

RECOGNIZING AND REALIZING INDUCTIVELY PIERCED CODES

Ryan Curry,* R. Amzi Jeffs,† Nora Youngs,‡ and Ziyu Zhao*

ABSTRACT. We prove algebraic and combinatorial characterizations of the class of inductively pierced codes, resolving a conjecture of Gross, Obatake, and Youngs. Starting from an algebraic invariant of a code called its canonical form, we explain how to compute a piercing order in polynomial time, if one exists. Given a piercing order of a code, we explain how to construct a realization of the code using a well-formed collection of open balls, and classify the minimal dimension in which such a realization exists.

1 Introduction

Given a collection $\mathcal{U} = \{U_1, \dots, U_n\}$ of convex sets in \mathbb{R}^d , one may combinatorially record how these sets intersect and cover one another by computing the *code* of the collection \mathcal{U} , defined as

$$\begin{aligned} \text{code}(\mathcal{U}) &\stackrel{\text{def}}{=} \left\{ \sigma \subseteq [n] \mid \bigcap_{i \in \sigma} U_i \setminus \bigcup_{j \in [n] \setminus \sigma} U_j \neq \emptyset \right\} \\ &= \{ \sigma \subseteq [n] \mid \text{there exists } p \in \mathbb{R}^d \text{ with } p \in U_i \text{ if and only if } i \in \sigma \}. \end{aligned}$$

Above, $[n] \stackrel{\text{def}}{=} \{1, 2, \dots, n\}$. Observe that $\text{code}(\mathcal{U})$ is a subset of the Boolean lattice $2^{[n]}$. The collection \mathcal{U} is called a *realization* of $\text{code}(\mathcal{U})$, and the region $\bigcap_{i \in \sigma} U_i \setminus \bigcup_{j \in [n] \setminus \sigma} U_j$ is called the *atom* of σ in the realization \mathcal{U} . Elements of a code are called *codewords*. With this terminology, the codewords of $\text{code}(\mathcal{U})$ are exactly the subsets of $[n]$ whose atoms are nonempty in the realization \mathcal{U} .

In 2013 Curto, Itskov, Veliz-Cuba, and Youngs [6] initiated the study of *open convex codes*, which are the codes that have realizations consisting of convex open sets. Their work was motivated by the neuroscientific study of place cells, and for this reason it is typical to refer to the indices in $[n]$ as *neurons*, and think of \mathbb{R}^d as a stimulus space in which the various U_i are the place fields where each neuron is active. Codes are a granular tool for capturing the intersection and covering patterns of a collection of sets, and so the study of codes that have realizations consisting of convex sets is of purely mathematical interest as well. Indeed, work in this direction has led to novel theorems in discrete geometry, interesting constructive results, and rich families of extremal examples [4, 5, 14, 3, 10, 12, 2, 7, 11].

*Department of Mathematics, Colby College, Waterville, ME 04901, USA.

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

amzij@cmu.edu. Supported by the National Science Foundation through Award No. 2103206.

‡Department of Mathematics, Colby College, Waterville, ME 04901, USA. nora.youngs@colby.edu. Supported by the Clare Boothe Luce Program and the Luce Foundation.

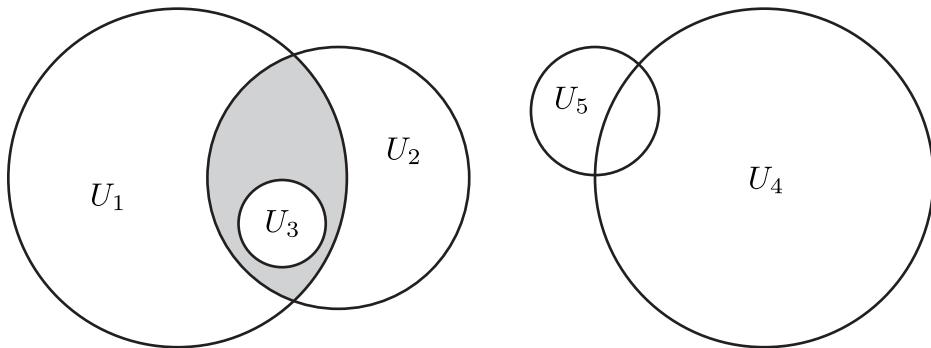


Figure 1: A realization of code C in the plane, with the atom of 12 highlighted in gray.

Example 1. Consider the code $C = \{123, 45, 12, 1, 2, 4, 5, \emptyset\}$ on five neurons. Here we are eliminating brackets and commas in codewords to make our notation more concise—for example, 123 represents $\{1, 2, 3\}$. We will adopt this convention elsewhere as well. Code C is convex; Figure 1 shows a realization of C using open balls in the plane.

Our work focuses exclusively on realizations consisting of convex open sets, and for this reason we use the more concise term *convex code* in place of “open convex code.” An efficient characterization of all convex codes is unfortunately out of the question; recent work by Kunin, Lienkaemper, and Rosen [13] indicates that recognizing whether or not a code is convex is NP-hard, even in \mathbb{R}^2 . One way around this problem is to restrict to a simpler class of codes. For example Cruz, Giusti, Itskov, and Kronholm [4] gave a construction showing that every intersection complete code¹ is convex.

In this vein, our work treats the class of *k-inductively pierced* codes, which can be built up iteratively from certain intervals of codewords (see Definition 2 below). We give an algebraic and geometric characterization of *k*-inductively pierced codes (Theorem 6), proving a conjecture of Gross, Obatake, and Youngs [9]. Furthermore, we exactly characterize the minimum dimension in which an inductively pierced code has a well-formed realization by open balls (Theorem 13); in particular, this shows that all inductively pierced codes are certainly convex. Importantly, all of our characterizations are efficient in the following sense: given the canonical form of a code, we can compute a piercing order in polynomial time if one exists, and a piercing order allows us to quickly compute the smallest dimension in which a code can be realized by a well-formed collection of open balls. This is discussed in Section 4. To present our results formally, we first require some additional background.

Inductively Pierced Codes. For any $\sigma \subseteq \tau \subseteq [n]$, the *interval* between σ and τ is the set

$$[\sigma, \tau] \stackrel{\text{def}}{=} \{\gamma \subseteq [n] \mid \sigma \subseteq \gamma \subseteq \tau\}.$$

The *rank* of an interval $[\sigma, \tau]$ is equal to $|\tau \setminus \sigma|$. Given a code $C \subseteq 2^{[n]}$ and a neuron $i \in [n]$, we will often consider the code $C \setminus i$, called the *deletion* of i , which is obtained by deleting i from every codeword in C where it appears. For any $\sigma \subseteq [n]$, we let $C \setminus \sigma$ denote the code obtained

¹A code is *intersection complete* if for any $\sigma_1, \sigma_2 \in C$, we have $\sigma_1 \cap \sigma_2 \in C$ also.

by successively deleting the neurons $i \in \sigma$ from \mathcal{C} . Geometrically, if $\mathcal{U} = \{U_1, \dots, U_n\}$ is a realization of \mathcal{C} then deleting U_i from this realization for all $i \in \sigma$ yields a realization of $\mathcal{C} \setminus \sigma$.

Informally, a k -inductively pierced code is one which can be built up as a union of intervals with rank at most k that fit together appropriately.

Definition 2. Let $\mathcal{C} \subseteq 2^{[n]}$ be a code and let $i \in [n]$ be a neuron. We say that i is an (*abstract*) k -piercing of \mathcal{C} if there exist $\sigma \subseteq \tau \subseteq [n] \setminus \{i\}$ so that

- (i) $[\sigma, \tau]$ has rank k ,
- (ii) $[\sigma, \tau]$ is contained in $\mathcal{C} \setminus i$, and
- (iii) $\mathcal{C} = (\mathcal{C} \setminus i) \cup [\sigma \cup \{i\}, \tau \cup \{i\}]$.

We say that $[\sigma, \tau]$ is the interval *associated to* the piercing, and every element of this interval must appear both with and without i ; these must also be the only appearances of i . Informally, neuron i only occurs within σ , and only interacts with the neurons in $\tau \setminus \sigma$ - but must interact with those neurons in every possible combination. A code $\mathcal{C} \subseteq 2^{[n]}$ is *k -inductively pierced* if $\mathcal{C} = \{\emptyset\}$, or there exists a neuron $i \in [n]$ which is a k' -piercing of \mathcal{C} for some $k' \leq k$, and $\mathcal{C} \setminus i$ is k -inductively pierced. A code is called simply *inductively pierced* if it is k -inductively pierced for some value of k .

From the definition, every k -inductively pierced code admits an ordering $i_1 < i_2 < \dots < i_n$ of the neurons in $[n]$ so that for each j , the neuron i_j is a k' -piercing of the code $\mathcal{C} \setminus \{i_{j+1}, \dots, i_n\}$ for some $k' \leq k$. We call such an order a *k -piercing order* (or merely *piercing order*); note that there is no requirement that such an ordering be unique.

Example 3. Consider again the code \mathcal{C} from Example 1. Here, neuron 3 is a 0-piercing of \mathcal{C} , with associated interval $[12, 12]$. Neuron 5 is a 1-piercing of \mathcal{C} with associated interval $[\emptyset, 4]$, and symmetrically, 4 is a 1-piercing of \mathcal{C} with interval $[\emptyset, 5]$. Neither neuron 1 nor neuron 2 are a k -piercing for any k . However, either would be a valid 1-piercing of $\mathcal{C} \setminus 3$. The code \mathcal{C} is inductively pierced, with many possible orderings. For example, both $1 < 2 < 3 < 4 < 5$ and $1 < 4 < 2 < 5 < 3$ are possible piercing orders.

Inductively pierced codes were previously studied by Gross, Obatake, and Youngs [9]. Motivated by the study of Euler diagrams, their work focused on 0-, 1-, and 2-inductively pierced codes and their possible realizations in the plane. We note that our definition of inductively pierced codes differs from theirs in notation, but is equivalent. Our notation will allow us to more easily characterize inductively pierced codes, and describe their realizations in higher dimensions.

The Neural Ideal. Curto, Itskov, Veliz-Cuba, and Youngs [6] introduced an algebraic approach to understanding codes and their realizations. Their approach used *pseudo-monomials* in the polynomial ring $\mathbb{F}_2[x_1, \dots, x_n]$, which are polynomials that have the form $\prod_{i \in \sigma} x_i \prod_{j \in \tau} (1 - x_j)$ where σ and τ are disjoint subsets of $[n]$. It will be useful to partially order pseudo-monomials by divisibility, so that in particular we can pick out *minimal*

pseudo-monomials from a given set. Each $\sigma \subseteq [n]$ can be associated to an *indicator pseudo-monomial*

$$\rho_\sigma \stackrel{\text{def}}{=} \prod_{i \in \sigma} x_i \prod_{j \in [n] \setminus \sigma} (1 - x_j).$$

These indicators allow us to uniquely associate each code to an ideal in $\mathbb{F}_2[x_1, \dots, x_n]$.

Definition 4 ([6]). Given a code $\mathcal{C} \subseteq 2^{[n]}$, we define the *neural ideal* $J_{\mathcal{C}} \subseteq \mathbb{F}_2[x_1, \dots, x_n]$ as follows:

$$J_{\mathcal{C}} \stackrel{\text{def}}{=} \langle \rho_\sigma \mid \sigma \notin \mathcal{C} \rangle.$$

The *canonical form* of $J_{\mathcal{C}}$, denoted $\text{CF}(J_{\mathcal{C}})$, is the set of minimal pseudo-monomials in $J_{\mathcal{C}}$.

Although the definition of $J_{\mathcal{C}}$ uses indicator pseudo-monomials, which have degree equal to n , the elements of the canonical form can in general have much lower degree. The pseudo-monomials in $J_{\mathcal{C}}$ have the following important geometric interpretation: $\prod_{i \in \sigma} x_i \prod_{j \in \tau} (1 - x_j)$ is an element of $J_{\mathcal{C}}$ if and only if $\bigcap_{i \in \sigma} U_i \subseteq \bigcup_{j \in \tau} U_j$ in every (possibly not convex or open) realization $\mathcal{U} = \{U_1, \dots, U_n\}$ of \mathcal{C} . Thus $\text{CF}(J_{\mathcal{C}})$ can be regarded as a minimal description of the intersection and covering information in any realization of \mathcal{C} .

Additionally, $J_{\mathcal{C}}$ provides a generalization of the well-studied Stanley-Reisner ideal [15, 17] of a simplicial complex: if \mathcal{C} is a simplicial complex then $\text{CF}(J_{\mathcal{C}})$ consists of monomials corresponding to minimal non-faces of \mathcal{C} , and the neural ideal $J_{\mathcal{C}}$ is the Stanley-Reisner ideal of \mathcal{C} (see [6, Section 4.4]). The ideal $J_{\mathcal{C}}$ and the canonical form are both unique to a particular code, and can be used to characterize codewords. Using a slight abuse of notation, we can evaluate polynomials in $J_{\mathcal{C}}$ on elements of $2^{[n]}$, where $f(\sigma)$ is computed by setting $x_i = 1$ if $i \in \sigma$, and $x_i = 0$ otherwise. We make use of the following useful facts: given $\sigma \subseteq [n]$, we have $\sigma \in \mathcal{C}$ if and only if $f(\sigma) = 0$ for all $f \in J_{\mathcal{C}}$ and if and only if $f(\sigma) = 0$ for all $f \in \text{CF}(J_{\mathcal{C}})$.

Conventions. We are interested in using the canonical form to study inductively pierced codes and their realizations. To streamline our work, we will adopt several conventions, which we explain below.

- (1) **Every code contains \emptyset as a codeword.** This means that sets in a realization do not cover all of \mathbb{R}^d , and in particular we may restrict our attention to realizations consisting of bounded sets. Algebraically, this means that $\prod_{i \in [n]} (1 - x_i)$ is not in $J_{\mathcal{C}}$ (corresponding geometrically to the convention that $\bigcap_{i \in \emptyset} U_i = \mathbb{R}^d$) and hence $\text{CF}(J_{\mathcal{C}})$ does not contain any pseudo-monomial that is entirely a product of $(1 - x_i)$ terms.
- (2) **Every neuron appears in some codeword.** This means that no set in a realization is empty. Algebraically, this means that x_i is not in $J_{\mathcal{C}}$ for any $i \in [n]$.
- (3) **No two neurons appear in exactly the same codewords.** This means that no two sets in a realization are equal. Note that if two neurons did have identical behavior, deleting one does not change whether or not a code is realizable, and so this assumption does not affect our analysis. Algebraically, this means that for any $i \neq j$ we do not have both $x_i(1 - x_j)$ and $x_j(1 - x_i)$ in $J_{\mathcal{C}}$.

Note that every code $\mathcal{C} \subseteq 2^{[n]}$ can be associated to a code \mathcal{D} satisfying (2) and (3) above, so that \mathcal{D} is convex if and only if \mathcal{C} is convex. Essentially, up to permutation, any neurons that do not appear could be forgotten, and likewise if two neurons have identical behavior one of them can be discarded.

Characterizing Inductively Pierced Codes. Note that our conventions (1) and (2) imply that every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ has degree at least two. We say that a code \mathcal{C} is *degree two* if every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ has degree exactly two. Gross, Obatake, and Youngs [9] showed that under the conventions (1)–(3), every 2-inductively pierced code is degree two. Their techniques imply the more general result that any inductively pierced code is degree two. We will postpone the proof of this and our other results until Section 2.

Proposition 5. *If $\mathcal{C} \subseteq 2^{[n]}$ is inductively pierced, then \mathcal{C} is degree two—that is, all elements of $\text{CF}(J_{\mathcal{C}})$ have degree two.*

Following [9], we associate each degree two code $\mathcal{C} \subseteq 2^{[n]}$ to a graph $G(\mathcal{C})$ called the *general relationship graph* of \mathcal{C} , which has vertex set $[n]$ and an edge ij whenever $\text{CF}(J_{\mathcal{C}})$ does not contain any pseudo-monomial whose two variables are x_i and x_j . Geometrically, ij is an edge in $G(\mathcal{C})$ if and only if U_i and U_j intersect and neither contains the other in every realization of \mathcal{C} . It turns out that the general relationship graph determines whether or not \mathcal{C} is inductively pierced.

Below, recall that a graph is *chordal* if it can be built up by successively adding *simplicial vertices*, whose neighborhood among previously added vertices is a clique. The reverse of the order in which these vertices are added is called a *perfect elimination order*.

Theorem 6. *A code $\mathcal{C} \subseteq 2^{[n]}$ is inductively pierced if and only if \mathcal{C} is degree two and $G(\mathcal{C})$ is chordal. If k is the smallest integer so that \mathcal{C} is k -inductively pierced, then the largest clique in $G(\mathcal{C})$ has size $k + 1$.*

Computing Piercing Orders. Theorem 6 gives a positive resolution to a conjecture of Gross, Obatake, and Youngs [9]. However, we would like to not only characterize inductively pierced codes, but also find a way to recognize them efficiently. To this end, we introduce a second auxiliary object associated to a degree two code. Given a degree two code \mathcal{C} , let $P(\mathcal{C})$ be the partially ordered set whose elements are $[n]$ and which has comparability relation defined by

$$i < j \text{ in } P(\mathcal{C}) \text{ if and only if } x_i(1 - x_j) \in \text{CF}(J_{\mathcal{C}}).$$

Recall that $\text{CF}(J_{\mathcal{C}})$ cannot contain both $x_i(1 - x_j)$ and $x_j(1 - x_i)$ by convention (3), so this relation is antisymmetric. Moreover, if we have $i < j < k$ in $P(\mathcal{C})$, then $x_i(1 - x_j)$ and $x_j(1 - x_k)$ both lie in $J_{\mathcal{C}}$, and so $x_i(1 - x_k) = (x_i(1 - x_j))(1 - x_k) + x_i(x_j(1 - x_k))$ lies in $J_{\mathcal{C}}$. In fact, $x_i(1 - x_k)$ lies in $\text{CF}(J_{\mathcal{C}})$ since \mathcal{C} is degree two, hence this relation is transitive, and defines a partial order. As previously noted in the discussion on the neural ideal, we have $x_i(1 - x_j)$ in $J_{\mathcal{C}}$ if and only if $U_i \subseteq U_j$ in every realization of \mathcal{C} , so $P(\mathcal{C})$ can be regarded as capturing pairwise containments of sets in any realization of \mathcal{C} .

We will use $P(\mathcal{C})$ and $G(\mathcal{C})$ to identify piercings of \mathcal{C} , and thus find piercing orders. It turns out that piercings of \mathcal{C} correspond to elimination neurons, defined below.

Definition 7. Let $\mathcal{C} \subseteq 2^{[n]}$ be a degree two code. We say that a neuron $i \in [n]$ is an *elimination neuron* if i is both a simplicial vertex in $G(\mathcal{C})$ and a minimal element in $P(\mathcal{C})$.

Proposition 8. Let $\mathcal{C} \subseteq 2^{[n]}$ be a degree two code. Then a neuron $i \in [n]$ is a k -piercing of \mathcal{C} if and only if i is an elimination neuron with k neighbors in $G(\mathcal{C})$.

Note that Proposition 8 already provides a novel characterization of piercing orders: they are the linear orders on $[n]$ whose reverse is simultaneously a perfect elimination order of $G(\mathcal{C})$ and a linear extension of $P(\mathcal{C})$.

Proposition 9. Let $\mathcal{C} \subseteq 2^{[n]}$ be a code, and suppose that a neuron $i \in [n]$ is a piercing of \mathcal{C} . Then \mathcal{C} is inductively pierced if and only if $\mathcal{C} \setminus i$ is inductively pierced.

Propositions 8 and 9 will let us argue that not only can we find piercings efficiently, but also that we will also never “get stuck” when trying to build a piercing order. In other words, a piercing order may be computed greedily. For a full discussion of this algorithm, see Section 4. For now, we summarize our realization results.

Realizing Inductively Pierced Codes. Since inductively pierced codes have a significant amount of combinatorial structure, it is reasonable to hope that they admit structured realizations. Rather than seek realizations by arbitrary convex open sets, we seek realizations by open balls. This can be achieved from any piercing order, and we can easily characterize the dimension in which such realizations exist. Below we make the additional requirement that our realizations are well-formed, in the sense that the boundaries of sets intersect generically.

Definition 10. A collection $\mathcal{S} = \{S_1, S_2, \dots, S_n\}$ of $(d-1)$ -dimensional spheres in \mathbb{R}^d is called *well-formed* if for every $k \leq d$ the intersection of k -many spheres from \mathcal{S} is either empty, or a $(d-k)$ -dimensional sphere, and the intersection of $(d+1)$ -many spheres from \mathcal{S} is empty. A collection of open balls is called well-formed if their boundaries comprise a well-formed collection of spheres.

The above definition of well-formed spheres mirrors the definition of well-formed curves used by Gross, Obatake, and Youngs [9] and the literature on Euler diagrams. Indeed, a collection of well-formed circles in \mathbb{R}^2 is a special case of a collection of well-formed curves. To characterize the exact dimensions in which an inductively pierced code admits a well-formed realization by open balls, we require two combinatorial definitions.

Definition 11. Let $\mathcal{C} \subseteq 2^{[n]}$ be a degree two code, and suppose that the largest clique in $G(\mathcal{C})$ has size $k+1$. For every $\sigma \subseteq [n]$ of size k on which $G(\mathcal{C})$ is a clique, the *attaching set* of σ consists of all $i \in [n] \setminus \sigma$ such that $\sigma \cup \{i\}$ is a $(k+1)$ -clique.

Note that the attaching set of σ is always an independent set, as an edge between two of its elements would yield a clique of size $k+2$ in $G(\mathcal{C})$.

Definition 12. Let $\mathcal{C} \subseteq 2^{[n]}$ be a degree two code, and suppose that the largest clique in $G(\mathcal{C})$ has size $k+1$. We say that \mathcal{C} is *splittable* if the following holds for every $\sigma \subseteq [n]$ of size k on which $G(\mathcal{C})$ is a clique: the attaching set of σ can be partitioned as $A \sqcup B$ so that $P(\mathcal{C})$ restricted to the attaching set consists of two disjoint total orders on A and B respectively.

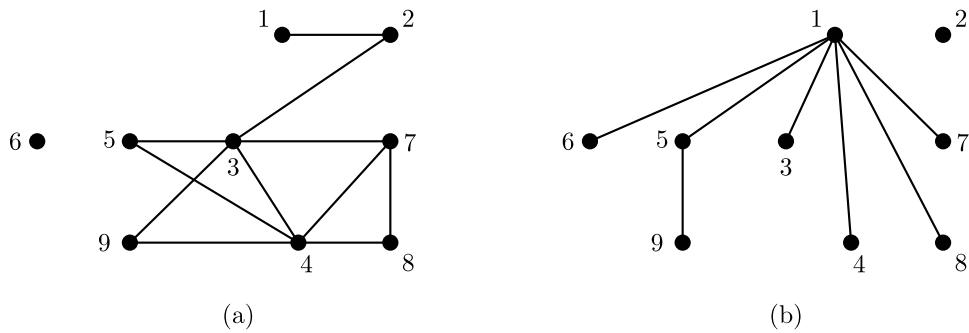


Figure 2: (a) The graph $G(\mathcal{C})$ associated to \mathcal{C} . (b) The Hasse diagram of $P(\mathcal{C})$.

Theorem 13. *Let $\mathcal{C} \subseteq 2^{[n]}$ be an inductively pierced code, and let $k \geq 1$. If the largest clique in $G(\mathcal{C})$ has size $k + 1$, then*

- (I) \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^{k+1} ,
- (II) \mathcal{C} has such a realization in \mathbb{R}^k if and only if \mathcal{C} is splittable, and
- (III) \mathcal{C} does not have such a realization in \mathbb{R}^{k-1} .

Note that Theorem 13 disregards 0-inductively pierced codes. These codes are simple to realize since their realizations consist of sets that are either pairwise disjoint or nested, which can always be managed in \mathbb{R}^1 .

Example 14. Consider the following code:

$$\begin{aligned} \mathcal{C} = & \{13459, 1347, 1478, 123, 16, \\ & 1345, 1359, 1459, 134, 135, 145, 137, 137, 147, 147, 148, 159, 178, 12, 13, 14, 15, 17, 18, 1, 2, \emptyset\}. \end{aligned}$$

Above, the maximal codewords appear on the first line and the remaining codewords are listed on the second line. This code is 2-inductively pierced, and the natural linear order on $\{1, 2, \dots, 9\}$ is a piercing order. It further turns out that \mathcal{C} is splittable. Note, for example, that the attaching set of $\{3, 4\}$ is $\{5, 7, 9\}$, and the partition $\{5, 9\} \sqcup \{7\}$ has $P(\mathcal{C})$ a total order on each part. Figure 2 shows $G(\mathcal{C})$ and $P(\mathcal{C})$. Item (II) of Theorem 13 implies that \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^2 , and Figure 3 shows such a realization.

Example 15. Consider the following code:

$$\begin{aligned} \mathcal{C} = & \{123, 124, 125, \\ & 12, 13, 14, 15, 23, 24, 25, 1, 2, 3, 4, 5, \emptyset\}. \end{aligned}$$

This code is 2-inductively pierced; indeed, the natural linear order is a piercing order, and the largest clique in $G(\mathcal{C})$ has size three. However, this code is not splittable. The attaching set of $\sigma = \{1, 2\}$ is $\{3, 4, 5\}$ and this set cannot be partitioned into two linearly ordered subsets in $P(\mathcal{C})$, since $P(\mathcal{C})$ is an antichain on this set. Thus, Theorem 13 implies that this code cannot be realized with open balls in \mathbb{R}^2 , but such a realization is possible in \mathbb{R}^3 . One such realization in \mathbb{R}^3 is shown in Figure 4.

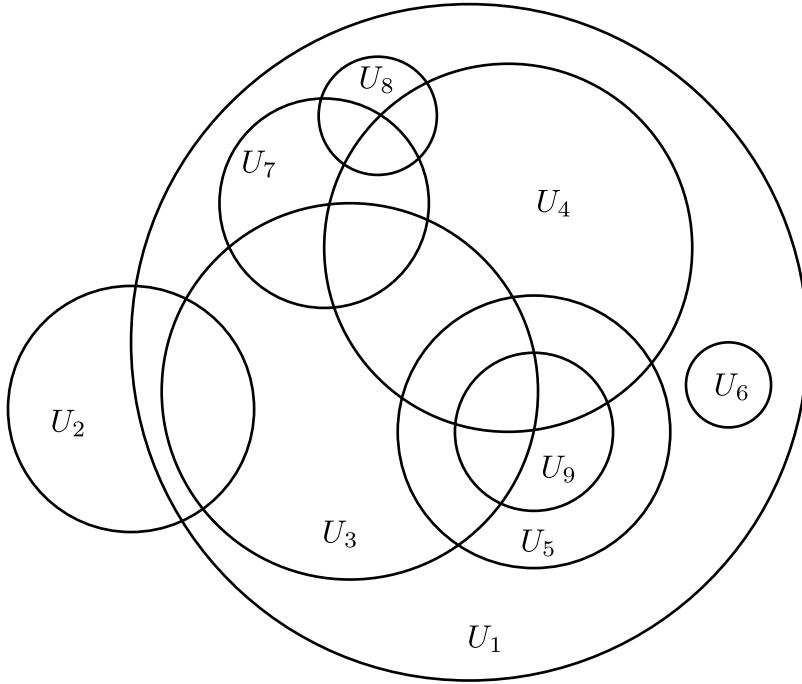


Figure 3: A well-formed realization of \mathcal{C} by open balls in \mathbb{R}^2 .

2 Characterizing Inductively Pierced Codes

To begin, we justify several important statements about the canonical form, which may be of general interest. First, we describe the effect of eliminating a neuron in any (not necessarily inductively pierced) code. We then characterize how a piercing impacts the canonical form, again without assuming that the code in question is inductively pierced. This yields a short proof of Proposition 5, establishing that every inductively pierced code is degree two.

Lemma 16. *For any code $\mathcal{C} \subseteq 2^{[n]}$, the canonical form of $\mathcal{C} \setminus n \subseteq 2^{[n-1]}$ consists of exactly the pseudo-monomials in $\text{CF}(J_{\mathcal{C}})$ which do not depend on the variable x_n ; that is,*

$$\text{CF}(J_{\mathcal{C} \setminus n}) = \text{CF}(J_{\mathcal{C}}) \setminus \{f \in \text{CF}(J_{\mathcal{C}}) \mid x_n \text{ or } (1 - x_n) \text{ divides } f\}.$$

Proof. Let f be any pseudo-monomial that does not depend on x_n . Observe that $f(\sigma) = 0$ if and only if $f(\sigma \setminus \{n\}) = 0$ for every $\sigma \subseteq [n]$. Thus f vanishes on every codeword of \mathcal{C} if and only if f vanishes on every codeword of $\mathcal{C} \setminus n$. Since the pseudo-monomials in the neural ideal of a code are exactly those which vanish on every codeword, we conclude that $f \in J_{\mathcal{C} \setminus n}$ if and only if $f \in J_{\mathcal{C}}$. Consequently, the minimal pseudo-monomials in $J_{\mathcal{C} \setminus n}$ are the minimal pseudo-monomials in $J_{\mathcal{C}}$ that do not depend on x_n , proving the result. \square

Lemma 17. *Let $\mathcal{C} \subseteq 2^{[n]}$ and let n be a k -piercing of \mathcal{C} with associated interval $[\sigma, \tau]$. Then*

$$\text{CF}(J_{\mathcal{C}}) = \text{CF}(J_{\mathcal{C} \setminus n}) \sqcup \{x_n x_i \mid i \in [n-1] \setminus \tau\} \sqcup \{x_n (1 - x_j) \mid j \in \sigma\}. \quad (*)$$

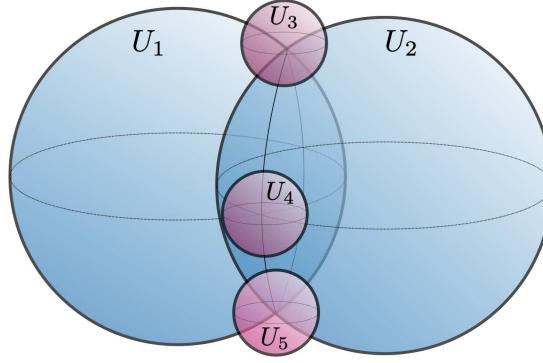


Figure 4: A realization of \mathcal{C} in \mathbb{R}^3 , consisting of two intersecting unit balls U_1 and U_2 , with U_3 , U_4 , and U_5 as smaller, disjoint balls centered at points on the circle where the boundaries of U_1 and U_2 meet.

Proof. By definition of k -piercing, the codewords of \mathcal{C} containing n form an interval $[\sigma \cup \{n\}, \tau \cup \{n\}]$. Hence n only appears in codewords that contain σ , and contain no neuron $i \in [n] \setminus \tau$. Consequently, $x_n(1 - x_i) \in J_{\mathcal{C}}$ for all $i \in \sigma$ and $x_n x_i \in J_{\mathcal{C}}$ for every $i \notin \tau$. Our conventions (1)–(3) forbid $J_{\mathcal{C}}$ from containing any degree one polynomials, and so these pseudo-monomials are minimal in $J_{\mathcal{C}}$, and hence lie in $\text{CF}(J_{\mathcal{C}})$. Lemma 16 shows that $\text{CF}(J_{\mathcal{C} \setminus n})$ is obtained from $\text{CF}(J_{\mathcal{C}})$ by deleting any pseudo-monomials in $\text{CF}(J_{\mathcal{C}})$ that depend on x_n . Putting all of this together, we have

$$\text{CF}(J_{\mathcal{C}}) \supseteq \text{CF}(J_{\mathcal{C} \setminus n}) \sqcup \{x_n x_i \mid i \in [n-1] \setminus \tau\} \sqcup \{x_n(1 - x_j) \mid j \in \sigma\}.$$

For the reverse containment, it will suffice to show that every pseudo-monomial in $J_{\mathcal{C}}$ is a multiple of one of the pseudo-monomials in the union of the three sets above. Equivalently, we must show the following.

Claim: If f is a pseudo-monomial that is *not* a multiple of some pseudo-monomial in

$$\text{CF}(J_{\mathcal{C} \setminus n}) \sqcup \{x_n x_i \mid i \in [n-1] \setminus \tau\} \sqcup \{x_n(1 - x_j) \mid j \in \sigma\}$$

then $f(c) \neq 0$ for some codeword $c \in \mathcal{C}$.

Case 1: f does not depend on x_n . Here Lemma 16 implies that f does not lie in $J_{\mathcal{C} \setminus n}$. Hence there is a codeword $c' \in \mathcal{C} \setminus n$ with $f(c') \neq 0$. By the definition of a k -piercing, every $c' \in \mathcal{C} \setminus n$ is also a codeword of \mathcal{C} .

Case 2: $(1 - x_n)$ divides f . We may write $f = (1 - x_n)g$ for some pseudo-monomial g that does not depend on x_n . Since f is not a multiple of any pseudo-monomial in $\text{CF}(J_{\mathcal{C} \setminus n})$, we must have $g \notin J_{\mathcal{C} \setminus n}$. Hence there is a codeword $c' \in \mathcal{C} \setminus n$ with $g(c') \neq 0$. But by the definition of a k -piercing, c' is also a codeword of \mathcal{C} , and furthermore we have $f(c') = g(c') \neq 0$ since c' does not contain n .

Case 3: x_n divides f . We can write $f = x_n g$ where g is a pseudo-monomial that does not depend on x_n . Since f is not a multiple of $x_n x_i$ for any $i \in [n-1] \setminus \tau$, we know that x_i does not divide g for any $i \in [n-1] \setminus \tau$. Similarly, since f is not a multiple of any $x_n(1-x_j)$ for $j \in \sigma$, we know that $(1-x_j)$ does not divide g for any $j \in \sigma$. Consequently, for some disjoint $\tau' \subseteq \tau$ and $\sigma' \subseteq [n-1] \setminus \sigma$ we have

$$g = \prod_{i \in \tau'} x_i \prod_{j \in \sigma'} (1-x_j).$$

By the definition of a k -piercing, \mathcal{C} contains the entire interval of codewords $[\sigma \cup \{n\}, \tau \cup \{n\}]$. In particular, $c \stackrel{\text{def}}{=} \sigma \cup \tau' \cup \{n\}$ is a codeword of \mathcal{C} . We see directly that $f(c) \neq 0$, so the claim follows in this case as well, establishing the lemma. \square

Proposition 5. *If $\mathcal{C} \subseteq 2^{[n]}$ is inductively pierced, then \mathcal{C} is degree two—that is, all elements of $\text{CF}(J_{\mathcal{C}})$ have degree two.*

Proof. This follows from Lemma 17 by induction. Note that the base case is easy as there is a unique code on one neuron that satisfies conventions (1)–(3) in the introduction, namely $\mathcal{C} = \{1, \emptyset\}$, and in this case $\text{CF}(J_{\mathcal{C}}) = \emptyset$, so \mathcal{C} is (vacuously) degree two. \square

We now prove Proposition 8, to establish that piercings correspond to elimination neurons. As we will see, this implies that the graph $G(\mathcal{C})$ associated to an inductively pierced code is chordal. We are then able to prove Theorem 6, which shows that, in fact, any degree two code whose corresponding graph is chordal must be inductively pierced.

Below we will work with degree two codes \mathcal{C} , and the associated graph $G(\mathcal{C})$ and partial order $P(\mathcal{C})$. For any neuron $i \in [n]$, we use the notation $N(i)$ for the neighborhood of i in $G(\mathcal{C})$, and $I(i)$ for the order ideal of i in $P(\mathcal{C})$. That is,

$$\begin{aligned} N(i) &\stackrel{\text{def}}{=} \{j \in [n] \mid i \text{ and } j \text{ adjacent in } G(\mathcal{C})\}, \text{ and} \\ I(i) &\stackrel{\text{def}}{=} \{j \in [n] \mid j < i \text{ in } P(\mathcal{C})\}. \end{aligned}$$

Proposition 8. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a degree two code. Then a neuron $i \in [n]$ is a k -piercing of \mathcal{C} if and only if i is an elimination neuron with k neighbors in $G(\mathcal{C})$.*

Proof. Up to permutation, it suffices to establish that n is a k -piercing of \mathcal{C} if and only if n is an elimination neuron with k neighbors in $G(\mathcal{C})$. First suppose that n is a k -piercing of \mathcal{C} with associated interval $[\sigma, \tau]$. By (\star) from Lemma 17, $(1-x_n)$ is not a factor of any pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$, and so n is minimal in $P(\mathcal{C})$. The equation (\star) also implies that the neighborhood of n in $G(\mathcal{C})$ is exactly $\tau \setminus \sigma$, which has k elements. It remains to argue that these neighbors form a clique. Let $i, j \in \tau \setminus \sigma$, and note from the definition of a k -piercing that $\sigma, \sigma \cup \{i\}, \sigma \cup \{j\}$, and $\sigma \cup \{i, j\}$ are all codewords of \mathcal{C} . The pseudo-monomials $(1-x_i)(1-x_j)$, $x_i(1-x_j)$, $x_j(1-x_i)$, and $x_i x_j$ respectively do not vanish on these codewords, and hence do not lie in $\text{CF}(J_{\mathcal{C}})$. Since these are the only degree two pseudo-monomials involving x_i and x_j , we conclude that no pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ has variables indexed by i and j . Hence i and j are adjacent in $G(\mathcal{C})$, and the neighbors of n form a clique as desired.

We now argue that if n is an elimination neuron with k neighbors in $G(\mathcal{C})$, then n is a k -piercing of \mathcal{C} . Let $\sigma = I(n)$ and let $\tau = I(n) \cup N(n)$. Note that $\tau \setminus \sigma = N(n)$, which has size k . Since $N(n)$ is a clique, for every pair of neurons in $N(n)$ there is no pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ whose variables are indexed by these neurons. To show that n is a k -piercing, we must verify two claims, which together show parts (ii) and (iii) of Definition 2 are satisfied.

Claim 1: $[\sigma, \tau \cup \{n\}] \subseteq \mathcal{C}$. Let $\gamma \in [\sigma, \tau \cup \{n\}]$. It will suffice to show that every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ vanishes on γ . First consider the pseudo-monomials that depend on x_n . Since n is minimal in $P(\mathcal{C})$, every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ that depends on x_n is of the form $x_n x_i$ with $i \notin \tau$, or $x_n(1 - x_j)$ with $j \in \sigma$. In both cases the second term in the pseudo-monomial vanishes since $\sigma \subseteq \gamma \subseteq \tau \cup \{n\}$.

It remains to consider pseudo-monomials that do not depend on x_n . Here it suffices to consider the case where $n \notin \gamma$, since adding or removing n will not affect the vanishing of the pseudo-monomial in question. First consider those of the form $x_i x_j$. Since the neighbors of n in $G(\mathcal{C})$ form a clique on $\tau \setminus \sigma$, one of i or j lies outside $\tau \setminus \sigma$. If i lies in σ , then $x_n(1 - x_i)$ lies in $\text{CF}(J_{\mathcal{C}})$, and hence

$$x_n(x_i x_j) + (x_n(1 - x_i))x_j = x_n x_j$$

lies in $J_{\mathcal{C}}$. This implies that $j \notin \tau$ by (*), and in particular $j \notin \gamma$. Hence $x_i x_j$ vanishes on γ .

Finally, consider a pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ that has the form $x_i(1 - x_j)$ where $i, j \in [n - 1]$. Suppose for contradiction that this pseudo-monomial does not vanish on γ . Then $i \in \gamma$ and $j \notin \gamma$. In particular, $i \in \tau$ and $j \notin \sigma$. We will argue that i and j both lie in $\tau \setminus \sigma$, contradicting the fact that $G(\mathcal{C})$ is a clique on these neurons. It suffices to show that $i \notin \sigma$ and $j \in \tau$. Note that if $i \in \sigma$ then $x_n(1 - x_i) \in J_{\mathcal{C}}$, and so

$$x_n(1 - x_i)(1 - x_j) + x_n x_i(1 - x_j) = x_n(1 - x_j)$$

lies in $\text{CF}(J_{\mathcal{C}})$, contradicting the fact that $j \notin \sigma$. Similarly, if $j \notin \tau$ then $x_n x_j$ lies in $J_{\mathcal{C}}$, and so

$$x_n x_i x_j + x_n x_i(1 - x_j) = x_n x_i$$

lies in $\text{CF}(J_{\mathcal{C}})$ contradicting the fact that $i \in \tau$. Hence i and j both lie in $\tau \setminus \sigma$, contradicting the fact that $G(\mathcal{C})$ is a clique on this set. We conclude that $[\sigma, \tau \cup \{n\}] \subseteq \mathcal{C}$.

Claim 2: If $c \in \mathcal{C}$ and $n \in c$, then $c \in [\sigma \cup \{n\}, \tau \cup \{n\}]$. Here we use the fact that every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ vanishes on c , together with (*). For all $j \in \sigma$, the pseudo-monomial $x_n(1 - x_j)$ vanishes on c , so we must have $\sigma \subseteq c$. Likewise, for every $j \notin \tau \cup \{n\}$ the monomial $x_n x_j$ vanishes on c , and so $c \subseteq \tau \cup \{n\}$. This proves the result. \square

Lemma 18. *Let \mathcal{C} be a degree two code, and let $i, j \in [n]$ be such that $j < i$ in $P(\mathcal{C})$. Then*

$$N(j) \subseteq I(i) \cup N(i).$$

Proof. Suppose for contradiction that j is adjacent to some neuron ℓ that lies in neither $I(i)$ nor $N(i)$. Since ℓ does not lie in $N(i)$, there must be a pseudo-monomial $f \in \text{CF}(J_{\mathcal{C}})$ that

depends on x_i and x_ℓ . Since ℓ does not lie in $I(i)$, the only possibilities are $f = x_i x_\ell$ or $f = x_i(1 - x_\ell)$. But $j < i$, and so $x_j(1 - x_i)$ lies in $\text{CF}(J_{\mathcal{C}})$. Multiplying $x_j(1 - x_i)$ by either x_ℓ or $1 - x_\ell$ and adding it to $x_j f$, we obtain a pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ that depends on j and ℓ , contradicting the fact that j and ℓ are adjacent in $G(\mathcal{C})$. \square

We are now ready to prove Theorem 6, and establish our characterization of inductively pierced codes.

Theorem 6. *A code $\mathcal{C} \subseteq 2^{[n]}$ is inductively pierced if and only if \mathcal{C} is degree two and $G(\mathcal{C})$ is chordal. If k is the smallest integer so that \mathcal{C} is k -inductively pierced, then the largest clique in $G(\mathcal{C})$ has size $k + 1$.*

Proof. If \mathcal{C} is k -inductively pierced, then our previous work immediately indicates the required conditions hold. Indeed, Proposition 5 guarantees that \mathcal{C} is degree two. Furthermore, Proposition 8 shows that a k -piercing order corresponds (in reverse) to a perfect elimination order on the vertices of $G(\mathcal{C})$ in which each simplicial vertex has at most k neighbors, implying that $G(\mathcal{C})$ is chordal with maximum clique size at most $k + 1$.

It remains to prove the converse: if \mathcal{C} is degree two and $G(\mathcal{C})$ is chordal with all cliques having size at most $k + 1$, we must argue that \mathcal{C} is k -inductively pierced. Using Proposition 8, we will find an elimination neuron $i \in [n]$ with at most k neighbors in $G(\mathcal{C})$, and prove that $\mathcal{C} \setminus i$ is still degree two with $G(\mathcal{C} \setminus i)$ chordal and all cliques having size at most $k + 1$. This allows us to inductively obtain a k -piercing order for \mathcal{C} .

Since $G(\mathcal{C})$ is chordal, $G(\mathcal{C})$ contains at least one simplicial vertex. Choose such a vertex $i \in [n]$ that is minimal amongst all simplicial vertices in $P(\mathcal{C})$ —that is, there is no simplicial vertex $j < i$. We will show that in fact i is minimal in $P(\mathcal{C})$; that is, $I(i) = \emptyset$.

Let G' be the induced subgraph of $G(\mathcal{C})$ on the vertex set $N(i) \cup I(i)$. Since every $j \in I(i)$ is not adjacent to i in $G(\mathcal{C})$, it will suffice to show that G' is a clique. For contradiction, suppose that G' is not a clique. Since $G(\mathcal{C})$ is chordal, any induced subgraph is chordal. Every chordal graph that is not a clique contains two non-adjacent simplicial vertices² and hence there is a vertex j that is simplicial in G' and not adjacent to i .

We must have $j < i$, and by Lemma 18 we have $N(j) \subseteq N(i) \cup I(i)$. In particular, the neighborhood of j in G' is the same as its neighborhood in G . But since this neighborhood is a clique, this contradicts our choice of i as a minimal simplicial vertex.

Hence we can find a neuron $i \in [n]$ that is simplicial in $G(\mathcal{C})$ and minimal in $P(\mathcal{C})$. Up to permutation, we can assume $i = n$. Since the largest clique in $G(\mathcal{C})$ is size $k + 1$, we conclude that n has $k' \leq k$ neighbors in $G(\mathcal{C})$. By Proposition 8, n is a k' -piercing of \mathcal{C} . By (★) we see that $\mathcal{C} \setminus n$ is degree two, that $G(\mathcal{C} \setminus n)$ is obtained from $G(\mathcal{C})$ by deleting n , and that $P(\mathcal{C} \setminus n)$ is obtained from $P(\mathcal{C})$ by deleting n . Hence we may conclude inductively that $\mathcal{C} \setminus n$ is k -inductively pierced, and the result follows. \square

We conclude with a proof of Proposition 9, establishing that piercing orders may be constructed greedily.

²This result is well-known in the chordal graph literature; see for example the proof of Theorem 8.3.27 in [18].

Proposition 9. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a code, and suppose that a neuron $i \in [n]$ is a piercing of \mathcal{C} . Then \mathcal{C} is inductively pierced if and only if $\mathcal{C} \setminus i$ is inductively pierced.*

Proof. It suffices to consider the case $i = n$. The reverse implication is clear, and it remains to argue that if \mathcal{C} is inductively pierced then so is $\mathcal{C} \setminus n$. By Theorem 6, it suffices to show that $\mathcal{C} \setminus n$ is degree two, and that $G(\mathcal{C} \setminus n)$ is chordal. By (★), $\text{CF}(J_{\mathcal{C} \setminus n}) \subseteq \text{CF}(J_{\mathcal{C}})$, and so $\mathcal{C} \setminus n$ is degree two. Moreover, $G(\mathcal{C} \setminus n)$ is the induced subgraph of $G(\mathcal{C})$ on the vertex set $[n - 1]$. Since $G(\mathcal{C})$ is chordal so is $G(\mathcal{C} \setminus n)$, and the result follows. \square

3 Realizing Inductively Pierced Codes

Below, we will make use of existing results regarding *set-theoretically independent* collections of open balls in \mathbb{R}^d —in our language, a collection $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of balls is set-theoretically independent whenever $\text{code}(\mathcal{U})$ is the full power set $2^{[n]}$. For concision, we will use the simpler term *independent*. We will need the following facts regarding collections of independent balls:

- (i) The largest size of a collection of independent balls in \mathbb{R}^d is $d + 1$. This was first established by Rényi, Rényi, and Surányi [16].
- (ii) Let $\{U_1, \dots, U_k\}$ be a collection of $k \leq d$ independent balls in \mathbb{R}^d . Then the intersection of the boundaries of the U_i is a $(d - k)$ -dimensional sphere S . In particular, the U_i are a well-formed collection of balls. This result appears as Theorem 1 in a paper by Anusiak [1].
- (iii) In the situation of item (ii), S is the set of common limit points of all atoms in the realization \mathcal{U} (Theorem 3 of [1]).
- (iv) In the situation of item (ii), suppose additionally that $k = d$, and let U_{d+1} be any ball. Then $S = \{p, q\}$ is a 0-dimensional sphere, and the collection of balls $\{U_1, \dots, U_d, U_{d+1}\}$ is independent if and only if the closure of U_{d+1} contains exactly one of p or q (Theorem 2 of [1]).

To build well-formed realizations of inductively pierced codes, we will use the following geometric analogue of an abstract piercing.

Definition 19. Let $\mathcal{U} = \{U_1, \dots, U_n\}$ be a realization of $\mathcal{C} \subseteq 2^{[n]}$ in \mathbb{R}^d . An interval $[\sigma, \tau] \subseteq \mathcal{C}$ is called *pierceable* at a point $p \in \mathbb{R}^d$ if the codewords which arise in any small neighborhood of p are exactly those in $[\sigma, \tau]$. Equivalently, $[\sigma, \tau]$ is pierceable at p if and only if p is a limit point of the atom of every codeword in $[\sigma, \tau]$ and no others.

Given a collection $\{U_1, \dots, U_n\}$ of well-formed (but not necessarily independent) balls in \mathbb{R}^d , and a point $p \in \mathbb{R}^d$, it will be convenient to assign p to a vector $\rho(p)$ of length n on the symbols $\{+, 0, -\}$, as follows. The i -th coordinate of $\rho(p)$ will record whether p lies in the interior, boundary, or complement of the closure of U_i , with $+$, 0 , and $-$ denoting these respective possibilities. With this notation, an interval $[\sigma, \tau]$ is pierceable at p if and only if $\rho(p)$ has $+$'s in the coordinates of σ , 0 's in the coordinates of $\tau \setminus \sigma$, and $-$'s everywhere else

(note that to establish this we need to use item (iii) above on the collection of balls indexed by $\tau \setminus \sigma$). With this observation, we can prove the following technical lemma.

Lemma 20. *Let $1 \leq k \leq d$, and let $\{U_1, \dots, U_k\}$ be a collection of independent balls in \mathbb{R}^d . Then every interval $[\sigma, \tau] \subseteq 2^{[k]}$ is pierceable at some point.*

Proof. Let p be a point in the intersection of the boundaries of all U_i , noting that such a point exists because $k \leq d$. Observe that $\rho(p)$ is the all zeroes vector. We wish to find a point whose vector has + for each index in σ , 0 for each index in $\tau \setminus \sigma$, and – for all other indices (that is, a pierceable point). To achieve this, for every $i = 1, 2, \dots, k$, we will replace p by a new point, so that our new point has the correct symbols in the first i indices, and 0's in the remaining indices.

Let $i \in [k]$, and assume that $\rho(p)$ has the correct symbols in indices less than i , and 0's for indices equal to i and larger. Then p lies in the intersection of the boundaries of all U_j with $j > i$, as well as the U_j with $j < i$ and $\rho(p)$ equal to zero in the j -th index. Let S denote the intersection of these boundaries (setting $S = \mathbb{R}^d$ if there are no boundaries being intersected). Observe that $\partial U_i \cap S$ is a sphere with codimension one in S , and that it contains p . Thus ∂U_i divides S into two nonempty halves, one contained in U_i and the other contained in the complement of the closure of U_i , and p is a limit point of each of these halves. By moving p slightly inside S , we do not change $\rho(p)$ except possibly at the i -th index. If $i \in \sigma$, we move p so that it lies in the interior of U_i ; if $i \in [k] \setminus \tau$ we move p so that it lies in the complement of the closure of U_i . Otherwise, we do not move p . This yields a point whose first i indices have the correct symbols, and by performing this operation for all $i \in [k]$ we obtain a point with the desired $\rho(p)$, proving the result. \square

The following lemma allows us to perform a (geometric) k -piercing in \mathbb{R}^d so that all previously pierceable intervals remain pierceable, and all new intervals are also pierceable, provided that $k < d$. By applying this lemma repeatedly, we will prove Lemma 22, establishing a slightly stronger version of part (I) of Theorem 13.

Lemma 21. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a code, and let $1 \leq k < d$. Suppose that the neuron n is a k -piercing of \mathcal{C} with associated interval $[\sigma, \tau]$, and let $\mathcal{D} = \mathcal{C} \setminus n$. If \mathcal{D} has a well-formed realization $\{U_1, \dots, U_{n-1}\}$ by open balls in \mathbb{R}^d , and $[\sigma, \tau]$ is pierceable at point p in this realization, then for any sufficiently small open ball U_n centered at p the following hold:*

- (i) *$\{U_1, \dots, U_n\}$ is a well-formed realization of \mathcal{C} by open balls,*
- (ii) *every interval in \mathcal{D} that was pierceable in the original realization is also pierceable in this new realization, and*
- (iii) *every interval in \mathcal{C} whose top element contains n is pierceable in this new realization.*

Proof. Item (i) is immediate since $[\sigma, \tau]$ is pierceable at p . The point p must lie on the common intersection of the boundaries of all U_i with $i \in \tau \setminus \sigma$. This common intersection is a sphere S of dimension at least one since $k < d$. The point p lies in the interior of all U_j with $j \in \sigma$, and in the complement of the closure of all U_j with $j \in [n] \setminus \tau$. The same is true

for all points in a small neighborhood of p , and hence S contains an infinite set of points at which $[\sigma, \tau]$ is pierceable. For every interval $[\sigma', \tau']$ in \mathcal{D} that was pierceable in our original realization (including the interval $[\sigma, \tau]$ itself), choose a point $p' \neq p$ at which this interval was pierceable. Choosing U_n small enough that its closure avoids all such p' guarantees item (ii); this can be done as there are only finitely many p' .

For item (iii), define $\gamma = (\tau \cup \{n\}) \setminus \sigma$ and observe that the U_i with i in γ comprise a collection of $k+1 \leq d$ independent balls. Each interval $[\sigma', \tau']$ in \mathcal{C} with $n \in \tau'$ must be a subset of $[\sigma, \tau \cup \{n\}]$. Up to removing the indices in σ , such an interval corresponds to an interval in 2^γ , which is pierceable by Lemma 20. Since $n \in \tau$, the point where this interval is pierceable lies in the closure of U_n . In particular, this point lies in the interior of all U_i with $i \in \sigma$, and in the complement of the closure of all U_j with $j \in [n-1] \setminus \tau$. Hence the interval $[\sigma', \tau']$ is pierceable at this point in our new realization, proving the result. \square

Lemma 22. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a k -inductively pierced code. Then \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^{k+1} , with the following property: every interval of rank at most k in \mathcal{C} is pierceable at some point.*

Proof. We may assume that the natural linear order on n is a piercing order, and proceed by induction on n , applying Lemma 21 at each step. The base case $n = 1$ is immediate by choosing any open ball U_1 in \mathbb{R}^{k+1} . For the inductive step, one must only note that the intervals described in items (ii) and (iii) account for all intervals of rank at most k in \mathcal{C} . Indeed, when performing a k -piercing, the only new intervals of rank at most k that are introduced to the code are those of the form $[\sigma', \tau' \cup \{n\}]$, which are covered by case (iii), while case (ii) covers all previously introduced intervals of rank at most k . \square

We now turn our attention to the task of performing geometric k -piercings in \mathbb{R}^k . Since the intersection of k -many independent spheres in \mathbb{R}^k is a 0-dimensional sphere consisting of only two points, not every k -piercing can be realized geometrically in \mathbb{R}^k . However, when a code is splittable (recall Definition 12), certain k -piercings can be achieved. We say that an interval $[\sigma, \tau]$ in a splittable code \mathcal{C} is *accessible* if it has rank less than k , or if it has rank exactly k and the following holds: the intersection of σ with the attaching set $A \sqcup B$ of $\tau \setminus \sigma$ is equal to A or B .

Lemma 23. *Let $k \geq 1$, let $\mathcal{C} \subseteq 2^{[n]}$ be an inductively k -pierced code, and suppose that \mathcal{C} is splittable. Then \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^k , with the following property: every accessible interval in \mathcal{C} is pierceable at some point.*

Proof. By relabeling, we can reduce to the case where the natural linear order on $[n]$ is a piercing order. Similarly to the proof of Lemma 22, we will proceed by induction on n . For the base case $n = 1$, any open ball U_1 in \mathbb{R}^k has every interval pierceable, and the result follows. For $n \geq 2$, define $\mathcal{D} \stackrel{\text{def}}{=} \mathcal{C} \setminus n$, let $[\sigma, \tau]$ be the interval associated to the k -piercing of \mathcal{C} by n , and let $\{U_1, \dots, U_{n-1}\}$ be a realization of \mathcal{D} satisfying the conclusion of the lemma. If $[\sigma, \tau]$ has rank less than k (i.e. n is a k' -piercing of \mathcal{C} with $k' < k$) then Lemma 21 yields a realization of \mathcal{C} with the desired properties.

This leaves the case that $[\sigma, \tau]$ has rank exactly k . In this case, let $A \sqcup B$ be the attaching set of $\tau \setminus \sigma$ in \mathcal{C} , and assume without loss of generality that $n \in A$. Since \mathcal{C} is splittable and n is minimal in $P(\mathcal{C})$, n is less than every $a \in A$ and is not comparable to any $b \in B$. It follows that σ contains A and no elements of B . Noting that the attaching set of $\tau \setminus \sigma$ in \mathcal{D} is $A \setminus \{n\} \sqcup B$, we see that the interval $[\sigma, \tau]$ is accessible in \mathcal{D} , and by inductive hypothesis there exists a point p where this interval is pierceable in the realization $\{U_1, \dots, U_{n-1}\}$.

Consider a small open ball U_n centered at p . By making U_n small enough, we can guarantee that $\{U_1, \dots, U_n\}$ is a well-formed realization of \mathcal{C} by open balls, and that every interval in \mathcal{D} that was pierceable remains pierceable, except possibly the interval $[\sigma, \tau]$ itself. Moreover, Lemma 20 guarantees that any new intervals $[\sigma', \tau']$ of rank k or less with $n \in \tau'$ are pierceable in our new realization as well. The only remaining interval in \mathcal{C} which could be accessible is $[\sigma, \tau]$. Observe that this interval is accessible in \mathcal{C} if and only if $B = \emptyset$. In this case, the 0-dimensional sphere which is the intersection of the boundaries of all U_i indexed by $\tau \setminus \sigma$ will consist of p and a point q , and the interval $[\sigma, \tau]$ will be pierceable at both these points in the realization of \mathcal{D} . The set U_n must avoid the point q , lest the spheres indexed by $(\tau \setminus \sigma) \cup \{n\}$ not form an independent collection, and so $[\sigma, \tau]$ remains pierceable at q in our constructed realization of \mathcal{C} . This proves the result. \square

The above lemmas give the required constructive results for Theorem 13, and it remains to establish the theorem's restrictions on well-formed realizations of inductively pierced codes—namely that if \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^k then \mathcal{C} must be splittable, and that realizations in \mathbb{R}^{k-1} are not possible.

Lemma 24. *Let $\mathcal{C} \subseteq 2^{[n]}$ be an inductively pierced code, suppose that the largest clique in $G(\mathcal{C})$ has size $k+1$, and suppose that \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^k . Then \mathcal{C} is splittable.*

Proof. Let σ be the set of vertices of a clique of size k in $G(\mathcal{C})$. In any well-formed realization of \mathcal{C} in \mathbb{R}^k by balls, the intersection of the boundaries of the balls indexed by σ will be a 0-dimensional sphere consisting of two points, p and q . The balls indexed by the attaching set of σ must each contain exactly one of p or q in their closure. Let A be the set of neurons whose associated open balls contain p , and let B be the neurons whose associated balls contain q . Clearly $S = A \sqcup B$.

It remains to argue that $P(\mathcal{C})$ is a total order on A and B respectively, and that no element of A is comparable to any element of B . To see that $P(\mathcal{C})$ is a total order on A , first recall that the attaching set of σ is an independent set in $G(\mathcal{C})$. Hence if $i, j \in A$, then $\text{CF}(J_{\mathcal{C}})$ contains one of the pseudo-monomials $x_i(1 - x_j)$, $x_j(1 - x_i)$, or $x_i x_j$. Since U_i and U_j share the point p , the pseudo-monomial $x_i x_j$ does not occur in $\text{CF}(J_{\mathcal{C}})$, and we conclude that i and j are comparable in $P(\mathcal{C})$. Since any pair of elements in A are comparable, $P(\mathcal{C})$ is a total order on A . A symmetric argument establishes that $P(\mathcal{C})$ is a total order on B . Finally, note that if $i \in A$ and $j \in B$, then i and j are not comparable in $P(\mathcal{C})$ because U_i contains p while U_j does not, and symmetrically U_j contains q while U_i does not. \square

Lemma 25. *Let $\mathcal{C} \subseteq 2^{[n]}$ be an inductively pierced code, and suppose $G(\mathcal{C})$ contains a clique of size $k + 1$. Then \mathcal{C} does not have a well-formed realization by open balls in \mathbb{R}^{k-1} .*

Proof. By relabeling we may assume that $G(\mathcal{C})$ has a clique on $[k + 1]$. The fact that \mathcal{C} is inductively pierced implies that the restriction of \mathcal{C} to $[k + 1]$ is the full power set $2^{[k+1]}$. Indeed, up to ignoring the indices outside of $[k + 1]$, any piercing order successively builds sub-intervals of $2^{[k+1]}$ with ranks $0, 1, \dots, k + 1$, each of which are contained in \mathcal{C} . We conclude that if \mathcal{C} had a well-formed realization by open balls in \mathbb{R}^{k-1} , then the balls indexed by $[k + 1]$ would form an independent collection. But the largest collection of independent balls in \mathbb{R}^{k-1} has size k . This proves the result. \square

Theorem 13. *Let $\mathcal{C} \subseteq 2^{[n]}$ be an inductively pierced code, and let $k \geq 1$. If the largest clique in $G(\mathcal{C})$ has size $k + 1$, then*

- (I) \mathcal{C} has a well-formed realization by open balls in \mathbb{R}^{k+1} ,
- (II) \mathcal{C} has such a realization in \mathbb{R}^k if and only if \mathcal{C} is splittable, and
- (III) \mathcal{C} does not have such a realization in \mathbb{R}^{k-1} .

Proof. Lemma 22 established a realization of the desired form for (I). Lemmas 23 and 24 established the reverse and forward implications needed for (II) respectively. Finally, Lemma 25 established (III). \square

4 Computing Piercing Orders and Concluding Remarks

Motivated by the work of Gross, Obatake, and Youngs [9], we have provided a complete characterization of inductively pierced codes. Additionally, we have shown that inductively pierced codes can be realized with open balls (and hence are convex), and we have exactly characterized the smallest dimension in which such realizations exist. A consequence of our characterization is that an inductively pierced code $\mathcal{C} \subseteq 2^{[n]}$ is not only convex, but can be realized with convex sets in dimension $n - 1$ or less. This strongly contrasts the general situation for convex codes: Jeffs [10] showed that there are codes $\mathcal{C} \subseteq 2^{[n]}$ which can only be realized in dimensions that are *exponential* in terms of n .

Within our results and their proofs, we have developed an implicit algorithm for a) recognizing if a code is inductively pierced, b) computing a piercing order, c) determining the minimal dimension in which the code may be realized with open balls, and d) producing such a realization. We summarize this algorithm below. Recall that the codes we work with satisfy the mild requirements (1)–(3) outlined in the introduction.

Algorithm for computing a piercing order.

1. Compute $\text{CF}(J_{\mathcal{C}})$.

Note: Current algorithms for this are not necessarily efficient—see for example the discussion by Garcia, Puente, Kruse, Liu, Miyata, Petersen, Phillipson, and Shiu [8,

Page 1]. We conjecture below that $\text{CF}(J_{\mathcal{C}})$ can be computed efficiently when \mathcal{C} is inductively pierced, even if we do not know this a priori.

2. Verify that \mathcal{C} is degree two. If not, then \mathcal{C} is not inductively pierced.
3. Form $G(\mathcal{C})$ and $P(\mathcal{C})$ from $\text{CF}(J_{\mathcal{C}})$.
4. Search for an elimination neuron $i \in [n]$ (recall Definition 7). This can be done in $O(n^2)$ time by computing the neighborhood of each neuron in $G(\mathcal{C})$ and checking whether it is simplicial, and then comparing all simplicial vertices according to $P(\mathcal{C})$ to find a minimal one. If no elimination neuron exists, then \mathcal{C} is not inductively pierced.
5. Perform a permutation of $[n]$ so that n is our elimination neuron. Record this permutation and the clique consisting of n and its neighbors.
6. Delete n from $G(\mathcal{C})$ and $P(\mathcal{C})$ to obtain $G(\mathcal{C} \setminus n)$ and $P(\mathcal{C} \setminus n)$ respectively. Repeat from Step 4 until no neurons remain, tracking the permutations from each time that we perform in Step 5.
7. Compose the recorded permutations and apply to the original linear ordering to obtain a piercing order for \mathcal{C} . The smallest value of k for which \mathcal{C} is k -inductively pierced is size of the largest clique recorded in an instance of Step 5, minus one.
8. Determine if \mathcal{C} is splittable. Each clique of size k will be contained in one of the $(k+1)$ -cliques recorded in Step 5, so there are at most $n(k+1)$ such cliques to check. For each, we can compute its attaching set and check whether it is possible to partition it as in Definition 12.
9. If \mathcal{C} is splittable, we form a realization in \mathbb{R}^k . Otherwise, we form a realization in \mathbb{R}^{k+1} . To realize \mathcal{C} by open balls, sequentially add open balls according to the piercing order. Lemmas 21, 22, and 23 determine the points on the boundaries of these balls where subsequent balls should be added to form the appropriate code.

Conjecture 26. *Fix $d \geq 2$. Given a code $\mathcal{C} \subseteq 2^{[n]}$, there is a polynomial-time algorithm to determine whether or not every pseudo-monomial in $\text{CF}(J_{\mathcal{C}})$ has degree at most d . Here, the input to the algorithm is a list of the codewords in \mathcal{C} , with each codeword given as a binary vector of length n .*

Note that when $\text{CF}(J_{\mathcal{C}})$ consists of pseudo-monomials with degree bounded by a fixed constant d , there is a polynomial-time algorithm for determining $\text{CF}(J_{\mathcal{C}})$. Indeed, the total number of pseudo-monomials in $\mathbb{F}_2[x_1, \dots, x_n]$ of degree at most d is equal to $2^d \binom{n}{d} + 2^{d-1} \binom{n}{d-1} + \dots + 2d + 1 = O(n^d)$, and so we could simply evaluate each of these pseudo-monomials on all codewords to determine which ones lie in $J_{\mathcal{C}}$, and then compare them pairwise to find the minimal ones, which will comprise $\text{CF}(J_{\mathcal{C}})$. Hence if this conjecture is true, then our results imply that there is a polynomial-time algorithm for determining whether or not a code is inductively pierced, even if we are not initially given $\text{CF}(J_{\mathcal{C}})$.

Concluding Remarks. Our work proves that inductively pierced codes must be realizable using open balls (and thus are convex), but there are codes that are not inductively pierced but nevertheless admit a realization by open balls. Indeed, there are such codes that are not degree two, and there are also such codes that are degree two but have $G(\mathcal{C})$ non-chordal. The code $\mathcal{C} = \{12, 13, 23, 1, 2, 3, \emptyset\}$ is a simple example of the first case, as its canonical form is $\{x_1x_2x_3\}$, but it can be realized in two dimensions using open balls, as in Figure 5. For an example of the second case, consider $\mathcal{D} = \{12, 14, 23, 34, 1, 2, 3, 4, \emptyset\}$, which has canonical form $\{x_1x_3, x_2x_4\}$ and hence its associated graph is a 4-cycle, which is not chordal. However, \mathcal{D} has a realization by open balls in \mathbb{R}^2 as shown in Figure 5.

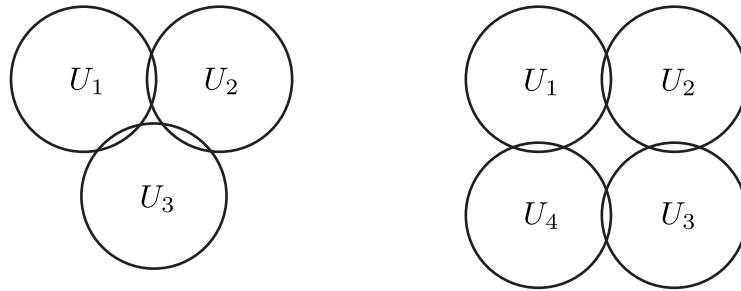


Figure 5: Two codes that are not inductively pierced, but can be realized with well-formed collections of open balls in the plane. (Left) The code \mathcal{C} is not degree two. (Right) The graph $G(\mathcal{D})$ is not chordal.

It seems plausible to seek larger classes of codes that admit realizations that are constructed iteratively from open balls. However, codes such as \mathcal{D} above, whose realizations cover non-contractible subsets of \mathbb{R}^d , will pose a challenge to such an approach. Had we arranged U_1, U_2 , and U_3 in a vertical line, then there would have been no way to add an appropriate open ball (or any convex set) U_4 at the final step. This highlights the special structure of inductively pierced codes, whose realizations do not create loops or other topologically interesting features.

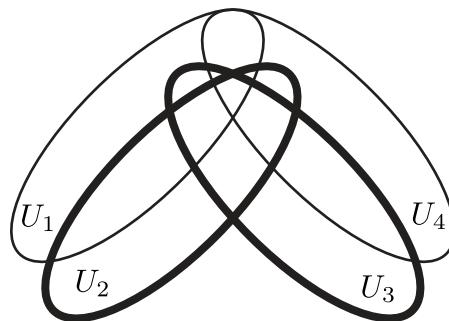


Figure 6: An inductively pierced code realized by ellipsoidal regions in the plane. This code cannot be realized by open balls in the plane.

Finally, note that when realizations are not restricted to open balls, our characterization of realization dimension does not apply, even when a code is inductively pierced. For example, consider the complete code on four neurons, $\mathcal{C} = 2^{[4]}$. This code is inductively pierced—its canonical form is empty and thus trivially meets the degree two condition, and its graph $G(\mathcal{C})$ is the complete graph K_4 and thus is chordal. Theorem 13 implies that \mathcal{C} is not realizable in \mathbb{R}^2 using open balls, but it does have realization in \mathbb{R}^2 using ellipsoidal regions, as shown in Figure 6. The principal difference between this realization and any realization using open balls is that the intersection of two ellipses need not be a lower-dimensional ellipse. In the figure, the sets U_2 and U_3 “pass through” one another, and their boundaries share four common points. Accounting for such possibilities renders it more difficult to formulate simple inductive constructions for realizations by ellipses. A characterization of minimal dimension for realizations of inductively pierced codes using ellipses (or other types of convex regions) remains an open question.

References

- [1] J. Anusiak. On set-theoretically independent collections of balls. *Colloquium Mathematicum*, 13(2):223–233, 1964.
- [2] Patrick Chan, Katherine Johnston, Joseph Lent, Alexander Ruys de Perez, and Anne Shiu. Nondegenerate neural codes and obstructions to closed-convexity. *SIAM Journal on Discrete Mathematics*, 37(1):114–145, 2023.
- [3] Aaron Chen, Florian Frick, and Anne Shiu. Neural codes, decidability, and a new local obstruction to convexity. *SIAM Journal on Applied Algebra and Geometry*, 3(1):44–66, 2019.
- [4] Joshua Cruz, Chad Giusti, Vladimir Itskov, and Bill Kronholm. On open and closed convex codes. *Discrete & Computational Geometry*, 61:247–270, 2016.
- [5] Carina Curto, Elizabeth Gross, Jack Jeffries, Katherine Morrison, Mohamed Omar, Zvi Rosen, Anne Shiu, and Nora Youngs. What makes a neural code convex? *SIAM Journal on Applied Algebra and Geometry*, 1(1):222–238, 2017.
- [6] Carina Curto, Vladimir Itskov, Alan Veliz-Cuba, and Nora Youngs. The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes. *Bulletin of Mathematical Biology*, 75(9):1571–1611, 2013.
- [7] Brianna Gambacini, R. Amzi Jeffs, Sam Macdonald, and Anne Shiu. Non-monotonicity of closed convexity in neural codes. *Vietnam Journal of Mathematics*, 2021.
- [8] Rebecca Garcia, Luis Garcia-Puente, Ryan Kruse, Jessica Liu, Dane Miyata, Ethan Petersen, Kaitlyn Phillipson, and Anne Shiu. Gröbner bases of neural ideals. *International Journal of Algebra and Computation*, 28(4):553–571, 2018.
- [9] Elizabeth Gross, Nida Kazi Obatake, and Nora Youngs. Neural ideals and stimulus space visualization. *Advances in Applied Mathematics*, 95:65–95, 2018.

- [10] R. Amzi Jeffs. Embedding dimension phenomena in intersection complete codes. *Selecta Math. (N.S.)*, 28(1):Paper No. 18, 37, 2022.
- [11] R. Amzi Jeffs. Open, closed, and non-degenerate embedding dimensions of neural codes. *Discrete & Computational Geometry*, 2023. (Accepted to appear).
- [12] R. Amzi Jeffs, Caitlin Lienkaemper, and Nora Youngs. Order-forcing in neural codes. *arXiv e-prints: 2008.13192*, 2020.
- [13] Alexander Kunin, Caitlin Lienkaemper, and Zvi Rosen. Oriented matroids and combinatorial neural codes. *Combinatorial Theory*, 3(1), 2023.
- [14] Caitlin Lienkaemper, Anne Shiu, and Zev Woodstock. Obstructions to convexity in neural codes. *Advances in Applied Mathematics*, 85:31–59, 2017.
- [15] Ezra Miller and Bernd Sturmfels. *Combinatorial commutative algebra*, volume 227 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.
- [16] A. Rényi, C. Rényi, and J. Surányi. Sur l’indépendance des domaines simples dans l’espace Euclidien à n dimensions. *Colloq. Math.*, 2:130–135, 1951.
- [17] Richard P. Stanley. *Combinatorics and commutative algebra*, volume 41 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 1996.
- [18] Douglas B. West. *Combinatorial Mathematics*. Cambridge University Press, 2021.