Open and closed convex codes and their embedding dimensions

R. Amzi Jeffs
University of Washington, Seattle

CoSP Seminar
July 23rd 2020
Part I: Convex Codes

\[ C = \{123, 12, 23, 2, 3, \emptyset\} \]
What is a Code?

Definition

A code is any subset of $2^{[n]}$. We will call the indices $i \in [n]$ neurons. Elements of a code are called codewords. We will often partially order codewords by inclusion, and speak of maximal codewords.

Every code $C$ has an associated simplicial complex

$$\Delta(C) = \{ \sigma \subseteq [n] \mid \sigma \subseteq c \in C \}.$$ 

Consider the code $C = \{123, 12, 23, 2, 3, \emptyset\}$

- It has 3 neurons.
- It has 6 codewords.
- $\Delta(C)$ is a 2-simplex generated by the unique maximal codeword 123.
Convex Codes and Realizations

**Definition**

Let $\mathcal{U} = \{U_1, \ldots, U_n\}$ be a collection of convex open (resp. closed) sets in $\mathbb{R}^d$. Define

$$\text{code}(\mathcal{U}) = \{\sigma \subseteq [n] \mid p \in U_i \iff i \in \sigma \text{ for some } p \in \mathbb{R}^d\}.$$ 

The collection $\mathcal{U}$ is a open (resp. closed) realization of $\mathcal{C} = \text{code}(\mathcal{U})$. 

\[ \mathcal{C} = \{123, 12, 23, 2, 3, \emptyset\} \]
When Do Realizations Exist?

Definition

Let $C \subseteq 2^{\lfloor n\rfloor}$ be a code. The *open embedding dimension* of $C$ is

$$\text{odim}(C) \overset{\text{def}}{=} \min\{d \mid C \text{ has an open realization in } \mathbb{R}^d\}.$$  

When no realization exists $\text{odim}(C) = \infty$. The *closed embedding dimension* $\text{cdim}(C)$ is defined analogously.

Example: Recall we can realize $C = \{123, 12, 23, 2, 3, \emptyset\}$ in $\mathbb{R}^2$. In fact, $\text{odim}(C) = \text{cdim}(C) = 1$ since we can “flatten” our realization.
When Do Realizations Exist?

Definition

Let \( C \subseteq 2^{[n]} \) be a code. The open embedding dimension of \( C \) is

\[
\text{odim}(C) \overset{\text{def}}{=} \min\{d \mid C \text{ has an open realization in } \mathbb{R}^d\}.
\]

When no realization exists \( \text{odim}(C) = \infty \). The closed embedding dimension \( \text{cdim}(C) \) is defined analogously.

Example: Can you find a realization of \( C = \{123, 1, 2, 3, \emptyset\} \) in \( \mathbb{R}^2 \)? How about \( \mathbb{R}^1 \)?
In this case \( \text{odim}(C) = \text{cdim}(C) = 2 \).
When Do Realizations Exist?

**Definition**

Let \( C \subseteq 2^{[n]} \) be a code. The *open embedding dimension* of \( C \) is

\[
\text{odim}(C) \overset{\text{def}}{=} \min\{d \mid C \text{ has an open realization in } \mathbb{R}^d\}.
\]

When no realization exists \( \text{odim}(C) = \infty \). The *closed embedding dimension* \( \text{cdim}(C) \) is defined analogously.

**This talk:** Aim to understand embedding dimension of *intersection complete* codes.
Part II: Past Work and Appetizers
Lemma (Borsuk’s Nerve Lemma)

Let $\mathcal{U} = \{U_1, \ldots, U_n\}$ be a (open or closed) realization of $\mathcal{C}$. Then $\bigcup_{i \in [n]} U_i$ is homotopy equivalent to $\Delta(\mathcal{C})$.

Corollary

Whenever $\Delta(\mathcal{C})$ has a $d$-dimensional hole, $\text{odim}(\mathcal{C}) \geq d$ and $\text{cdim}(\mathcal{C}) \geq d$.

There are further topological lower bounds via Betti numbers in [CV16], and connections to Helly’s theorem.
A Few Upper Bounds

**Theorem**

If $\mathcal{C}$ is intersection complete and has $k$ codewords, then $\text{odim}(\mathcal{C}) \leq k - 1$ and $\text{cdim}(\mathcal{C}) \leq k - 1$.

**Theorem (CGIK16)**

If $\mathcal{C}$ is intersection complete and has $m \geq 3$ maximal codewords, then $\text{odim}(\mathcal{C}) \leq m - 1$ and $\text{cdim}(\mathcal{C}) \leq m - 1$.

**Theorem (CGIK16)**

If $\mathcal{C} \subseteq 2^{[n]}$ is intersection complete, then $\text{cdim}(\mathcal{C}) \leq n - 1$. 
Comparing $\text{odim}$ and $\text{cdim}$

**Theorem** ("Folklore", proof in J19)

If $\mathcal{C} = \Delta(\mathcal{C})$, then $\text{odim}(\mathcal{C}) = \text{cdim}(\mathcal{C})$.

**Proof.**

When $\mathcal{C} = \Delta(\mathcal{C})$, perturbing sets in a realization is okay.

- Open realization $\rightarrow$ shrink sets $\rightarrow$ take closures.
- Closed realization $\rightarrow$ assume compactness $\rightarrow$ add $\varepsilon$-ball to sets.
Comparing $\text{odim}$ and $\text{cdim}$

**Theorem (J19)**

If $\mathcal{C}$ is intersection complete, then $\text{cdim}(\mathcal{C}) \leq \text{odim}(\mathcal{C})$.

**Proof.**

**Key fact:** If $\mathcal{C}$ is intersection complete, then $\bigcap_{i \in \sigma} U_i \subseteq \bigcup_{i \in \tau} U_i$ implies $U_\sigma \subseteq U_i$ for some $i \in \tau$.

This allows us shrink our $U_i$ and then take closures.

Implicit in [LSW15]: Sometimes $\text{cdim}(\mathcal{C}) < \text{odim}(\mathcal{C})$. But how big can the gap be? We’ll see later on.
Part III: Closed Embedding Dimension
Theorem (J19)

If $\mathcal{C} \subseteq 2^{[n]}$ is intersection complete and $d = \dim(\Delta(\mathcal{C}))$, then

$$\text{cdim}(\mathcal{C}) \leq 2d + 1.$$ 

Note: Compare to bound $\text{cdim}(\mathcal{C}) \leq n - 1$ from [CGIK16].

When $\mathcal{C} = \Delta(\mathcal{C})$, this is almost immediate from classical (circa 70s) discrete geometry results about nerve complexes and $d$-representability. See survey paper [Tan11].

We will re-apply these techniques.
Reminder: Polytopes

A polytope is the convex hull of finitely many points. Polytopes have faces which form a face poset. Faces that are points are called vertices. Every polytope has a dual, which turns the face poset upside down.

Definition

A polytope is *d-neighborly* if every *d* vertices form a face.
Building a Closed Realization

$(d + 1)$-neighborly in $\mathbb{R}^{2d+2}$

Dual in $\mathbb{R}^{2d+2}$

Schlegel diagram in $\mathbb{R}^{2d+1}$

Trim sets to get closed realization in $\mathbb{R}^{2d+1}$
Part IV: Open Embedding Dimension and Sunflowers of Convex Open Sets
Sunflowers of Convex Open Sets

**Definition**

\( U \) is a sunflower if \( \text{code}(U) = \{[n], \text{singletons}, \emptyset\} \). The \( U_i \) are called *petals* and \( U_1 \cap U_2 \cap \cdots \cap U_n \) is called the *center*. 
Lemma (LSW15)

If $\mathcal{U} = \{U_1, U_2, U_3\}$ is a sunflower in $\mathbb{R}^2$, and $p_i \in U_i$, then $\text{conv}\{p_1, p_2, p_3\}$ contains a point in the center of $\mathcal{U}$. This fails in $\mathbb{R}^3$. 
Rigidity in General

**Theorem (J18)**

If $\mathcal{U} = \{U_1, U_2, \ldots, U_{d+1}\}$ is a sunflower in $\mathbb{R}^d$, and $p_i \in U_i$, then $\text{conv}\{p_1, p_2, \ldots, p_{d+1}\}$ contains a point in the center of $\mathcal{U}$. This fails in $\mathbb{R}^{d+1}$.

**A word on the proof:** Original proof quite messy. A nicer proof uses Radon’s theorem, as suggested by Zvi Rosen.

**Note:** Result fails badly when the sets are not open.
Application to Codes

**Theorem (J19)**

Let $S_n \subseteq 2^{[n+1]}$ be the code whose codewords are every singleton, every 2-set containing $n+1$, the empty set, and $[n]$. Then $\odim(S_n) = n$.

**Proof.**

The collection $\{U_1, \ldots, U_n\}$ is a sunflower. The set $U_{n+1}$ touches every petal but not the center. Impossible when $d < n$.

On the other hand, can get a realization in $\mathbb{R}^n$ by thickening coordinate axes and hyperplane with normal vector $1$.

Example: $S_3$ in $\mathbb{R}^3$
More Sunflowers

Definition (J19)

Let $\Delta \subseteq 2^{[n]}$ be a simplicial complex. Define a code $S_\Delta \subseteq 2^{[n+1]}$ by

$$S_\Delta \overset{\text{def}}{=} (\Delta \ast (n + 1)) \cup \{[n]\}.$$

Theorem (J19)

Suppose $\Delta$ has $m$ facets. Then we may use a realization of $S_\Delta$ in $\mathbb{R}^d$ to build a realization of $S_m$ in $\mathbb{R}^d$. 
Embedding Dimension for $S_\Delta$

**Corollary (J19)**

If $\Delta$ has $m \geq 2$ facets, then $\text{odim}(S_\Delta) = m$.

**Proof.**

Results of [CGIK16] let us construct a realization of $S_\Delta$ in $\mathbb{R}^m$. On the other hand, $\text{odim}(S_\Delta) \geq \text{odim}(S_m) = m$. □

**Corollary (J19)**

Open embedding dimension may grow *exponentially* in terms of number of neurons $n$, as large as $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$. Surprising!

**Punchline:** Open and closed embedding dimensions differ hugely. $\min\{2d + 1, n - 1\}$ as compared to $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$. 
Part V: The Landscape Ahead, Open Questions
Code Minors

Notation: If $\sigma \subseteq [n]$, then $U_\sigma = \bigcap_{i \in \sigma} U_i$.

**Definition**

Let $C$ be a code with a realization $\mathcal{U} = \{U_1, \ldots, U_n\}$. A minor of $C$ is a code that has a realization using sets of the form $U_\sigma$.

**Nice fact:** Minors can be described totally combinatorially.

**Proposition**

If $D$ is a minor of $C$, then $\text{odim}(D) \leq \text{odim}(C)$ and $\text{cdim}(D) \leq \text{cdim}(C)$.

**Theorem (J18b)**

The relation “is a minor of” forms a partial order on all (appropriate equivalence classes of) codes, yielding a poset $\mathbb{P}_{\text{Code}}$. 
Theorem (J19)
Suppose that $\Delta$ has $m$ facets. Then $S_m$ is a minor of $S_\Delta$.

Observation: odim and cdim are monotone in $P_{\text{Code}}$, but number of neurons is not!
The Terrain for Future Work in P\text{Code}

- cdim < odim = \infty
  - Many examples via "local obstructions"
  - Many examples via sunflowers

- odim < cdim = \infty
  - Six known examples

- cdim < odim
  - Intersection complete codes
    - Simplicial complexes

- odim = cdim

- odim < \infty
- cdim < \infty

- cdim = \infty
Open Questions

Question
Is $\text{cdim}(\mathcal{C})$ ever larger than $n - 1$?

Question
Can you find $\mathcal{C}$ with $\text{odim}(\mathcal{C}) < \text{cdim}(\mathcal{C}) < \infty$?

Question
Characterize the embedding dimension of “tangled sunflowers”? (See [J19]).

$t_1 = 1$  $t_2 = 2$  $t_3 = 3$  $t_4 = 3$
References

This talk is based on [J19]:

“Embedding dimension phenomena in intersection complete codes”

Other papers referenced:

- [CGIK16] “On open and closed convex codes”
  https://arxiv.org/abs/1609.03502
- [CV16] “The leruley dimension of a convex code”
  https://arxiv.org/abs/1612.07797
- [J18] “Sunflowers of convex open sets”
  https://arxiv.org/abs/1810.03741
- [LSW15] “Obstructions to convexity in neural codes”
  https://arxiv.org/abs/1509.03328
- [Tan11] “Intersection patterns of convex sets via simplicial complexes, a survey”
  https://arxiv.org/abs/1102.0417