A Homological Theory of Functions: Nonuniform Boolean Complexity Separation and VC Dimension Bound Via Algebraic Topology, and a Homological Farkas Lemma*

Greg Yang†

MSR AI, Microsoft Research, Redmond, WA, USA
gregyang@microsoft.com

Abstract

In computational complexity, a complexity class is given by a set of problems or functions, and a basic challenge is to show separations of complexity classes $\mathtt{A} \neq \mathtt{B}$ especially when $\mathtt{A}$ is known to be a subset of $\mathtt{B}$. In this paper we introduce a homological theory of functions that can be used to establish complexity separations, while also providing other interesting consequences. We propose to associate a topological space $\mathcal{S}_{\mathtt{A}}$ to each class of functions $\mathtt{A}$, such that, to separate complexity classes $\mathtt{A} \subseteq \mathtt{B}'$, it suffices to observe a change in "the number of holes", i.e. homology, in $\mathcal{S}_{\mathtt{A}}$ as a subclass $\mathtt{B} \subseteq \mathtt{B}'$ is added to $\mathtt{A}$. In other words, if the homologies of $\mathcal{S}_{\mathtt{A}}$ and $\mathcal{S}_{\mathtt{A}\cup\mathtt{B}}$ are different, then $\mathtt{A} \neq \mathtt{B}'$. We develop the underlying theory of functions based on homological commutative algebra and Stanley-Reisner theory, and prove a "maximal principle" for polynomial threshold functions that is used to recover Aspnes, Beigel, Furst, and Rudich's characterization of the polynomial threshold degree of symmetric functions. A surprising coincidence is demonstrated, where, roughly speaking, the maximal dimension of "holes" in $\mathcal{S}_{\mathtt{A}}$ upper bounds the VC dimension of $\mathtt{A}$, with equality for common computational cases such as the class of polynomial threshold functions or the class of linear functionals over $\mathbb{F}_2$, or common algebraic cases such as when the Stanley-Reisner ring of $\mathcal{S}_{\mathtt{A}}$ is Cohen-Macaulay. As another interesting application of our theory, we prove a result that a priori has nothing to do with complexity separation: it characterizes when a vector subspace intersects the positive cone, in terms of homological conditions. By analogy to Farkas' result doing the same with *linear conditions*, we call our theorem the Homological Farkas Lemma.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes, I.2.6 Learning, G.2.1 Combinatorics, G.2.m Miscellaneous

Keywords and phrases Homology, Stanley-Reisner, Cellular resolution, VC dimension, Homological Farkas

Digital Object Identifier 10.4230/LIPIcs.ITCS.2018.56

1 Introduction

Computational complexity is one of the most important areas of theoretical computer science, within which complexity lower bounds is the aspect that is least understood. Basic questions such as P vs NP, P vs BPP, L vs P, and so on still remain open. In this work, we propose the following method for proving nonuniform Boolean lower bounds. For every class $\mathtt{C}$ of

* A full version of this paper is available at [19], https://arxiv.org/abs/1701.02302

† Work done while at Harvard University

9th Innovations in Theoretical Computer Science Conference (ITCS 2018).
Editor: Anna R. Karlin; Article No. 56; pp. 56:1–56:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

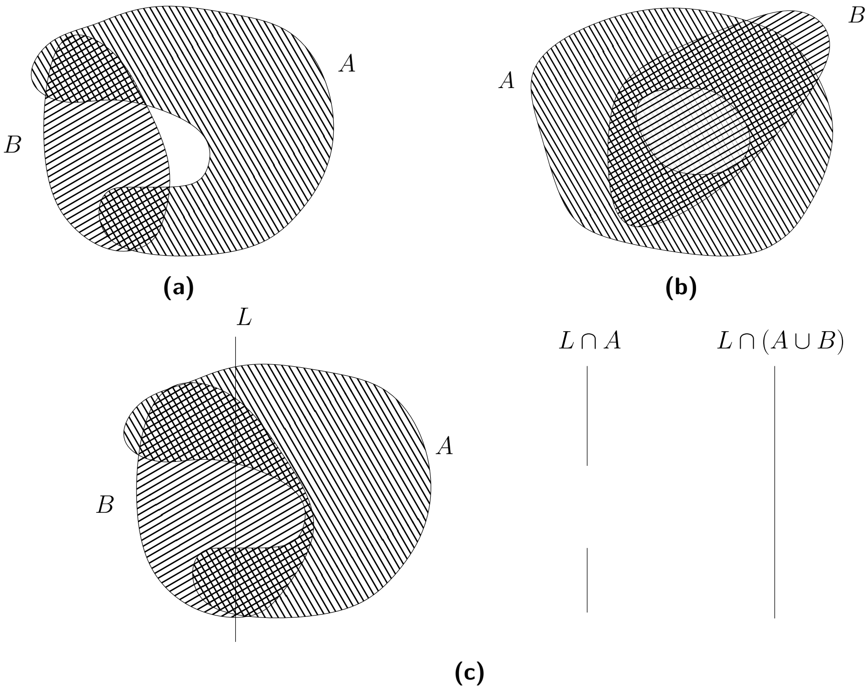

Figure 1 Suppose $A = \mathcal{S}_{\mathtt{C}}$ and $B = \mathcal{S}_{\{f\}}$. We can hope to certify $f \notin \mathtt{C} \iff A \cup B \neq A$ by noting that the numbers of 1-dimensional holes are different between $A \cup B$ and A. There can be many scenarios: (a) A and B are both contractible (do not have holes), but their union $A \cup B$ has a hole. Or (b) A has a hole in its center, but B covers it, so that $A \cup B$ is now contractible. Or (c) $A \cup B$ and A are both contractible, but if we look at a section L of $A \cup B$, we see that $L \cap A$ has 2 connected components, but $L \cap (A \cup B)$ has only 1. This shows why we need to look at homologies of subspaces or subcomplexes.

functions, we associate a simplicial complex $\mathcal{S}_{\mathtt{C}}$ to $\mathtt{C}$ in a way to be described in a moment; to show that a function f is not in $\mathtt{C}$, it suffices to show that $\mathcal{S}_{\mathtt{C}}$ is different from $\mathcal{S}_{\mathtt{C} \cup \{f\}}$ [1]. In this paper, we attempt to do so by showing that the (co)homologies [2] of $\mathcal{S}_{\mathtt{C}}$ and those of $\mathcal{S}_{\mathtt{C} \cup \{f\}}$ (or those of corresponding subcomplexes) are different. Figure 1 illustrates this idea.

▶ **Definition 1.1.** Write $[n] = \{0, 1, \ldots, n-1\}$. For a class $\mathtt{C} \subseteq [2]^{[n]}$ consisting of Boolean functions on a common domain $[n]$, $\mathcal{S}_{\mathtt{C}}$ is constructed as follows: There are $|[n] \times [2]| = 2n$ vertices, each labeled by an input/output pair $u \mapsto b$ with $u \in [n]$ and $b \in [2]$. For each function $f \in \mathtt{C}$, one adds to $\mathcal{S}_{\mathtt{C}}$ a maximal simplex on the n vertices with labels of the form $u \mapsto f(u)$. Vertices that don't belong to any such simplices are deleted.

In theoretical computer science, we will be mostly interested in the case when $[n] = [2^d] \cong [2]^d$ is the set of length d boolean strings. See Section 5 for examples. We use this idea to achieve several results in this paper.

---

[1] For readers unfamiliar with simplicial complexes, Appendix A provides a quick introduction.
[2] or in particular the Betti numbers, i.e. ranks of (co)homologies. For readers unfamiliar with homology, this is roughly speaking the "number of holes" of different dimensions in the simplicial complex. Intuitively speaking, the "dimension" of a hole is 1 if the hole "looks like" a circle, 2 if it "looks like" a sphere, and so on.

1.1 Aspnes-Beigel-Furst-Rudich Bound via Homology

Let $\textsc{polythr}_d^k$ denote the class of polynomial threshold functions of degree k on input space $\{-1,1\}^d$. The *polynomial threshold degree* of a Boolean function f is the smallest k such that $f \in \textsc{polythr}_d^k$. We give a new proof of Aspnes et al.'s result [2] that gives the polynomial threshold degrees of symmetric functions. The primary conduit is a general "maximal principle for polynomial thresholds," which is derived by looking at how adding a function to a degree bounded polynomial threshold class changes low dimensional Betti numbers. It says

▶ **Theorem 1.2** (Maximal Principle for Polynomial Threshold). *Let* $\mathtt{C} := \textsc{polythr}_d^k$*, and let* $f : \{-1,1\}^d \to \{-1,1\}$ *be a function. We want to know whether* $f \in \mathtt{C}$*.*

Suppose there exists a function $g \in \mathtt{C}$ *(a "local maximum" for approximating* f*) such that: for each* $h \in \mathtt{C}$ *that differs from* g *on exactly one input* u*, we have* $g(u) = f(u) = \neg h(u)$*. If* $g \neq f$*, then* $f \notin \mathtt{C}$*. (i.e., if* $f \in \mathtt{C}$*, then the "local maximum"* g *must be a "global maximum").*

We furthermore prove in the full paper that adding $\textsc{parity}$ to $\textsc{polythr}_d^k, k < d$, "covers up" the only hole in $\mathcal{S}_{\textsc{polythr}_d^k}$. In general, $\mathcal{S}_{\textsc{polythr}_d^k \cup \{f\}}$ "has no holes" iff f has no weak representation by degree k polynomials [19]. [3]

1.2 VC Dimension Bound via Homology

We exhibit a surprising connection of our framework to classical learning theory. **VC dimension** of a class $\mathtt{C}$ is defined as the size of the largest subset U of the input space such that $\mathtt{C} \upharpoonright U$ contains all Boolean functions on U. It is roughly the number of samples needed to learn an unknown function f from a known class $\mathtt{C}$, up to multiplicative constants [10]. The **homological dimension** of a class $\mathtt{C}$, written $\dim_{\mathrm{h}} \mathtt{C}$, is defined precisely in the full paper [19], but intuitively, for most cases, it is one plus the highest dimension of any nontrivial homology group in $\mathcal{S}_{\mathtt{C}}$ [4]. Then we prove that

▶ **Theorem 1.3.**

$$\dim_{\mathrm{VC}} \mathtt{C} \leq \dim_{\mathrm{h}} \mathtt{C}.$$

The equality cases include when $\mathtt{C}$ *is the class of parity functions (i.e. linear functionals over* $\mathbb{F}_2$*), the class of degree* $\leq k$ *polynomial threshold functions (for any fixed* k*), and the class of monotone conjunctions. This inequality cannot be improved to an equality, because for the class of conjunctions the gap between the two sides is 1, and for the class of delta functions on* $\{0,1\}^d$ *the homological dimension is* 2^d *but the VC dimension is 1.*

We also introduce an algebraic property of function classes called *Cohen-Macaulayness*, which is related to the corresponding notion in commutative algebra. We show that all Cohen-Macaulay classes satisfy this inequality with equality. It is nevertheless a major open problem to characterize the equality cases and the cases where the gap between the two sides is small (say, polynomial in d).

This beautiful result suggest that our homological theory captures something essential about computation, that it's not a coincidence that we can use "holes" to prove complexity separation.

[3] A polynomial p *weakly represents* a Boolean function f is for every input x such that $p(x) \neq 0$, we have $\mathrm{sgn}(p(x)) = f(x)$.

[4] i.e. the "highest dimension of any hole in $\mathcal{S}_{\mathtt{C}}$." Intuitively speaking, the "dimension" of a hole is 1 if the hole "looks like" a circle, 2 if it "looks like" a sphere, and so on.

1.3 Homological Farkas Lemma

Farkas lemma [22] characterizes when a linear subspace intersects the positive cone using linear algebraic conditions. From our study of threshold function classes via the lens of homology, we obtain easily a *Homological Farkas* lemma which characterizes such situations using homological conditions. It roughly says that

▶ **Theorem 1.4** (Homological Farkas Lemma (Informal)). *Either a linear subspace intersects the positive cone, or its intersection with a part of the boundary of a neighboring cone has "holes," but not both.*

Section 6 provides a brief exposition on the precise statement and the intuition why it should be true.

In addition to the main results described above, we also provide a probabilistic interpretation of algebraic data, called Hilbert functions, derived from our theory and elucidate a connection to (co)sheaf theory, in the full paper [19]. We believe that these results are just the tip of a large, hidden (so far) iceberg that forms a multi-directional connection between computer science, algebra, and topology.

2 Related Works

2.1 Distributed Computability via Topology

Herlihy and Shavit [9] famously used topological techniques to characterize decision problems solvable in the basic shared memory model by asynchronous, wait-free protocols. While their work associates simplicial complexes to individual functions, we associate simplicial complexes to *classes* of functions. In addition, in contrast to their clever applications of elementary techniques of combinatorial topology, we leverage the more modern Stanley-Reisner theory and cellular resolutions heavily. They also focus on *continuous maps* much more than we do here, which is something our future work could possibly benefit from.

2.2 Algebraic Decision Tree Lower Bounds via Betti Numbers

A long line of work yielded lower bounds on algebraic decision trees via topological techniques [5, 6, 21, 20]. Typically these techniques first show that a set $A \subseteq \mathbb{R}^d$ of interest has high complexity in terms of some topological aspect, and then show that shallow algebraic decision trees cannot compute sets of too high complexity. Even disregarding the difference in domains ($\mathbb{R}^d$ vs a discrete set), these methods operate on a different level than what we propose in this paper. Here we compute the Betti numbers of function classes, not of functions themselves, and we prove lower bounds by observing that adding the function in question to the class of low complexity functions changes the Betti numbers of the class. In addition, we are not concerned with only Betti numbers graded by dimension, but also Betti numbers graded by partial functions (which correspond to Betti numbers of filtered subcomplexes $\mathcal{S}_{\mathtt{C}|\mathtt{g}}$; see Section 5 for definitions).

2.3 Geometric Complexity Theory

There is a superficial similarity of our work to Mulmuley's Geometric Complexity program [14] in that both associate mathematical objects to complexity classes and focus on finding obstructions to equality of complexity classes. In the case of geometric complexity, each class is associated to a variety, and the obstructions sought are of representation-theoretic nature. In our case, each class is associated to a labeled simplicial complex, and the obstructions sought are of homological nature. But beyond this similarity, the inner workings of the two techniques are quite distinct. Whereas geometric complexity focuses on using algebraic geometry and representation theory to shed light on algebraic complexity classes (such as the permanent vs determinant question), our approach uses combinatorial algebraic topology and has a framework general enough to reason about any class of functions, not just algebraic functions. This generality allowed, for example, the unexpected connection to VC dimension. Thus there is no obvious relationship between GCT and our homological theory. However, there is a spiritual link. Indeed, Mulmuley and Sohoni proposed looking at higher dimension cohomology of the associated varieties in [14]. One possible direction for our future work is also to note that many classes have action by a symmetry group (see, e.g., [8]) and study how the Betti numbers break up into irreducible representations.

2.4 Homotopy Type Theory

A recent breakthrough in understanding the connection between algebraic topology and computer science is Homotopy Type Theory (HoTT) [17]. This theory concerns itself with rebuilding the foundation of mathematics via a homotopic interpretation of type theoretic semantics. Some of the key observations were that dependent sum types in type theory correspond to fibrations in homotopy theory, and equality types correspond to homotopies.

While HoTT only concerns itself with the B side (logic and semantics) of TCS, in this paper we primarily apply algebraic topology to the A side (complexity and learning theory). As such there really is no common ground between us in the technical details. However, early phases of our homological theory were inspired by the "fibration" philosophy of HoTT. In fact, the simplicial complex $\mathcal{S}_{\mathsf{C}}$ was first constructed as a sort of "fibration" (which turned out to be a cosheaf, and not a fibration) as explained in the full paper [19]. It remains to be seen if other aspects of HoTT could be illuminating in future research.

2.5 Computable Analysis and Topology

Computable analysis and topology study how topological spaces and functions on topological spaces can be represented in digital computers [18]. The theory builds a beautiful correspondence between computability via type II Turing machines on one hand and continuity of functions on the other hand that corroborates discoveries made in descriptive set theory [11, 13]. The initial spark for this paper was when the author realized that polylogarithmic time computation of a point in a topological space in the framework of computable analysis corresponds to polynomial time approximation schemes and also, roughly speaking, PAC learning: given more time, an algorithm should be able to pinpoint the desired point in a space more and more accurately, similar to how learning algorithms should be able to achieve better and better generalization errors with more samples and more computation time. But this correspondence to PAC learning ignores the probability of failure, which depends on the underlying data distribution. This initialized a search for a topological space encoding both the data distribution and the concept classes. The canonical suboplexes described in this paper turned out to be the right objects; see the cosheaf construction in the full paper for more details [19].

3 Does Our Proposal Run into Known Barriers?

Some of the most remarkable results in theoretical computer science in the last few decades are explicit "no-go" theorems that show that a common technique used in the past for proving complexity lower bounds cannot be extended to prove $\mathsf{P} \neq \mathsf{NP}$. These include relativization [3], algebrization [1], and natural proofs [15].

A priori, our framework is not blocked by the relativization or algebrization barriers because there is no reason to expect homology computations to relativize.

3.1 Razborov-Rudich Natural Proofs

Based on the methods presented in this paper, one might try to show $\mathsf{NP} \not\subseteq \mathsf{P/poly}$ by showing that the Betti numbers of $\mathcal{S}_{\mathsf{SIZE}(d^c)}$ differ from those of $\mathcal{S}_{\mathsf{SIZE}(d^c)\cup\{\mathsf{3SAT}_d\}}$, for any c and large enough d. Would this be a natural proof in the sense of Razborov and Rudich [15]?

A predicate $\mathcal{P}$ on functions with d-bit inputs is called **natural** if it satisfies

- (Constructiveness) It is polynomial time in its input size: there is an $2^{O(d)}$-time algorithm that on input the graph of a function f, outputs $\mathcal{P}(f)$.
- (Largeness) A random function f satisfies $\mathcal{P}(f) = 1$ with probability at least $\frac{1}{d}$.

▶ **Theorem 3.1** (Razborov-Rudich [15]). *Suppose there is no subexponentially strong one-way functions. Then there exists a constant c such that no natural predicate $\mathcal{P}$ maps $\mathsf{SIZE}(d^c)$ to 0.*

In our case, since $\mathsf{SIZE}(d^c)$ has $2^{\mathsf{poly}(d)}$ functions, naively computing the dimension-$(2^d - k)$ homology of $\mathcal{S}_{\mathsf{SIZE}(d^c)\cup\{\mathsf{3SAT}_d\}}$ for any constant k requires computing the ranks of two $2^{\mathsf{poly}(d)}$-sized matrix, which is already superpolynomial time in 2^d, violating the "constructiveness" of natural proofs. It is unknown whether the "largeness" condition is also violated, but, for any fixed dimension r, we conjecture that the probability a random total function f changes the dimension r homology of $\mathcal{S}_{\mathsf{SIZE}(d^c)}$ is exponentially small. Thus a priori this homological technique is not natural (barring the possibility that in the future, advances in the structure of $\mathcal{S}_{\mathsf{SIZE}(d^c)}$ yield efficient algorithms for its homology). [5]

4 Discussion

We anticipate several questions about our approach and provide corresponding retorts.

4.1 The Aspnes-Beigel-Furst-Rudich bound is an easy result; is your technique really new?

We agree that we are proving old results which are not particularly difficult, but we contend that the proofs really are different and serve as proof of concepts for future endeavors.

There is a *local-global philosophy* of our homological approach to complexity, inherited from algebraic topology. If we are interested in showing $f \notin \mathsf{C}$, we first examine the intersections of f with certain fragments of functions in C, determined by the Betti numbers of C (this is the local step), and then piece together these fragments with nontrivial intersections with

[5] In general, given the contents of the full paper, one may also want to show that the ideal $I^{\star}_{\mathsf{SIZE}(d^c)\boxtimes\{\mathsf{3SAT}_d\}}$ is principal by showing that its Betti numbers are all zero except at dimension 0. Computing the Betti numbers of an arbitrary ideal is NP-hard in the number of generators [4], which is $\Omega(2^d)$ in this case. Thus a priori it seems unlikely this algebraic method is constructive.

f to draw conclusions about "holes" f creates or destroys (this is the global step). This is markedly different from conventional wisdom in computer science, which seeks to show that a function, such as $f = 3\text{SAT}$, has some property that no function in a class, say $\mathtt{C} = \mathsf{P}$, has. In that method, there is no *global* step that argues that some nontrivial global property of $\mathtt{C}$ changes after adding f into it.

This philosophy is evident in our maximal principle, where the "local maximum" condition is saying that when one looks at the intersections with f of g and its "neighbors" (local), these intersections together form a hole that f creates when added to $\mathtt{C}$ (global). [6] It certainly does not look like the maximal principle can be reduced to a "separation by property", as it seems to depend fundamentally on the function f and the class $\mathtt{C}$ at the same time.

The original proof of the Aspnes-Beigel-Furst-Rudich bound primarily operates through a theory of "strong" and "weak" degrees of functions built via linear programming duality, and contains no notion of "locally maximal approximating polynomial thresholds" that is central to the maximal principle. It is not clear if it is possible to reduce our proof to theirs.

4.2 Your Betti number results seem to follow from previous work on algebraic decision trees.

As explained in Section 2, the Betti number bounds on algebraic decision trees are for the Betti numbers of *semialgebraic* sets represented by *individual* functions, while we compute graded Betti numbers of the *simplicial complex* induced by a *class* of functions. All of our Betti number results for nontrivial classes such as polynomial thresholds and linear functionals are new.

4.3 Do we always have a "homological certificate" for complexity lower bounds?

We won't necessarily be able to spot a difference in homology between $\mathcal{S}_{\mathtt{C}}$ and $\mathcal{S}_{\mathtt{C}\cup\{f\}}$ (though this is the case for, for example, $\mathtt{C} = \text{POLYTHR}_d^k$ and $f = \text{PARITY}_d$). But, assuming the definitions in Section 5.2, we will always be able to spot a difference between pairs of subcomplexes $\mathcal{S}_{\mathtt{C}\restriction\mathtt{g}} \subseteq \mathcal{S}_{\mathtt{C}}$ and $\mathcal{S}_{\mathtt{C}\cup\{f\}\restriction\mathtt{g}} \subseteq \mathcal{S}_{\mathtt{C}\cup\{f\}}$ for some partial function $\mathtt{g}$. For example, trivially take $\mathtt{g} = f$; in general there is always some non-total partial function $\mathtt{g}$ that works (see full paper [19] for precise statement and proof).

4.4 It seems hard to scale your method to larger classes because it becomes too combinatorial too quickly.

This exponentiality in fact already occurs with the classes studied in this paper, but by using the structures of each class we were still able to obtain Betti numbers. Note that homology is polynomial time in the number of bits encoding the simplicial complex (think of it as the size of the corresponding hypergraph). So if the complexes were just polynomially large, then such an approach would probably run into the natural proof barrier.

[6] The homological intuition, in more precise terms, is that a local maximum $g \neq f \in \mathtt{C}$ implies that the filtered class $\mathtt{C} \downharpoonright (f \cap g)$ consists of a single point with label g, so that when f is added to $\mathtt{C}$, a zero-dimensional hole is created.

4.5 Would your method apply to less algebraic complexity classes?

It is important to note that there are two distinct levels of algebra involved in this paper: the algebra in the algebraic topology used on the *class* level (ex: Stanley-Reisner theory), and the algebraic structure used on the *function* level (ex: vector space structure in the class of linear functional). They are independent usages of algebra, and one can apply regardless of the other. For example, we have computed Betti numbers for classes that are not so algebraic such as conjunctions (and by symmetry disjunctions); it seems plausible to build upon these results to obtain results on the Betti numbers of circuit classes. Even for the more algebraic classes, the techniques in computing Betti numbers in the paper don't quite use all of the algebraic structures; for example the polynomial threshold computation really only depends on its linear structure, but not its multiplicative structure. So the lack of *algebraic* structure does not seem like the most pressing obstacle, but the lack of *any* structure whatsoever is probably more worrying. For most complexity classes the key seems to be to pick up an approximating class that 1) is representative of the difficulties of the class and 2) has enough structure to give rise to simplicial complexes amenable to analysis. In any case, the intuition for the homological angle of complexity is quite undeveloped at this stage; a priori, there is no reason to even think that we can compute the Betti numbers of polynomial thresholds, linear functionals, conjunctions, etc, but it is done nevertheless. So there is cause for optimism.

5 Warmups

To illustrate the main ideas of this paper without being mired in the algebraic details, we walk through some examples that require only comfort with combinatorics, basic knowledge of topology and simplicial complexes, and a geometric intuition for "holes." A short introduction to simplicial complexes is included as Appendix A. A brief note about notation: $[n]$ denotes the set $\{0, \ldots, n-1\}$, and $[n \to m]$ denotes the set of functions from domain $[n]$ to codomain $[m]$. The notation $\mathsf{f} :\subseteq A \to B$ specifies a partial function from domain A to codomain B. $\dagger$ represents the partial function with empty domain.

5.1 The Complete Class

If $\mathtt{C} = [n \to 2]$, then one can see that $\mathcal{S}_\mathtt{C}$ is isomorphic to the 1-norm unit sphere $S_1^{n-1} := \{\|x\|_1 = 1 : x \in \mathbb{R}^n\}$ (also known as an *orthoplex*, shown in Figure 2a). Indeed, each function $f \in \mathtt{C}$ adds a facet to $\mathcal{S}_\mathtt{C}$ corresponding to the standard simplex in an orthant of $\mathbb{R}^n$, and together they generate the 1-norm unit sphere. For general $\mathtt{C}$, $\mathcal{S}_\mathtt{C}$ can be realized as a subcomplex of S_1^{n-1}. For this reason, $\mathcal{S}_\mathtt{C}$ is called the **canonical suboplex** of $\mathtt{C}$, where "suboplex" is short for "sub-orthoplex."

5.2 Delta Function is Not Linear

Let $\textsc{linfun}_d \cong (\mathbb{F}_2^d)^*$ be the class of linear functionals of a d-dimensional vector space V over $\mathbb{F}_2$. If $d \geq 2$, then $\textsc{linfun}_d$ does not compute the indicator function $\mathbb{I}_\mathbf{1}$ of the singleton set $\{\mathbf{1} := 11\cdots 1\}$. This is obviously true, but let's try to reason in a "homological way."

Define the partial function $\Upsilon : \mathbf{0} \mapsto 0, \mathbf{1} \mapsto 1$. Observe that for every partial linear functional $\mathsf{h} \supset \Upsilon$ *strictly extending* Υ, $\mathbb{I}_\mathbf{1}$ intersects h nontrivially. (Because $\mathbb{I}_\mathbf{1}$ is zero outside of Υ, and every such h must send at least one element to zero outside of Υ). I claim this completes the proof. *Combinatorially*, this is because if $\mathbb{I}_\mathbf{1}$ were a linear functional, then for any 2-dimensional subspace W of V containing $\{\mathbf{0}, \mathbf{1}\}$, the partial

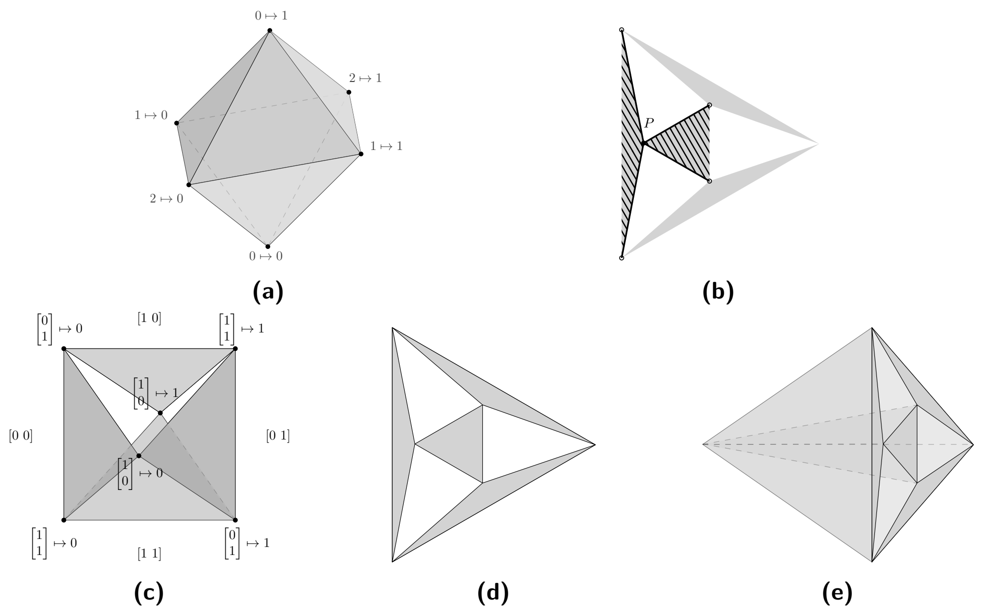

Figure 2 (a) The canonical suboplex of $[3 \to 2]$. (b) The open star $\operatorname{St} P$ of vertex P. (c) $\mathcal{S}_{\text{LINFUN}'_2}$ with vertices and facets labeled. (d) $\mathcal{S}_{\text{LINFUN}'_2}$ stretched flat. (e) $\mathcal{S}_{\text{LINFUN}_d}$ is just a cone over $\mathcal{S}_{\text{LINFUN}'_d}$.

function $\mathsf{h} :\subseteq V \to \mathbb{F}_2, \operatorname{dom} \mathsf{h} = W$,

$$\mathsf{h}(u) = \begin{cases} \Upsilon(u) & \text{if } u \in \operatorname{dom} \Upsilon \\ 1 - \mathbb{I}_{\mathbf{1}}(u) & \text{if } u \in \operatorname{dom} \mathsf{h} \setminus \operatorname{dom} \Upsilon \end{cases}$$

is a linear functional, and by construction, does not intersect $\mathbb{I}_{\mathbf{1}}$ on $W \setminus \{\mathbf{0}, \mathbf{1}\}$. *Homologically*, we are really showing the following

A section of $\mathcal{S}_{\text{LINFUN}_d}$ by an affine subspace corresponding to Υ "has a hole" that is "filled up" when $\mathbb{I}_{\mathbf{1}}$ is added to LINFUN_d. (*)

The meaning of this statement will seem cryptic right now, so let us elaborate.

Figure 2c exhibits the complex $\mathcal{S}_{\text{LINFUN}'_2}$, where $\text{LINFUN}'_d \subseteq [\mathbb{F}_2^d \setminus \{\mathbf{0}\} \to \mathbb{F}_2]$ is essentially the same class as LINFUN_d, except we delete $\mathbf{0}$ from the domain of every function. Notice that the structure of "holes" is not trivial at all: $\mathcal{S}_{\text{LINFUN}'_2}$ has 3 holes in dimension 1 but no holes in any other dimension. An easy way to visualize this is to pick one of the triangular holes; if you put your hands around the edge, pull the hole wide, and flatten the entire complex onto a flat plane, then you get Figure 2d.

It is easy to construct the canonical suboplex of LINFUN_d from that of LINFUN'_d: $\mathcal{S}_{\text{LINFUN}_d}$ is just a cone over $\mathcal{S}_{\text{LINFUN}'_d}$, where the cone vertex has the label $[0\ 0]^T \mapsto 0$ (Figure 2e). This is because every function in LINFUN_d shares this input/output pair. Note that a cone over any base has no hole in any dimension, because any hole can be contracted to a point in the vertex of the cone. This is a fact we will use again very soon.

Let $\mathtt{C} \subseteq [n \to 2]$, and let $\mathsf{f} :\subseteq [n] \to [2]$ be a partial function. Define the **filtered class** $\mathtt{C} \upharpoonright \mathsf{f}$ to be

$$\{g \setminus \mathsf{f} : g \in \mathtt{C}, g \supseteq \mathsf{f}\} \subseteq [[n] \setminus \operatorname{dom} \mathsf{f} \to [2]]$$

Unwinding the definition: $\mathtt{C} \upharpoonright \mathsf{f}$ is obtained by taking all functions of $\mathtt{C}$ that extend f and ignoring the inputs falling in the domain of f. The canonical suboplex $\mathcal{S}_{\mathtt{C} \upharpoonright \mathsf{f}}$ can be seen to be isomorphic to an affine section of $\mathcal{S}_{\mathtt{C}}$, when the latter is embedded as part of the L_1 unit sphere S_1^{n-1}. Figure 3a shows an example when f has a singleton domain. Indeed, recall LINFUN'_d is defined as $\text{LINFUN}_d \downharpoonright \{\mathbf{0} \mapsto 0\}$, and we may recover $\mathcal{S}_{\text{LINFUN}'_d}$ as an affine cut through the "torso" of $\mathcal{S}_{\text{LINFUN}_d}$ (Figure 3b). This explains the "affine section" part of (*).

To continue our elaboration, we need a "duality principle" in algebraic topology called the

▶ **Lemma 5.1** (Nerve Lemma (Informal)). *Let $\mathcal{U} = \{U_i\}_i$ be a* "nice" *(to be explained below) cover* [7] *of a topological space X. The* **nerve** *$\mathcal{N}_{\mathcal{U}}$ of $\mathcal{U}$ is defined as the simplicial complex with vertices $\{V_i : U_i \in \mathcal{U}\}$, and with simplices $\{V_i\}_{i\in S}$ for each index set S such that $\bigcap\{U_i : i \in S\}$ is nonempty.*

Then, for each dimension d, the set of d-dimensional holes in X is bijective with the set of d-dimensional holes in $\mathcal{N}_{\mathcal{U}}$.

What kind of covers are "nice?" Open covers in general spaces, or subcomplex covers in simplicial (or CW) complexes, are considered "nice", if in addition they satisfy the following requirements (*acyclicity*).

- Each set of the cover must have no holes.
- Each nontrivial intersection of a collection of sets must have no holes.

An example is the star cover: For vertex V in a complex, the **open star** $\operatorname{St} V$ of V is defined as the union of all open simplices whose closure meets V (see Figure 2b for an example). If the cover $\mathcal{U}$ consists of the open stars of every vertex in a simplicial complex X, then $\mathcal{N}_{\mathcal{U}}$ and X are isomorphic as complexes.

It turns out that $\mathcal{S}_{\text{LINFUN}'_d} = \mathcal{S}_{\text{LINFUN}_d \downharpoonright (\mathbf{0}\mapsto 0)}$ (a complex of dimension $2^d - 2$) has holes in dimension $d-1$ — in fact, these are the only holes in $\mathcal{S}_{\text{LINFUN}'_d}$ and the homological dimension of LINFUN'_d equals $d - 1 + 1 = d$, coinciding with its VC dimension. The proof is nontrivial and deferred to the full paper [19]. This can be clearly seen in our example when $d = 2$ (Figure 2d), which has 3 holes in dimension $d - 1 = 1$. Furthermore, for every partial linear functional h (a linear functional defined on a linear subspace), $\mathcal{S}_{\text{LINFUN}_d \downharpoonright \mathsf{h}}$ also has holes, in dimension $d-1-\dim(\operatorname{dom} \mathsf{h})$. Figure 3c show an example for $d = 2$ and $\mathsf{h} = [1\ 1]^T \mapsto 1$. This is in particular true for $\mathsf{h} = \Upsilon$. But when we add $\mathbb{I}_\mathbf{1}$ to LINFUN_d to obtain $\mathtt{D} := \text{LINFUN}_d \cup \{\mathbb{I}_\mathbf{1}\}$, $\mathcal{S}_{\mathtt{D}\downharpoonright\Upsilon}$ now does not have any hole! Figure 3d clearly demonstrates the case $d = 2$. For general d, note that $\mathcal{S}_{\text{LINFUN}_d \downharpoonright \Upsilon}$ has a "nice" cover by the open stars

$$\mathcal{C} := \{\operatorname{St} V : V \text{ has label } u \mapsto r \text{ for some } u \in \mathbb{F}_2^d \setminus \{\mathbf{0}, \mathbf{1}\} \text{ and } r \in \mathbb{F}_2\},$$

where the stars are with respect to $\mathcal{S}_{\text{LINFUN}_d \downharpoonright \Upsilon}$. When we added $\mathbb{I}_\mathbf{1}$ to form $\mathtt{D}$, the collection $\mathcal{C}' := \mathcal{C} \cup (\triangle_{\mathbb{I}_\mathbf{1} \setminus \Upsilon})$ is a "nice" cover of $\mathcal{S}_{\mathtt{D}\downharpoonright\Upsilon}$, where $\triangle_{\mathbb{I}_\mathbf{1}\setminus\Upsilon}$ is the face of $\mathbb{I}_\mathbf{1}$'s simplex generated by vertices with labels of the form $u \mapsto \mathbb{I}_\mathbf{1}(u), u \neq \mathbf{0}, \mathbf{1}$. Thus the nerve $\mathcal{N}_{\mathcal{C}'}$ has the same holes as $\mathcal{S}_{\mathtt{D}\downharpoonright\Upsilon}$, by the Nerve Lemma. But observe that $\mathcal{N}_{\mathcal{C}'}$ is a cone! …which is what our "combinatorial proof" of $\mathbb{I}_\mathbf{1} \notin \text{LINFUN}_d$ really showed.

More precisely,

1. a collection of stars $S := \{\operatorname{St} V : V \in \mathcal{V}\}$ has nontrivial intersection iff there is a partial linear functional extending the labels of each $V \in \mathcal{V}$.
2. We showed $\mathbb{I}_\mathbf{1} \setminus \Upsilon$ intersects every partial linear functional strictly extending Υ.
3. Therefore, a collection of stars S in $\mathcal{C}'$ intersects nontrivially iff $S \cup \{\triangle_{\mathbb{I}_\mathbf{1}\setminus\Upsilon}\}$ also intersects nontrivially.

[7] A cover of a space X is just a collection of sets whose union is X.

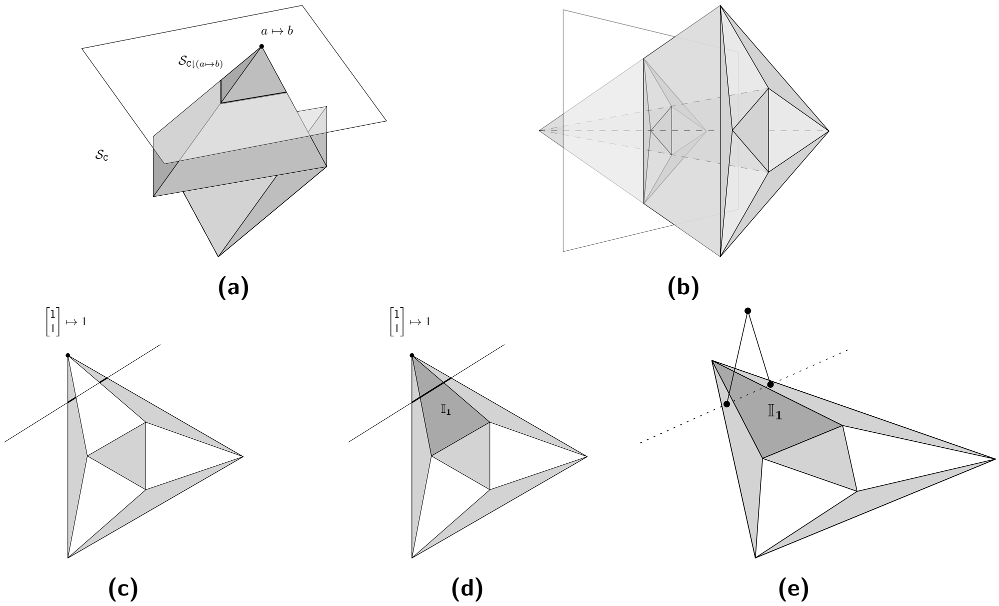

Figure 3 (a) $\mathcal{S}_{\mathsf{C}\downharpoonright(a\mapsto b)}$ is an affine section of $\mathcal{S}_{\mathsf{C}}$. (b) We may recover $\mathcal{S}_{\text{LINFUN}'_d}$ as a linear cut through the "torso" of $\mathcal{S}_{\text{LINFUN}_d}$. (c) $\mathcal{S}_{\text{LINFUN}_2\downharpoonright\{[0\ \ 0]^T\mapsto 0,[1\ \ 1]^T\mapsto 1\}}$ is isomorphic to the affine section as shown; it has "a single dimension zero hole." (d) When we add $\mathbb{I}_1$ to LINFUN_d to obtain $\mathtt{D} := \text{LINFUN}_d \cup \{\mathbb{I}_1\}$, $\mathcal{S}_{\mathtt{D}\downharpoonright\Upsilon}$ now does not have any hole! (e) The nerve $\mathcal{N}_{\mathcal{C}'}$ overlayed on $\mathtt{D} = \text{LINFUN}_2 \cup \{\mathbb{I}_1\}$. Note that $\mathcal{N}_{\mathcal{C}'}$ is a cone over its base of 2 points.

In other words, in the nerve of $\mathcal{C}'$, $\triangle_{\mathbb{I}_1\setminus\Upsilon}$ forms the vertex of a cone over all other $\operatorname{St} V \in \mathcal{C}$. In our example of LINFUN_2, this is demonstrated in Figure 3e.

Thus, to summarize,

- $\mathcal{N}_{\mathcal{C}'}$, being a cone, has no holes.
- By the Nerve Lemma, $\mathcal{S}_{\mathtt{D}\downharpoonright\Upsilon}$ has no holes either.
- Since $\mathcal{S}_{\text{LINFUN}_d\downharpoonright\Upsilon}$ has holes, we know $\mathtt{D} \downharpoonright \Upsilon \neq \text{LINFUN}_d \downharpoonright \Upsilon \implies \mathtt{D} \neq \text{LINFUN}_d$, i.e. $\mathbb{I}_1 \notin \text{LINFUN}_d$, as desired.

While this introduction took some length to explain the logic of our approach, much of this is automated in the theory we develop in this paper, which leverages existing works on Stanley-Reisner theory and cellular resolutions. The Nerve Lemma will in fact never be explicitly applied but rather is implicit in these machineries.

6 Homological Farkas Primer

We give a brief exposition on what the Homological Farkas Lemma says. Farkas Lemma is a simple result from linear algebra, but it is an integral tool for proving weak and strong dualities in linear programming [7], matroid theory [22], and game theory [12, chapter 7], among many other things.

▶ **Lemma 6.1** (Farkas Lemma). *Let $L \subseteq \mathbb{R}^n$ be a linear subspace not contained in any coordinate hyperplanes, and let $P = \{x \in \mathbb{R}^n : x > 0\}$ be the positive cone. Then either*

- *L intersects P, or*
- *L is contained in the kernel of a nonzero linear functional whose coefficients are all nonnegative.*

but not both.

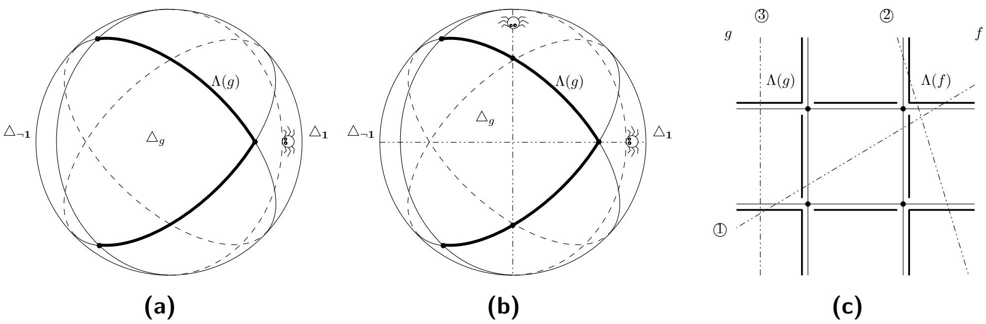

Figure 4 (a) An example of a $\Lambda(g)$. Intuitively, $\Lambda(g)$ is the part of $\partial\triangle_g$ that can be seen from an observer in $\triangle_{\mathbf{1}}$. (b) An illustration of Homological Farkas Lemma. The horizontal dash-dotted plane intersects the interior of $\triangle_{\mathbf{1}}$, but its intersection with any of the $\Lambda(f), f \neq \mathbf{1}, \neg\mathbf{1}$ has no holes. The vertical dash-dotted plane misses the interior of $\triangle_{\mathbf{1}}$, and we see that its intersection with $\Lambda(g)$ as shown has two disconnected components. (c) Example application of the affine version of homological Farkas lemma.. Let the hyperplanes (thin lines) be oriented such that the square S at the center is on the positive side of each hyperplane. The bold segments indicate the Λ of each region. Line 1 intersects S, and we can check that its intersection with any bold component has no holes. Line 2 does not intersect the closure $\overline{S}$, and we see that its intersection with $\Lambda(f)$ is two points, so has a "zeroth dimension" hole. Line 3 does not intersect $\overline{S}$ either, and its intersection with $\Lambda(g)$ consists of a point in the finite plane and another point on the circle at infinity.

Farkas Lemma is a characterization of when a linear subspace intersects the positive cone in terms of *linear conditions*. An alternate view important in computer science is that Farkas Lemma provides a *linear certificate* for when this intersection does not occur. Analogously, our Homological Farkas Lemma will characterize such an intersection in terms of *homological conditions*, and simultaneously provide a *homological certificate* for when this intersection does not occur.

Before stating the Homological Farkas Lemma, we first introduce some terminology.

For $g : [n] \to \{1, -1\}$, let $P_g \subseteq \mathbb{R}^n$ denote the open cone whose points have signs given by g. Consider the intersection $\triangle_g$ of $\overline{P_g}$ with the unit sphere S^{n-1} and its interior $\mathring{\triangle}_g$. $\mathring{\triangle}_g$ is homeomorphic to an open simplex. For $g \neq \neg\mathbf{1}$, define $\Lambda(g)$ to be the union of the facets F of $\triangle_g$ such that $\mathring{\triangle}_g$ and $\mathring{\triangle}_{\mathbb{1}}$ sit on opposite sides of the affine hull of F. Intuitively, $\Lambda(g)$ is the part of $\partial\triangle_g$ that can be seen from an observer in $\mathring{\triangle}_{\mathbf{1}}$ (illustrated by Figure 4a).

The following homological version of Farkas Lemma naturally follows from our homological technique of analyzing the complexity of threshold functions.

▶ **Theorem 6.2** (Homological Farkas Lemma). *Let $L \subseteq \mathbb{R}^n$ be a linear subspace. Then either*

- *L intersects the positive cone $P = P_{\mathbf{1}}$, or*
- *$L \cap \Lambda(g)$ for some $g \neq \mathbf{1}, \neg\mathbf{1}$ is nonempty and has holes.*

but not both.

Figure 4b illustrates an example application of this result.

One direction of the Homological Farkas Lemma has the following intuition. As mentioned before, $\Lambda(g)$ is essentially the part of $\partial\triangle_g$ visible to an observer Tom in $\mathring{\triangle}_{\mathbf{1}}$. Since the simplex is convex, the image Tom sees is also convex. Suppose Tom sits right on L (or imagine L to be a subspace of Tom's visual field). If L indeed intersects $\mathring{\triangle}_{\mathbf{1}}$, then for $L \cap \Lambda(g)$ he sees some affine space intersecting a convex body, and hence a convex body in itself. Since Tom sees everything (i.e. his vision is homeomorphic with the actual points), $L \cap \Lambda(g)$ has no holes, just as Tom observes. In other words, if Tom is inside $\mathring{\triangle}_{\mathbf{1}}$, then he cannot tell $\Lambda(g)$

is nonconvex by his vision alone, for any g. Conversely, the Homological Farkas Lemma says that if Tom is outside of $\mathring{\triangle}_{\mathbf{1}}$ and if he looks away from $\mathring{\triangle}_{\mathbf{1}}$, he will always see a nonconvex shape in some $\Lambda(g)$.

As a corollary to Theorem 6.2, we can also characterize when a linear subspace intersects a region in a linear hyperplane arrangement, and when an affine subspace intersects a region in an affine hyperplane arrangement, both in terms of homological conditions (see the full version of this paper [19] for details). A particular simple consequence, when the affine subspace either intersects the interior or does not intersect the closure at all, is illustrated in Figure 4c.

7 Overview of techniques and proofs

In this section we assume that the reader has the necessary algebraic and topological background. The complex $\mathcal{S}_{\mathtt{C}}$ is analyzed using Stanley-Reisner theory, which involves studying its face ideal $I_{\mathtt{C}}$ (i.e. the ideal consisting of monomials representing sets of vertices *not* in the complex $\mathcal{S}_{\mathtt{C}}$) and its Alexander dual $I^{\star}_{\mathtt{C}}$, primarily through the lens of free resolutions. It turns out that the Alexander dual is much easier to work with than the face ideal itself. The rank of each multigraded syzygy in its minimal resolution gives the Betti number of an appropriate dimension in the corresponding subcomplex of $\mathcal{S}_{\mathtt{C}}$ (more precisely, a certain link). This set of syzygy rank/Betti numbers is the principal topological invariant we use to separate classes in this work. Most resolutions here are computed as (co)cellular resolutions, i.e. we find labeled CW complexes whose (co)chain complexes resolve the ideals in question.

For the proof of Aspnes et al.'s theorem, we first obtain the cocellular resolutions of degree-bounded polynomial threshold classes $\textsc{polythr}^k_d$, which are supported on a natural CW decomposition of spheres. We yield the maximal principle (Theorem 1.2) by analyzing how dimension 1 Betti numbers change when a new function is added to $\textsc{polythr}^k_d$. We finish by constructing locally maximal approximating polynomial thresholds for symmetric functions. These local maxima are in general symmetric polynomial thresholds that encode the sign changes of the function in question as a polynomial of the sum of input bits.

Homological dimension is actually defined as the projective dimension of $I^{\star}_{\mathtt{C}}$, and as such it is the length of its minimal resolution. For the VC dimension bound, we give two proofs. The first observes that any minimal resolution of $I^{\star}_{\mathtt{C}}$ via relabeling supports a resolution of $I^{\star}_{\mathtt{C}\restriction U}$, for any subset U of the input space. When U is a largest shattering set, this shows that the homological dimension of $\mathtt{C}$ is at least the homological dimension of $\mathtt{C} \restriction U$, which we know is equal to $|U|$ and also equal to the VC dimension of $\mathtt{C}$. The second proof observes $\mathcal{S}_{\mathtt{D}}$ on an input space of size n has nontrivial homology in dimension $n-1$ iff $\mathcal{S}_{\mathtt{D}}$ contains every function. By applying this observation to $\mathcal{S}_{\mathtt{C}} \restriction U$ for a largest shattering set U, we get a lower bound of the regularity of $I_{\mathtt{C}}$ by the VC dimension plus one. Since regularity of $I_{\mathtt{C}}$ equals the projective dimension of $I^{\star}_{\mathtt{C}}$ plus one, we arrive at the desired result.

The much harder part of Theorem 1.3 is actually showing that equality holds in the various cases and that inequality is strict in other cases. This is done by deriving the (co)cellular resolutions of many function classes common in learning theory, such as conjunctions (supported on a pyramid over a pile of cubes) and monotone conjunctions (supported on a cube), degree bounded polynomial thresholds (supported on a CW decomposition of the sphere), linear functionals over finite fields, and so on. The topological dimension of such (co)cellular resolutions yield the homological dimension of the class itself, and we obtain the different equality and inequality cases listed in Theorem 1.3.

Finally, the Homological Farkas Lemma is obtained by drawing a parallel between, on the one hand, the intersection of a linear subspace with the positive cone, and on the other, the containment of a function in a generalized notion of a threshold class, which is possible due to the spherical structure of the cocellular resolution of such a class.

Acknowledgements. I would like to thank Madhu Sudan for hours of discussion, which were crucial in distilling the technical contents of this work into a clear exposition accessible to a general mathematical audience. I would also like to thank Josh Grochow for help fine-tuning the presentation toward a TCS audience. In addition, thanks are due to Günter Ziegler, Ezra Miller, Bernd Sturmfels, and Fatemeh Mohammadi, who provided references and discussion on algebra and combinatorics during the formative period of this theory; and to Leslie Valiant and Boaz Barak for listening to my babbles and provide encouragement.

References

1. Scott Aaronson and Avi Wigderson. Algebrization: A new barrier in complexity theory. *TOCT*, 1(1):2:1–2:54, 2009. doi:10.1145/1490270.1490272.
2. James Aspnes, Richard Beigel, Merrick L. Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994. doi:10.1007/BF01215346.
3. Theodore P. Baker, John Gill, and Robert Solovay. Relativizations of the P =? NP question. *SIAM J. Comput.*, 4(4):431–442, 1975. doi:10.1137/0204037.
4. Dave Bayer and Mike Stillman. Computation of Hilbert functions. *Journal of Symbolic Computation*, 14(1):31–50, 1992. URL: http://www.sciencedirect.com/science/article/pii/074771719290024X.
5. Michael Ben-Or. Lower bounds for algebraic computation trees (preliminary report). In David S. Johnson, Ronald Fagin, Michael L. Fredman, David Harel, Richard M. Karp, Nancy A. Lynch, Christos H. Papadimitriou, Ronald L. Rivest, Walter L. Ruzzo, and Joel I. Seiferas, editors, *Proceedings of the 15th Annual ACM Symposium on Theory of Computing, 25-27 April, 1983, Boston, Massachusetts, USA*, pages 80–86. ACM, 1983. doi:10.1145/800061.808735.
6. Anders Björner, László Lovász, and Andrew CC Yao. Linear decision trees: volume estimates and topological bounds. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 170–177. ACM, 1992. URL: http://dl.acm.org/citation.cfm?id=129730.
7. Stephen P. Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, UK ; New York, 2004.
8. Joshua A. Grochow. Unifying known lower bounds via geometric complexity theory. *Computational Complexity*, 24(2):393–475, 2015. doi:10.1007/s00037-015-0103-x.
9. Maurice Herlihy and Nir Shavit. The topological structure of asynchronous computability. *Journal of the ACM (JACM)*, 46(6):858–923, 1999. URL: http://dl.acm.org/citation.cfm?id=331529.
10. Michael Kearns and Umesh Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, jan 1994.
11. A. S. Kechris. *Classical descriptive set theory*. Number 156 in Graduate texts in mathematics. Springer-Verlag, New York, 1995.
12. Tjalling C Koopmans and others. *Activity analysis of production and allocation*. Wiley, 1951.
13. Yiannis N. Moschovakis. *Descriptive set theory*. Number v. 155 in Mathematical surveys and monographs. American Mathematical Society, Providence, R.I, 2nd ed edition, 2009.

14 Ketan Mulmuley and Milind A. Sohoni. Geometric complexity theory I: an approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2001. doi:10.1137/S009753970038715X.

15 Alexander A. Razborov and Steven Rudich. Natural proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997. doi:10.1006/jcss.1997.1494.

16 Joseph J. Rotman. *An introduction to algebraic topology.* Number 119 in Graduate texts in mathematics. Springer-Verlag, New York, 1988.

17 The Univalent Foundations Program. Homotopy Type Theory: Univalent Foundations of Mathematics. *arXiv preprint arXiv:1308.0729*, 2013. URL: https://arxiv.org/abs/1308.0729.

18 Klaus Weihrauch. *Computable Analysis.* Texts in Theoretical Computer Science. An EATCS Series. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000. DOI: 10.1007/978-3-642-56999-9. URL: http://link.springer.com/10.1007/978-3-642-56999-9.

19 Greg Yang. A Homological Theory of Functions. *arXiv:1701.02302 [cs, math]*, jan 2017. arXiv: 1701.02302. URL: http://arxiv.org/abs/1701.02302.

20 Andrew Chi-Chih Yao. Lower bounds for algebraic computation trees with integer inputs. *SIAM J. Comput.*, 20(4):655–668, 1991. doi:10.1137/0220041.

21 Andrew Chi-Chih Yao. Decision tree complexity and betti numbers. In Frank Thomson Leighton and Michael T. Goodrich, editors, *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 615–624. ACM, 1994. doi:10.1145/195058.195414.

22 Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 1995. URL: http://link.springer.com/10.1007/978-1-4613-8431-1.

A A Crash Course on Simplicial Complexes

Our presentation will follow [16].

▶ **Definition 1.1.** A d-dimensional **simplex** is just the convex hull of some affine independent subset $\{v_0, \dots, v_d\}$ of a Euclidean space. We denote such a simplex by $[v_0, \dots, v_d]$. The **vertex set** of a simplex $\triangle$ is denoted $\mathrm{Vrt}(\triangle)$.

▶ **Definition 1.2.** If $\triangle$ is a simplex, then a **face** of $\triangle$ is a simplex $\triangle'$ with $\mathrm{Vrt}(\triangle') \subseteq \mathrm{Vrt}(\triangle)$. $\triangle'$ is a **proper face** if the inclusion is strict.

Here is the main definition.

▶ **Definition 1.3.** A **simplicial complex** K is a finite collection of simplices in some Euclidean space such that

- (Hereditary) if $\triangle \in K$, then every face of $\triangle$ is also in K.
- (Regular intersection) if $\triangle, \triangle' \in K$, then $\triangle \cap \triangle'$ is either empty or a common face of $\triangle$ and of $\triangle'$.

The maximal faces of K, i.e. $\triangle \in K$ not properly contained in another $\triangle' \in K$, are called **facets** of K.

▶ **Definition 1.4.** If K is a simplicial complex, its **underlying space** $|K|$ is the subspace of the ambient Euclidean space given by the union of its simplexes:

$$|K| := \bigcup_{\triangle \in K} \triangle.$$

Figure 5 (a) Valid simplicial complex. (b) Invalid simplicial complex, because the intersection of the two simplices is not a face of the top simplex.

Most often we represent simplicial complexes pictorially as its underlying space. Figure 5 gives an example of a valid and of an invalid simplicial complex.

▶ **Definition 1.5.** The **boundary** $\partial\triangle$ of a simplex $\triangle$ is the collection of its proper faces.

▶ **Definition 1.6.** Suppose $\triangle$ is a d-dimensional simplex. If $d = 0$, define $\mathring{\triangle} = \triangle$. Otherwise, define $\mathring{\triangle} = \triangle \setminus |\partial\triangle|$. $\mathring{\triangle}$ is called an **open simplex**. For contrast, a plain simplex $\triangle$ is also called a **closed simplex**. The **closure** of an open simplex is just the corresponding closed simplex.

The above viewpoint of simplicial complexes is geometric and concrete. There is an equivalent combinatorial view that is sometimes more convenient to work with.

▶ **Definition 1.7.** Let V be a finite set. An **abstract simplicial complex** K is a family of nonempty subsets of V, called **simplices**, such that

- (Atomic) if $v \in V$, then $\{v\} \in K$;
- (Hereditary) if $\triangle \in K$ and $\triangle' \subseteq \triangle$, then $\triangle' \in K$.

V is called the **vertex set of** K and a simplex $\triangle \in K$ with $d + 1$ elements is called a d**-dimensional simplex**.

It is not hard to see that every simplicial complex has an associated abstract simplicial complex, simply by taking the vertex set of each simplex; call this **abstraction**. It is also true that every abstract simplicial complex has a topological space, called its **geometric realization**, that is a simplicial complex and is unique up to homeomorphism. These two operations are inverse in the sense that the geometric realization of an abstraction is homeomorphic to the original simplicial complex, and the abstraction of a geometric realization is isomorphic (in a suitable sense) to the original abstract simplicial complex. For details, see [16].

In this paper, we assume the equivalence of the geometric and combinatorial views, and use them interchangeably, as appropriate in different situations.

▶ **Definition 1.8.** Given a set of subsets $\nabla \subseteq 2^V$ of the vertex set, the (abstract) simplicial complex K generated by ∇ is the hereditary closure of ∇, i.e.

$$K = \{\triangle \subseteq V : \exists\triangle' \in \nabla, \triangle \subseteq \triangle'\}.$$

One can easily check that this is an (abstract) simplicial complex.

In this language, Theorem 1.1 says that $\mathcal{S}_{\mathtt{C}}$ is the simplicial complex on the vertex set $V = [2^d] \times [2]$ generated by $\{\operatorname{graph} f : f \in \mathtt{C}\}$, minus unused vertices.