(C,D)} ([Q] \cap [X]) \supseteq \mathcal{R}(C, R) \neq \varnothing$. Consequently, $\bigcap_{Q \in \Omega(C,D)} [Q] \neq \varnothing$. This concludes the proof of property (iii) and of the lemma. □

Definition 7 (Realizers)**.** For an equivalence class $\Omega_i = \Omega(C, D)$ of $(\mathbf{Samp}_f(\mathcal{L}), \sim)$, we call the nonempty intersection $\mathcal{R}(C, D) = \bigcap_{S \in \Omega_i} [S]$ the *realizer* of $\Omega(C, D)$.

4.3. **The localization lemma.** The *localization lemma* designates for any realizable sample S of a COM $\mathcal{M}$ the set of all potential covectors whose faces may contain topes of $\mathcal{M}$ which can be used by the reconstructor.

Let $\mathcal{M} = (U, \mathcal{L})$ be a COM of VC-dimension d and let $S \in \mathbf{Samp}(\mathcal{M})$ be a realizable sample. Consider the tope $T' = S \setminus S^0$ of the COM $\mathcal{M}' := \mathcal{M} \setminus S^0$ and let X' be a minimal covector of $\mathcal{M}'$ such that $T' \geq X'$. By Lemma 5, the OM $\mathcal{M}'(X') = \mathcal{M}' \setminus \underline{X'}$ has VC-dimension $\leq d$. Let

$$\mathcal{H}_{S,X'} := \{X \in \mathcal{L} : X \setminus S^0 = X' \text{ and VC-dim}(\mathcal{M}(X)) = \text{VC-dim}(\mathcal{M}'(X'))\}.$$

For a set $D \subseteq U$, let
$$\mathcal{H}_D := \{X \in \mathcal{L} : \mathcal{M}(X) \text{ maximally shatters } D\}.$$

Lemma 21. *Let $S \in \mathbf{Samp}(\mathcal{M})$, X' be a minimal covector of $\mathcal{M}' = \mathcal{M} \setminus S^0$ such that $S \setminus S^0 = T' \geq X'$, and let D be a subset of $\underline{S} = U \setminus S^0$ such that $|D| = \text{VC-dim}(\mathcal{M}'(X'))$ and D is shattered by $\mathcal{M}'(X')$. Then $\varnothing \neq \mathcal{H}_{S,X'} = \mathcal{H}_D$.*

Proof. First, we prove that $\mathcal{H}_{S,X'} \subseteq \mathcal{H}_D$. Pick any $X \in \mathcal{H}_{S,X'}$. Since $\mathcal{M}'(X')$ shatters D and $G(\mathcal{M}'(X'))$ is a pc-minor of $G(\mathcal{M}(X))$ because $X \setminus S^0 = X'$, $\mathcal{M}(X)$ also shatters D. Since $\text{VC-dim}(\mathcal{M}(X)) = \text{VC-dim}(\mathcal{M}'(X'))$, $\mathcal{M}(X)$ maximally shatters D, yielding $X \in \mathcal{H}_D$.

Now we prove that the set $\mathcal{H}_{S,X'}$ is nonempty. By Lemma 1 there exists at least one covector $X \in \mathcal{L}$ such that $X \setminus S^0 = X'$. For the same reason as above, $\mathcal{M}(X)$ shatters D. Suppose that $\mathcal{M}(X)$ shatters a superset of D. By Lemma 11(iii), there exists a covector $Y > X$ of $\mathcal{M}$ such that $\mathcal{M}(Y)$ maximally shatters D. Hence, $Y \setminus S^0 \geq X \setminus S^0 = X'$, but $\mathcal{M}'(Y \setminus S^0)$ and $\mathcal{M}'(X')$ have the same VC-dimension since they both maximally shatter the set D. By Lemma 5, $Y \setminus S^0 = X'$ and hence $Y \in \mathcal{H}_{S,X'}$. This proves that $\mathcal{H}_{S,X'} \neq \varnothing$.

It remains to prove that $\mathcal{H}_D \subseteq \mathcal{H}_{S,X'}$. Assume by way of contradiction that there exists $Y \in \mathcal{H}_D \setminus \mathcal{H}_{S,X'}$ and set $Y' = Y \setminus S^0$. Since $Y \notin \mathcal{H}_{S,X'}$ and $\mathcal{M}(Y)$ maximally shatters D, we have $X' \neq Y'$. Since $\mathcal{M}(Y)$ maximally shatters D and $D \subseteq \underline{S}$, also $\mathcal{M}'(Y')$ maximally shatters D. In particular, $D \subseteq X'^0 \cap Y'^0 = (X' \circ Y')^0$. By Lemma 10 the gates of $[Y']$ in $[X']$ are the topes of $\text{F}(X' \circ Y') \subseteq \text{F}(X')$. Thus, $[X' \circ Y']$ is a gated subgraph of $[X']$, and $[X' \circ Y']$ is crossed by D (since $D \subseteq (X' \circ Y')^0 = \text{cross}([X' \circ Y'])$), and D is shattered by $[X']$. By Lemma 7, the VC-dimension of $\mathcal{M}'(X' \circ Y')$ is at least $|D|$, which is the VC-dimension of $\mathcal{M}'(X')$. Then Lemmas 11(i) and 5 yield $X' \circ Y' = X'$. If $\text{Sep}(X', Y') = \varnothing$, then $\text{F}(X') = \text{F}(X' \circ Y') \subseteq \text{F}(Y')$. Since $\text{F}(X')$ is a maximal face of $\mathcal{M}'$, we get $X' = Y'$. Otherwise, if $\text{Sep}(X', Y') \neq \varnothing$, then by Lemmas 11(iii) and 10 there exists a geodesic gallery $(\text{F}(X') = \text{F}(X_0), \text{F}(X_1), \ldots, \text{F}(X_k) = \text{F}(Y'))$ with $k > 0$ from $\text{F}(X')$ to $\text{F}(Y')$ in $\mathcal{M}'$. By the definition of a gallery, the union of $\text{F}(X')$ and $\text{F}(X_1)$ is included in a face $\text{F}(Z) \supsetneq \text{F}(X')$ of $\mathcal{M}'$. Thus, $\text{F}(X')$ is not a maximal face of $\mathcal{M}'$, contradicting the assumption that X' is a minimal covector of $\mathcal{M}'$. □

4.4. **The labeled compression scheme.** Now, we describe the compression and the reconstruction and prove their correctness. The compression map generalizes the compression map for ample classes of [57]. However, the reconstruction map is much more involved than the reconstruction map for ample classes, since it uses both the distinguishing and the localization lemma.

Compression. Let $\mathcal{M} = (U, \mathcal{L})$ be a COM of VC-dimension d. For a sample $S \in \mathbf{Samp}(\mathcal{M})$ of $\mathcal{M}$, consider the tope $T' = S \setminus S^0$ of $\mathcal{M}' := \mathcal{M} \setminus S^0$ and let X' be the lexicographically minimal lower circuit for S, i.e., the lexicographically minimal support-minimal covector of $\mathcal{M}'$ such that $T' \geq X'$. Denote by $\mathcal{M}'(X') = \mathcal{M}' \setminus \underline{X'}$ the simple OM defined by the face $\text{F}(X')$ of $\mathcal{M}'$. Define $\alpha(S)_e = S_e$ if $e \in f_{\mathcal{M}'(X')}(T')$ and $\alpha(S)_e = 0$ otherwise. The map α is well-defined since T' is a tope of $\mathcal{M}'(X')$ and hence the sample T' is full in $\mathcal{M}'$. Moreover, by definition we have $\alpha(S) \leq S$, whence $\alpha(S) \in \mathbf{Samp}(\mathcal{M})$. Finally, by Lemma 5 the OM $\mathcal{M}'(X')$ has VC-dimension at most d and thus, by Lemma 20 $\alpha(S)$ has support of size $\leq d$.

Reconstruction. To define β, pick any $C \in \{\pm 1, 0\}^U$ in the image $\text{Im}(\alpha)$ of α and let $D := \underline{C}$. Let X be any covector from $\mathcal{H}_D$, i.e., X is a covector of $\mathcal{L}$ that maximally shatters D. By Lemma 21, X exists. Let $\Omega(C, D)$ be the set of all full samples $Q \in \mathbf{Samp}_f(\mathcal{M}(X))$ of the OM $\mathcal{M}(X)$ such that $f_{\mathcal{M}(X)}(Q) = D$ and $Q_{|f_{\mathcal{M}(X)}(Q)} = C$. Lemma 22 below shows that $\Omega(C, D)$ is nonempty. Thus $\Omega(C, D)$ is an equivalence class of $(\mathbf{Samp}_f(\mathcal{M}(X)), \sim)$. By Lemma 20(iii), the realizer $\mathcal{R}(C, D) = \bigcap_{Q \in \Omega(C,D)} [Q]$ of $\Omega(C, D)$ is a nonempty convex subgraph of $G(\mathcal{M}(X))$. Then, let $\beta(C)$ be any tope $\widetilde{T}$ of $\mathcal{M}$ of the form $\widetilde{T} = \widetilde{T}_0 \times X_{|\underline{X}}$, where $\widetilde{T}_0$ is any tope from $\mathcal{R}(C, D)$.

Correctness. We prove that (α,β) defines a proper labeled sample compression scheme, namely, we show that for all samples $S\in\mathbf{Samp}(\mathcal{M})$, we have (1) $\alpha(S)\le S$ and $\alpha(S)$ has support of size $\le d$ and (2) $\beta(\alpha(S))$ is well-defined and $\beta(\alpha(S))\ge S$. The assertion (1) has been already established. Let $C=\alpha(S)$ and $D=\underline{C}$. To prove that β is well-defined, we have to show that $\Omega(C,D)$ is nonempty. This follows from the following result:

Lemma 22. $\widehat{\widehat{S}}=\widehat{S}\setminus\underline{X}=(X\circ S)\setminus\underline{X}$ *belongs to* $\Omega(C,D)$.

Proof. By Lemma 14, $\widehat{\widehat{S}}\in\mathbf{Samp}(\mathcal{M}(X))$. Since $X\in\mathcal{H}_D$, by Lemma 21, X satisfies $X\setminus S^0=X'$, where X' is the minimal covector of $\mathcal{M}'=\mathcal{M}\setminus S^0$ chosen in the definition of $\alpha(S)$. Since $X\setminus S^0=X'\le T'=S\setminus S^0$, we have $\mathrm{Sep}(X,S)=\varnothing$. By Lemma 14 $[\widehat{S}]=[X]\cap[S]$ is a nonempty convex subgraph of $[X]$ and $[\widehat{\widehat{S}}]$ is U-isomorphic to $[\widehat{S}]$. Since $X\setminus\widehat{S}^0=X'$ and both $\mathcal{M}(X),\mathcal{M}'(X')$ have the same VC-dimension $|D|$, $\widehat{\widehat{S}}$ is a full sample of $\mathcal{M}(X)$ by the last assertion of Lemma 16.

By Lemma 17, $\mathcal{M}'(X')=\mathcal{M}(X)\setminus\widehat{S}^0=\mathcal{M}(X)\setminus\widehat{\widehat{S}}^0$. By definition of α and $f_{\mathcal{M}(X)}$, we have $\alpha(S)=D=f_{\mathcal{M}'(X')}(S)=f_{\mathcal{M}(X)}(\widehat{\widehat{S}})$. It remains to show that $\widehat{\widehat{S}}_{|D}=C_{|D}$. Pick any $e\in D$. Since $C_e\neq 0$ and $C=\alpha(S)\le S$, we get $S_e=C_e$. Since $D\subseteq X^0$ and $\widehat{\widehat{S}}=(X\circ S)\setminus\underline{X}$, we conclude that $\widehat{\widehat{S}}_e=C_e$, establishing the equality $\widehat{\widehat{S}}_{|D}=C_{|D}$. This shows that $\widehat{\widehat{S}}$ indeed belongs to $\Omega(C,D)$. $\square$

It remains to prove that $\beta(\alpha(S))\ge S$. Since Lemma 14 implies $[\widehat{S}]=[X]\cap[S]$, we conclude that $\mathrm{Sep}(X,S)=\varnothing$ and consequently that $\widehat{S}=X\circ S\ge S$ holds. By definition, $\beta(\alpha(S))=\beta(C)$ is any tope of the form $\widetilde{T}=\widetilde{T}_0\times X_{|\underline{X}}$ for a tope $\widetilde{T}_0$ of $\mathcal{M}(X)$ belonging to the realizer $\mathcal{R}(C,D)=\bigcap_{Q\in\Omega(C,D)}[Q]$. Since by Lemma 22, the sample $\widehat{\widehat{S}}$ belongs to $\Omega(C,D)$, the realizer $\mathcal{R}(C,D)$ is included in $[\widehat{\widehat{S}}]$. Consequently, $\widetilde{T}_0\ge\widehat{\widehat{S}}$. Since $\widehat{S}=\widehat{\widehat{S}}\times X_{|\underline{X}}$ and $\widetilde{T}=\widetilde{T}_0\times X_{|\underline{X}}$, we deduce that $\widetilde{T}\ge\widehat{S}$. Since $\widehat{S}\ge S$, we obtain $\beta(\alpha(S))=\widetilde{T}\ge S$. This concludes the proof of Theorem 3, the main result of the paper.

Remark 3. Note that by Lemma 22 any tope $T\ge\widehat{S}$ or any tope of the form $\widetilde{T}=\widetilde{T}_0\times X_{|\underline{X}}$ for a tope $\widetilde{T}_0\ge\widehat{\widehat{S}}$ would be feasible. However, S and henceforth $\widehat{S}$ and $\widehat{\widehat{S}}$ are not known to the reconstructor. Thus, we have to rely on the realizer $\mathcal{R}(C,D)\subseteq[\widehat{\widehat{S}}]$.

We conclude this section with two examples illustrating our compression scheme:

Example 2. Consider the tope graph G of a COM $\mathcal{M}$ of VC-dimension 3 and a realizable sample $S=(++-0-0+0)$ in Figure 5. $[S]$ is induced by 7 topes drawn as white vertices of G. Contracting the 3 dashed Θ-classes corresponding to $\{4,6,8\}=S^0$, yields the tope graph G' of $\mathcal{M}'=\mathcal{M}\setminus S^0$. Then $T'=S\setminus S^0=(++--+)$. The compressor picks $X'=(0+--+)$, the lexicographically minimal lower circuit for S; X' corresponds to the thick red edge in G', and in covector representation $\mathcal{M}'(X')=(\{1\},\{(0),(+),(-)\})$. The compressor returns $\alpha(S)=(+0000000)$ and $D=\{1\}$. The reconstructor receives $C=(+0000000)=\alpha(S)$, defines $D=\underline{C}=\{1\}$ and constructs the set $\mathcal{H}_D$. There are six covectors of $\mathcal{M}$ belonging to $\mathcal{H}_D$ corresponding to the thick red edges in G. By the localization lemma, they are the covectors which have the same VC-dimension as X' and agree with X' on $\{1,2,3,5,7\}=\underline{S}$. The reconstructor picks an arbitrary covector from $\mathcal{H}_D$, say $X=(0+----+-)$. The OM $\mathcal{M}(X)$ is composed of the covectors X and the ends T_1 and T_2 of the corresponding red edge. Then, we get $\Omega(C,D)=\{T_1\}$ and its realizer is $\mathcal{R}(C,D)=[T_1]$. Thus, $\beta(\alpha(S))$ is set to T_1, which is a white vertex of G.

The previous example might suggest that indeed $f_{\mathcal{M}}(S)=f_{\mathcal{M}}(S')$ and $S_{|f_{\mathcal{M}}(S)}=S'_{|f_{\mathcal{M}}(S')}$ together imply $[S]=[S']$. However, the next example shows that $[S]$ and $[S']$ might not even be contained in each other.

Example 3. Let $\mathcal{M}$ be the COM whose tope graph consists of a 4-cycle C with two edges pending on the same vertex T. Let $1,2$ be the Θ-classes of C and $3,4$ the Θ-classes of the other two edges.

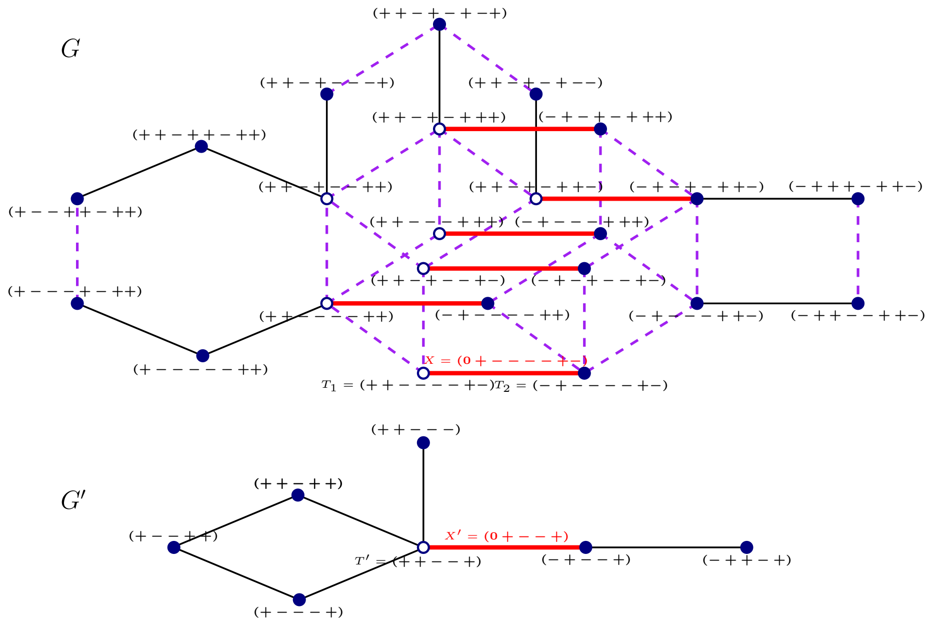

FIGURE 5. An illustration of Example 2.

Each of the two pending edges corresponds to a full sample $S = (+ + +0)$ and $S' = (+ + 0+)$ respectively. It is easy to see that $f_{\mathcal{M}}(S) = f_{\mathcal{M}}(S') = \{1, 2\}$ and $S_{|f_{\mathcal{M}}(S)} = S'_{|f_{\mathcal{M}}(S')} = (++)$ but $[S] \cap [S'] = \{T\}$. Further note that the tope graph of $\mathcal{M}$ can be easily embedded into a tope graph of a uniform OM $\mathcal{M}'$ of rank 3 in which C is a cocircuit, the samples S, S' encode the two pending edges (with possibly larger support) and still $f_{\mathcal{M}'}(S) = f_{\mathcal{M}'}(S')$ and $S_{|f_{\mathcal{M}'}(S)} = S'_{|f_{\mathcal{M}'}(S')}$ while $[S] \cap [S'] = \{T\}$ is a proper subset of both $[S]$ and $[S']$.

5. Conclusion

We have presented proper labeled compression schemes of size d for COMs of VC-dimension d. This is a generalization of the results of [57] for ample set systems, of [7] for affine arrangements of hyperplanes, and of our result [19] for complexes of uniform oriented matroids. Even though we made strong use of the structure of COMs, it is tempting to extend our approach to other classes, e.g., bouquets of oriented matroids [22], strong elimination systems [4], or CW left-regular-bands [54]. Our treatment of realizable samples as convex subgraphs suggests an angle at general partial cubes.

Our results together with the approach of [18, 19] suggest a new approach at *improper* labeled compression schemes of COMs. For this one needs to answer the question: Is it possible to extend a given set system or a partial cube to a COM without increasing the VC-dimension too much?

In unlabeled sample compression schemes, the compressor α is less expressive since its image is in 2^U and has to satisfy $\alpha(S) \subseteq \underline{S}$. Unlabeled compression schemes exist for realizable affine oriented matroids [7] and ample set systems with corner peelings [11, 49]. Recently, Marc [53] designed unlabeled sample compression schemes for OMs. His construction uses Oriented Matroid Programming and Lemma 20. Moreover, he shows there are unlabeled compression schemes for COMs with corner peelings – a recent notion introduced in [48].

Acknowledgments

We acknowledge the referees for their careful reading of the manuscript and useful suggestions and comments. Especially, we are very grateful to one referee who found a critical error in a previous version of the distinguishing lemma. We are also grateful to them for a new proof of Lemma 8 and to Emeric Gioan for finding a gap in a previous proof of this lemma. This research was supported by ANR project DISTANCIA (ANR-17-CE40-0015).

References

[1] M. Albenque and K. Knauer, Convexity in partial cubes: the hull number, Discr. Math. 339 (2016), 866–876.
[2] H.-J. Bandelt, Graphs with intrinsic S_3 convexities, J. Graph Th. 13 (1989), 215—238.
[3] H.-J. Bandelt, V. Chepoi, A. Dress, and J. Koolen, Combinatorics of lopsided sets, Eur. J. Comb. 27 (2006), 669–689.
[4] H.-J. Bandelt, V. Chepoi, and K. Knauer, COMs: complexes of oriented matroids, J. Comb. Th. Ser. A 156 (2018), 195–237.
[5] P.L. Bartlett, N. Harvey, C. Liaw, and A. Mehrabian, Nearly-tight VC-dimension and pseudodimension bounds for piecewise linear neural networks, JMLR 20 (2019), 1–17.
[6] A. Baum and Y. Zhu, The axiomatization of affine oriented matroids reassessed, J. Geom. 109, 11 (2018).
[7] S. Ben-David and A. Litman, Combinatorial variability of Vapnik-Chervonenkis classes with applications to sample compression schemes, Discr. Appl. Math. 86 (1998), 3–25.
[8] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler, Oriented Matroids, Encyclopedia of Mathematics and its Applications, vol. 46, Cambridge University Press, 1993.
[9] R. Bland, M. Las Vergnas, Orientability of matroids, J. Comb. Th. Ser. B 23 (1978), 94–123.
[10] B. Bollobás and A.J. Radcliffe, Defect Sauer results, J. Comb. Th. Ser. A 72 (1995), 189–208.
[11] J. Chalopin, V. Chepoi, S. Moran, and M.K. Warmuth, Unlabeled sample compression schemes and corner peelings for ample and maximum classes, J. Comput. Syst. Sci. 127 (2022), 1–28.
[12] J. Chalopin, V. Chepoi, F. Mc Inerney, S. Ratel, and Y. Vaxès, Sample compression schemes for balls in graphs, SIAM J. Discr. Math. (to appear).
[13] V. Chepoi, d-Convex sets in graphs, Dissertation, Moldova State Univ., Chişinău, 1986.
[14] V. Chepoi, Isometric subgraphs of Hamming graphs and d-convexity, Cybernetics 24 (1988), 6–10.
[15] V. Chepoi, Separation of two convex sets in convexity structures, J. Geometry 50 (1994), 30–51.
[16] V. Chepoi, Distance-preserving subgraphs of Johnson graphs, Combinatorica 37 (2017), 1039–1055.
[17] V. Chepoi, K. Knauer, and T. Marc, Hypercellular graphs: Partial cubes without Q_3^- as partial cube minor, Discr. Math. 343 (2020), 111678.
[18] V. Chepoi, K. Knauer, and M. Philibert, Two-dimensional partial cubes, Electr. J. Comb. 27 (2020), p3.29.
[19] V. Chepoi, K. Knauer, and M. Philibert, Ample completions of Oriented Matroids and Complexes of Uniform Oriented Matroids, SIAM J. Discr. Math. 36 (2022), 505–535.
[20] I. da Silva, Axioms for maximal vectors of an oriented matroid: a combinatorial characterization of the regions determined by an arrangement of pseudohyperplanes, Eur. J. Comb. 16 (1995), 125–145.
[21] P. Deligne, Les immeubles des groupes de tresses généralisés, Invent. Math. 17 (1972), 273–302.
[22] M. Deza, K. Fukuda, On bouquets of matroids and orientations, Kokyuroku 587 (1986), 110–129.
[23] E. Delucchi and K. Knauer, Finitary affine oriented matroids. arXiv:2011.13348, (2020).
[24] R. DeVore, B. Hanin, and G. Petrova, Neural network approximation, Acta Numer. 30 (2021), 327–444.
[25] D.Ž. Djoković, Distance–preserving subgraphs of hypercubes, J. Comb. Th. Ser. B 14 (1973), 263–267.
[26] G. Dorpalen-Barry, J. S. Kim, V. Reiner, Whitney numbers for poset cones. Order (2021), 1–40.
[27] A.W.M. Dress, Towards a theory of holistic clustering, DIMACS Ser. Discr. math. Theoret. Comput.Sci., 37 Amer. Math. Soc. 1997, pp. 271–289.
[28] A.W.M. Dress, R. Scharlau, Gated sets in metric spaces, Aequationes Math. 34 (1987), 112–120.
[29] P.H. Edelman, R.E. Jamison, The theory of convex geometries, Geom. Dedicata 19 (1985), 247–270.
[30] J. Edmonds and A. Mandel, Topology of Oriented Matroids, PhD Thesis of A. Mandel, University of Waterloo, 1982, 333 pages.
[31] D. Eppstein, Recognizing partial cubes in quadratic time, SODA 2008, pp. 1258–1266.
[32] D. Eppstein, Isometric diamond subgraphs, In: International Symposium on Graph Drawing, pp. 384-389. Springer, Berlin, Heidelberg, 2008.
[33] D. Eppstein, J.-C. Falmagne, and S. Ovchinnikov, Media theory: interdisciplinary applied mathematics, Springer Science & Business Media, 2007.
[34] S. Floyd, On space bounded learning and the Vapnik-Chervonenkis dimension, PhD Thesis, International Computer Science Institut, Berkeley, CA, 1989.
[35] S. Floyd and M.K. Warmuth, Sample compression, learnability, and the Vapnik-Chervonenkis dimension, Mach. Learn. 21 (1995), 269–304.
[36] J. Folkman and J. Lawrence, Oriented matroids, J. Comb. Th. Ser. B 25 (1978), 199–236.

[37] B. Gärtner and E. Welzl, Vapnik-Chervonenkis dimension and (pseudo-)hyperplane arrangements, Discr. Comput. Geom. 12 (1994), 399–432.
[38] R. L. Graham, and H. O. Pollak. On the addressing problem for loop switching. The Bell system technical journal, 50 (1971), 2495–2519.
[39] J.E. Grisby and K. Lindsey, On transversality of bent hyperplane arrangements and the topological expressiveness of ReLU neural networks, SIAM J. Appl. Algebra and Geometry, 6 (2022), 216–242.
[40] F. Haglund and F. Paulin, Simplicité de groupes d'automorphismes d'espaces à courbure négative, The Epstein birthday schrift, Geom. Topol. Monogr., 1 (1998), 181–248. Geom. Topol. Publ., Coventry.
[41] B. Hanin and D. Rolnick, Complexity of linear regions in deep networks, ICML 2019, pp. 2596–2604.
[42] B. Hanin and D. Rolnick, Deep ReLU networks have surprisingly few activation patterns, NeurIPS 2019, pp. 359–368.
[43] D. Helmbold, R. Sloan, and M.K. Warmuth, Learning nested differences of intersection-closed concept classes, Mach. Learn. 5 (1990), 165–196.
[44] W. Hochstättler, S. Keip, and K. Knauer, The signed Varchenko Determinant for Complexes of Oriented Matroids, arXiv:2211.13986, (2022).
[45] W. Hochstättler and V. Welker, The Varchenko determinant for oriented matroids, Math. Z. 293 (2019), 1415–1430.
[46] V. Itskov, A. Kunin, and Z. Rosen. Hyperplane neural codes and the polar complex. In Topological Data Analysis, pp. 343-369. Springer, Cham, 2020.
[47] K. Knauer and T. Marc, On tope graphs of complexes of oriented matroids, Discr. Comput. Geom. 63 (2020), 377–417.
[48] K. Knauer and T. Marc, Corners and simpliciality in oriented matroids and partial cubes, Eur. J. Comb. 112 (2023), 103714.
[49] D. Kuzmin and M.K. Warmuth. Unlabeled compression schemes for maximum classes, JMLR 8 no. 9 (2007).
[50] J. Lawrence, Lopsided sets and orthant-intersection of convex sets, Pac. J. Math. 104 (1983), 155–173.
[51] N. Littlestone, M.K. Warmuth, Relating data compression and learnability, Unpublished, 1986.
[52] T. Marc, Mirror graphs: Graph theoretical characterization of reflection arrangements and finite Coxeter groups, Eur. J. Comb. 63 (2017), 115–123.
[53] T. Marc, Unlabeled sample compression schemes for oriented matroids, arXiv:2203.11535, 2022.
[54] S. Margolis, F. Saliola, and B. Steinberg, Cell complexes, poset topology and the representation theory of algebras arising in algebraic combinatorics and discrete geometry, Mem. Amer. Math. Soc. 274 (2021), no 1345.
[55] S. Moran. Shattering Extremal Systems. Master's thesis, Saarland University, Saarbrücken, Germany, 2012.
[56] S. Moran, Personal communication, 2021.
[57] S. Moran, M.K. Warmuth, Labeled compression schemes for extremal classes, ALT 2016, 34–49.
[58] S. Moran, A. Yehudayoff, Sample compression schemes for VC classes, J. ACM 63 (2016), 1–21.
[59] T. Mészáros and L. Rónyai, Shattering-extremal set systems of VC dimension at most 2, Electr. J. Comb. 21 (2014), P 4.30.
[60] G.F. Montufar, R. Pascanu, K. Cho, and Y. Bengio, On the number of linear regions of deep neural networks, NIPS 2014, pp. 2924–2932.
[61] T. Neylon, Sparse solutions for linear prediction problems, PhD Thesis, New York Univ., Courant Institut of Mathematical Sciences, 2006.
[62] S. Ovchinnikov, Graphs and Cubes, Springer, New York, 2011.
[63] A. Padrol and E. Philippe, Sweeps, polytopes, oriented matroids, and allowable graphs of permutations, arXiv preprint arXiv:2102.06134 (2021).
[64] D. Pálvölgyi and G. Tardos, Unlabeled compression schemes exceeding the VC-dimension, Discr. Appl. Math. 276 (2020), 102–107.
[65] B.I.P. Rubinstein and J.H. Rubinstein, Geometric and topological representation of maximum classes with application to sample compression, COLT'08, pp. 299–310, 2008.
[66] B.I.P. Rubinstein and J.H. Rubinstein, A geometric approach to sample compression, JMLR 13 (2012), 1221–1261.
[67] B.I.P. Rubinstein, J.H. Rubinstein, and P.L. Bartlett, Bounding embeddings of VC classes into maximum classes, in: V. Vovk, H. Papadopoulos, A. Gammerman (Eds.), Measures of Complexity. Festschrift for Alexey Chervonenkis, Springer, 2015, pp. 303–325.
[68] P.W. Shor, Stretchability of pseudolines is NP-hard, The Victor Klee Festschrift, DIMACS Ser. Discr. Math. and Theor. Comput. Sci., 4 (1991), 534–554.
[69] V.N. Vapnik and A.Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Theory Probab. Appl. 16 (1971), 264–280.
[70] M.K. Warmuth, Compressing to VC dimension many points, COLT/Kernel 2003, 743–744.