

Realizing convex codes with axis-parallel boxes

Miguel Benitez* Siran Chen* Tianhui Han*
 R. Amzi Jeffs*[†] Kinapal Paguyo* Kevin A. Zhou*

March 2023

Abstract

Every ordered collection of sets in Euclidean space can be associated to a combinatorial code, which records the regions cut out by the sets in space. Given two ordered collections of sets, one can form a third collection in which the i -th set is the Cartesian product of the corresponding sets from the original collections. We prove a general “product theorem” which characterizes the code associated to the collection resulting from this operation, in terms of the codes associated to the original collections. We use this theorem to characterize the codes realizable by axis-parallel boxes, and exhibit differences between this class of codes and those realizable by convex open or closed sets. We also use our theorem to prove that a “monotonicity of open convexity” result of Cruz, Giusti, Itskov, and Kronholm holds for closed sets when some assumptions are slightly weakened.

1 Introduction

Given a collection $\mathcal{U} = \{U_1, \dots, U_n\}$ of sets in $\mathbb{R}^d$, one can record how they intersect and cover each other using a combinatorial code on the index set $[n]$, defined as follows:

$$\text{code}(\mathcal{U}) \stackrel{\text{def}}{=} \{\sigma \subseteq [n] \mid \text{there exists } p \in \mathbb{R}^d \text{ with } p \in U_i \text{ if and only if } i \in \sigma\}.$$

We say that $\mathcal{U}$ is a *realization* of $\text{code}(\mathcal{U})$, and elements of a code are called *codewords*. A codeword is *maximal* if it is not properly contained in any other codeword.

For notational convenience, we define the *intersection pattern* of $\mathcal{U}$ at a point p to be the set $\text{patt}_{\mathcal{U}}(p) \stackrel{\text{def}}{=} \{i \in [n] \mid p \in U_i\}$. For $\sigma \subseteq [n]$ the *atom* of σ in $\mathcal{U}$ is $\mathcal{U}^\sigma \stackrel{\text{def}}{=} \{p \in \mathbb{R}^d \mid \text{patt}_{\mathcal{U}}(p) = \sigma\}$. With this notation, $\text{code}(\mathcal{U})$ consists of the sets $\sigma \subseteq [n]$ with nonempty atoms. Figure 1 shows a realization of a code by four axis-parallel rectangles in $\mathbb{R}^2$. Throughout, we write maximal codewords in bold.

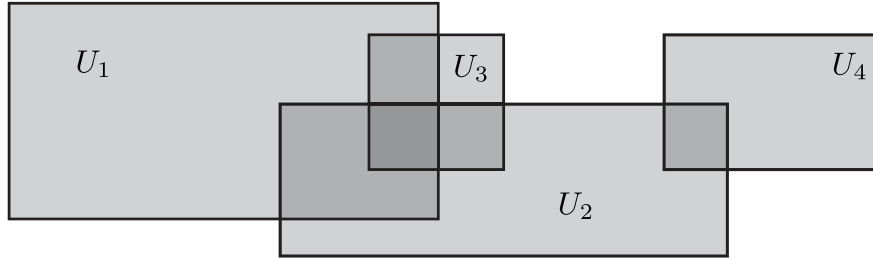


Figure 1: Four axis-parallel rectangles in $\mathbb{R}^2$ which realize the code $\{\mathbf{123}, \mathbf{24}, 12, 13, 23, 1, 2, 3, 4, \emptyset\}$.

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

[†]Supported by the National Science Foundation through Award No. 2103206.

Convex codes. Given a collection of sets, it is straightforward to compute its associated code. The inverse problem—finding a collection of sets that realize a given code—may be complicated, especially when geometric constraints are placed on the sets in the realization. In 2013, Curto, Itskov, Veliz-Cuba, and Youngs [5] initiated the study of *open convex codes*, which are codes with realizations consisting of convex open sets. One may also consider *closed convex codes*, which have realizations by closed convex sets. Curto, Itskov, Veliz-Cuba, and Youngs were motivated by neuroscientific phenomena, and the study of open and closed convex codes additionally generalizes the more classical study of d -representable complexes (see Tancer’s survey [15]). Since its inception the study of open and closed convex codes has thus led to rich results of purely mathematical interest, such as new discrete geometry theorems, families of extremal examples, and more.

Remark 1. Consistent with the convex codes literature, we only consider bounded sets in $\mathbb{R}^d$. This is equivalent to considering codes that contain $\emptyset$ as a codeword (if any sets are unbounded, we may intersect them with a large open or closed ball to obtain a bounded realization of the same code).

Beyond determining whether a code has an open or closed convex realization, it is natural to seek realizations that are minimal in dimension. To this end, it is useful to define the *open* (respectively, *closed*) *embedding dimension* of a code. These are

$$\begin{aligned} \text{odim}(\mathcal{C}) &\stackrel{\text{def}}{=} \min\{d \mid \mathcal{C} \text{ has a realization by open convex sets in } \mathbb{R}^d\}, \text{ and} \\ \text{cdim}(\mathcal{C}) &\stackrel{\text{def}}{=} \min\{d \mid \mathcal{C} \text{ has a realization by closed convex sets in } \mathbb{R}^d\}. \end{aligned}$$

By convention, the minimum over the empty set is ∞ .

A complete classification of open or closed convex codes and their embedding dimensions is currently out of reach, and an efficient classification seems to be out of the question: results of Kunin, Lienkaemper, and Rosen [12] posit that recognizing open convex codes in $\mathbb{R}^2$ is at least NP-hard. A further difficulty is the difference between open convex and closed convex realizations—Lienkaemper, Shiu, and Woodstock [13] described a code which has a closed convex realization, but no open convex realization, and Cruz, Giusti, Itskov, and Kronholm [2] described a code with the opposite behavior. Families of codes which exhibit arbitrary disparities between closed and open embedding dimension were given more recently by Jeffs [10], but general criteria for understanding and classifying these differences are lacking.

Box codes. Given the difficulties of classifying open and closed convex codes, we tackle the simpler problem of which codes can be realized by axis-parallel boxes. An *axis-parallel box* in $\mathbb{R}^d$ is either the empty set, or the Cartesian product of nonempty intervals $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$. For brevity, we use the term *box* from here on. Note that we are considering closed boxes, and we are allowing for boxes that are not full-dimensional—in particular, a point is a box. We say that a code is a *box code* or *box-convex* if it has a realization in $\mathbb{R}^d$ in which every set is an axis-parallel box. Lastly, we define the *box embedding dimension* of a code $\mathcal{C}$ to be

$$\text{bdim}(\mathcal{C}) \stackrel{\text{def}}{=} \min\{d \mid \mathcal{C} \text{ has a realization in } \mathbb{R}^d \text{ in which each set is a box}\}.$$

Note immediately that $\text{cdim}(\mathcal{C}) \leq \text{bdim}(\mathcal{C})$, as every box is a closed convex set. Later results will also imply that $\text{odim}(\mathcal{C}) \leq \text{bdim}(\mathcal{C})$ (see Theorem 16).

Our results. Our main result is Theorem 3, which applies to any collections of sets, but proves particularly useful for analyzing box-convex codes and their embedding dimensions. To state the result we first require a definition.

Definition 2. The *intersection product* of two codes $\mathcal{C} \subseteq 2^{[n]}$ and $\mathcal{D} \subseteq 2^{[n]}$ is the code

$$\mathcal{C} \boxtimes \mathcal{D} \stackrel{\text{def}}{=} \{c_1 \cap c_2 \mid c_1 \in \mathcal{C} \text{ and } c_2 \in \mathcal{D}\}.$$

Theorem 3 (Product theorem). *Let $\mathcal{U} = \{U_1, \dots, U_n\}$ and $\mathcal{V} = \{V_1, \dots, V_n\}$ be collections of (not necessarily open, closed, or convex) sets in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ respectively. Define $\mathcal{W} \stackrel{\text{def}}{=} \{W_1, \dots, W_n\}$ where $W_i \stackrel{\text{def}}{=} U_i \times V_i$. Then*

$$\text{code}(\mathcal{W}) = \text{code}(\mathcal{U}) \boxtimes \text{code}(\mathcal{V}).$$

Since boxes can be decomposed as Cartesian products of intervals, Theorem 3 will let us reduce the study of box-convex codes to the study of *interval codes* which have realizations in $\mathbb{R}^1$ in which every set is a closed interval. We give two immediate corollaries of Theorem 3 below.

Corollary 4. *For any codes $\mathcal{C} \subseteq 2^{[n]}$ and $\mathcal{D} \subseteq 2^{[n]}$, we have $\text{odim}(\mathcal{C} \boxtimes \mathcal{D}) \leq \text{odim}(\mathcal{C}) + \text{odim}(\mathcal{D})$. The analogous inequalities hold for $\text{cdim}(\mathcal{C} \boxtimes \mathcal{D})$ and $\text{bdim}(\mathcal{C} \boxtimes \mathcal{D})$.*

Proof. Note that if U_i and V_i are both open and convex, then so is $U_i \times V_i$. The same is true when the sets are closed and convex, or when they are boxes. Therefore, if $\mathcal{U}$ is an open (respectively, closed or box) convex realization of $\mathcal{C}$ in $\mathbb{R}^{d_1}$ and $\mathcal{V}$ is an open (respectively, closed or box) convex realization of $\mathcal{D}$ in $\mathbb{R}^{d_2}$, then the collection $\mathcal{W}$ of Theorem 3 is a realization of $\mathcal{C} \boxtimes \mathcal{D}$ in $\mathbb{R}^{d_1+d_2}$ of the same type. $\square$

Corollary 5. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a code. Then $\mathcal{C}$ is box-convex in $\mathbb{R}^d$ if and only if $\mathcal{C}$ is the intersection product of d -many interval codes.*

Proof. Let $\mathcal{B} = \{B_1, B_2, \dots, B_n\}$ be a realization of $\mathcal{C}$ by boxes in $\mathbb{R}^d$. Each set B_i can be represented as a Cartesian product of d -many intervals: $B_i = I_i^{(1)} \times I_i^{(2)} \times \dots \times I_i^{(d)}$. By Theorem 3, we know that $\text{code}(\mathcal{B}) = \text{code}(\mathcal{I}^{(1)}) \boxtimes \text{code}(\mathcal{I}^{(2)}) \boxtimes \dots \boxtimes \text{code}(\mathcal{I}^{(d)})$ where $\mathcal{I}^{(j)} = \{I_1^{(j)}, I_2^{(j)}, \dots, I_n^{(j)}\}$. Therefore, we can conclude that $\mathcal{C}$ is the intersection product of d -many interval codes. The converse follows similarly from Theorem 3, as one may form a realization of $\mathcal{C}$ by taking Cartesian products of intervals that realize the interval codes whose intersection product is $\mathcal{C}$. $\square$

Note that a decomposition of a code $\mathcal{C}$ as an intersection product of interval codes need not be unique. An important feature of Corollary 5 is that it yields a brute-force algorithm for recognizing whether or not a code is box-convex in $\mathbb{R}^d$. First, we enumerate all interval codes on $[n]$, which is possible because each such code is determined by the relative order of the $2n$ endpoints of the intervals in question (we explain this fact formally in Section 5). We then take all possible d -fold intersection products of these interval codes, and check whether or not $\mathcal{C}$ arises in this process.

Although the algorithm described above is far from efficient, this situation already stands in contrast to open and closed convex codes—Bukh and Jeffs [1] recently proved that open and closed convex codes realizable in $\mathbb{R}^2$ can be recognized algorithmically, but no such decidability result is

known in dimension three or higher. It remains an open question whether or not there is an *efficient* algorithm to recognize box-convex codes in $\mathbb{R}^d$.

The remainder of the paper is devoted to proving Theorem 3 in Section 2, and then exploring a variety of applications. In Section 3 we classify all codes $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ which are convex, but not box-convex—it turns out that there are only three such codes up to symmetry. We use one of these codes to show that a “monotonicity of open convexity” result proved by Cruz, Giusti, Itskov, and Kronholm [2] does not hold for box-convexity. In Section 4 we describe a family of codes whose box embedding dimension ranges over all integer values greater than or equal to two, while the open and closed embedding dimensions remain fixed and equal to two. We prove in Section 5 that codes $\mathcal{C} \subseteq 2^{[n]}$ realizable by boxes in $\mathbb{R}^d$ always have realizations in which the corners of the boxes have integer coordinates between 1 and $2n$, and additionally the interiors of the boxes realize the same code. In particular, there is no difference between open and closed box codes. Finally, Section 6 applies Theorem 3 beyond box codes by showing that a weak analogue of the aforementioned “monotonicity of open convexity” result holds for closed convex codes. The analogue we prove is necessarily weak, as Gambacini, Jeffs, Macdonald, and Shiu [6] proved that the strong analogue does not hold.

2 Proof and examples of the product theorem (Theorem 3)

We start by proving Theorem 3, and then give two examples. Let $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$ be as in the statement of Theorem 3. Observe that the nonempty atoms $\mathcal{U}^\sigma$ arising in the realization $\mathcal{U}$ cover all of $\mathbb{R}^{d_1}$, and likewise the nonempty atoms of the form $\mathcal{V}^\tau$ cover $\mathbb{R}^{d_2}$. Therefore, all the Cartesian products $\mathcal{U}^\sigma \times \mathcal{V}^\tau$ as σ and τ range over $\text{code}(\mathcal{U})$ and $\text{code}(\mathcal{V})$ respectively cover the entirety of $\mathbb{R}^{d_1+d_2}$. Fix codewords $\sigma \in \text{code}(\mathcal{U})$ and $\tau \in \text{code}(\mathcal{V})$. Each point in $\mathcal{U}^\sigma \times \mathcal{V}^\tau$ can be written as a pair (p, q) where $p \in \mathcal{U}^\sigma$ and $q \in \mathcal{V}^\tau$. We claim that $\text{patt}_{\mathcal{W}}(p, q) = \sigma \cap \tau$.

The inclusion $\text{patt}_{\mathcal{W}}(p, q) \subseteq \sigma \cap \tau$. Let $k \in \text{patt}_{\mathcal{W}}(p, q)$. Then $(p, q) \in W_k = U_k \times V_k$, which implies that $p \in U_k$ and $q \in V_k$. This in turn implies that $k \in \text{patt}_{\mathcal{U}}(p) = \sigma$, and $k \in \text{patt}_{\mathcal{V}}(q) = \tau$. Thus $k \in \sigma \cap \tau$.

The inclusion $\text{patt}_{\mathcal{W}}(p, q) \supseteq \sigma \cap \tau$. Let $k \in \sigma \cap \tau$. Then $p \in U_k$ and $q \in V_k$. Thus $(p, q) \in U_k \times V_k = W_k$, and so $k \in \text{patt}_{\mathcal{W}}(p, q)$, as desired.

Summary. Since $\text{patt}_{\mathcal{W}}(p, q) = \sigma \cap \tau$ for all $(p, q) \in \mathcal{U}^\sigma \times \mathcal{V}^\tau$, we see that the codewords of $\text{code}(\mathcal{W})$ are exactly those of the form $\sigma \cap \tau$, as σ ranges over $\text{code}(\mathcal{U})$ and τ ranges over $\text{code}(\mathcal{V})$. This is equivalent to the statement that $\text{code}(\mathcal{W}) = \text{code}(\mathcal{U}) \boxtimes \text{code}(\mathcal{V})$, proving the result. $\square$

Example 6. The left side of Figure 2 shows five boxes $\mathcal{U} = \{U_1, U_2, U_3, U_4, U_5\}$ in the plane, and their corresponding decompositions into intervals $\mathcal{I} = \{I_1, I_2, I_3, I_4, I_5\}$ and $\mathcal{J} = \{J_1, J_2, J_3, J_4, J_5\}$ in the x -direction and the y -direction respectively. We have $\text{code}(\mathcal{U}) = \{\mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, 1, \emptyset\}$, which is the intersection product of $\text{code}(\mathcal{I}) = \{\mathbf{124}, \mathbf{135}, 1, \emptyset\}$, and $\text{code}(\mathcal{J}) = \{\mathbf{123}, \mathbf{145}, 1, \emptyset\}$, as guaranteed by Theorem 3. The right side of the figure shows the atoms of nonempty codewords in the x -direction and y -direction, and the Cartesian products of these atoms, which we used in the proof of Theorem 3. For example, the top left square $\mathcal{I}^{124} \times \mathcal{J}^{123}$ is contained in the atom $\mathcal{U}^{12}$ since $12 = 124 \cap 123$.

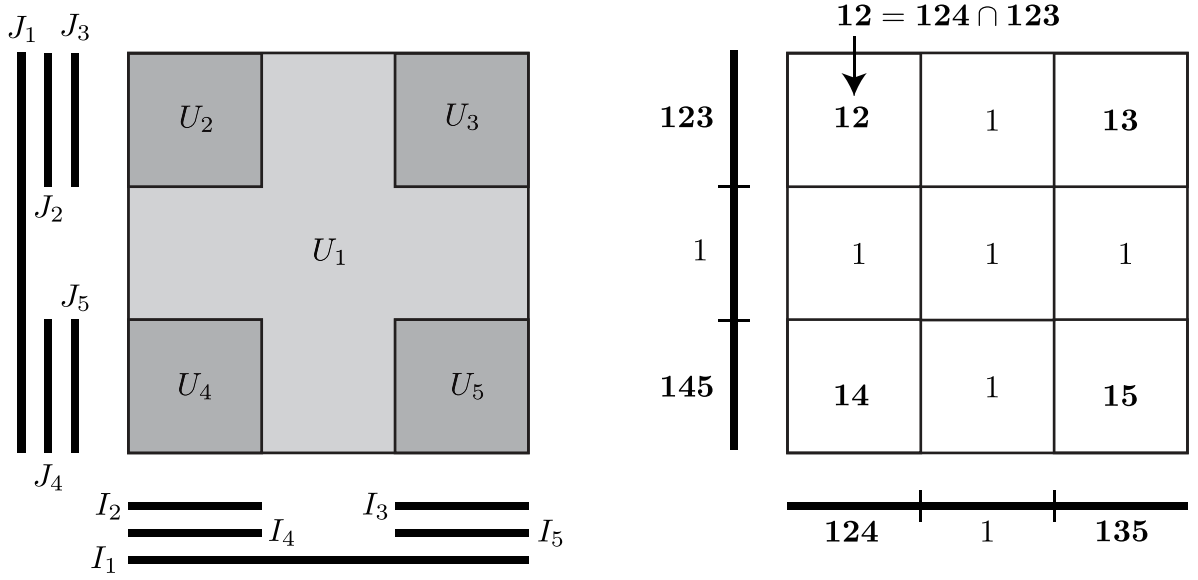


Figure 2: (Left) A realization of a box code in the plane, and its decomposition into interval codes in each coordinate direction. (Right) The atoms in the interval codes, and their products in the plane.

Example 7. Recall the box-convex code $\mathcal{C} = \{\mathbf{123}, \mathbf{24}, 12, 13, 23, 1, 2, 3, 4, \emptyset\}$ from Figure 1. We can decompose each box in this realization as a product of two intervals. Letting $\mathcal{I} = \{I_1, I_2, I_3, I_4\}$ be the intervals in the x -direction and $\mathcal{J} = \{J_1, J_2, J_3, J_4\}$ be the intervals in the y -direction, we then have $\mathcal{C} = \text{code}(\mathcal{I}) \boxtimes \text{code}(\mathcal{J})$ by Theorem 3. It turns out that $\text{code}(\mathcal{I}) = \{\mathbf{123}, \mathbf{24}, 12, 23, 1, 2, 4, \emptyset\}$ and $\text{code}(\mathcal{J}) = \{\mathbf{1234}, 134, 12, 1, 2, \emptyset\}$ in this case, and Figure 3 shows this decomposition.

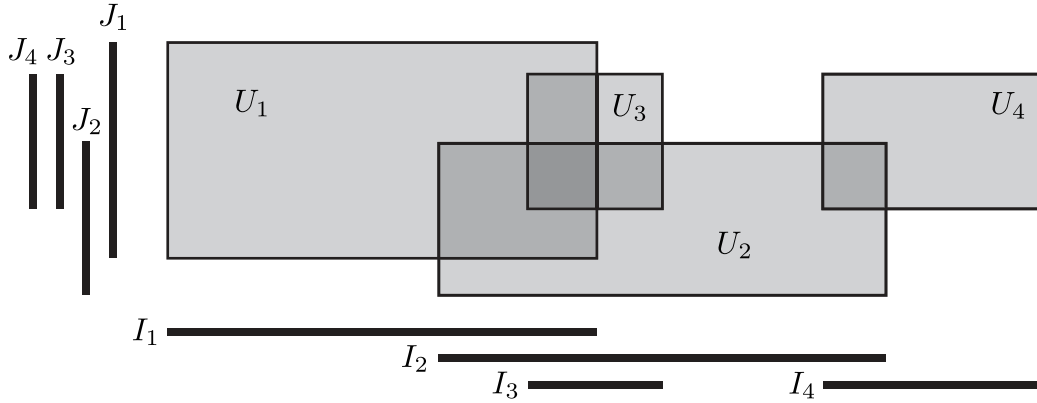


Figure 3: The boxes from Figure 1, decomposed into intervals in each coordinate direction.

3 Convex codes that are not box-convex

Every box-convex code is also convex, but the converse fails even for codes on three indices. Understanding the degree to which the converse fails helps us understand the difference between convex codes and box codes, and so we use Corollary 5 to characterize all codes on the index set $\{1, 2, 3\}$ that are convex, but not box-convex. There are three such codes, up to permuting the indices:

$$\begin{aligned}\mathcal{C}_1 &= \{\mathbf{123}, 12, 13, 23, \emptyset\}, \\ \mathcal{C}_2 &= \{\mathbf{123}, 12, 13, 23, 1, \emptyset\}, \text{ and} \\ \mathcal{C}_3 &= \{\mathbf{12}, \mathbf{13}, \mathbf{23}, 1, 2, 3, \emptyset\}.\end{aligned}$$

The codes $\mathcal{C}_1$ and $\mathcal{C}_2$ are closed and open convex in $\mathbb{R}^2$ since they have only one maximal codeword (see [4, Lemma 2.5]) and $\mathcal{C}_3$ is closed and open convex in $\mathbb{R}^2$ using three sets that pairwise intersect arranged in a triangular loop. Arguing that all other convex codes on $\{1, 2, 3\}$ are box-convex requires us to treat a large number of cases by hand, which we do in an appendix following the bibliography. In this section, we will establish that the codes $\mathcal{C}_1, \mathcal{C}_2$, and $\mathcal{C}_3$ are not box-convex, and also give an interesting example of such a code on four indices (Theorem 13).

The reason that $\mathcal{C}_1$ and $\mathcal{C}_3$ are not box-convex is that they have the following properties:

- (i) For $i = 1, 3$, the code $\mathcal{C}_i$ is not an interval code, and
- (ii) For $i = 1, 3$, if $\mathcal{C}_i = \mathcal{C} \boxtimes \mathcal{D}$ where $\mathcal{C}$ and $\mathcal{D}$ are convex codes, then $\mathcal{C} = \mathcal{C}_i$ or $\mathcal{D} = \mathcal{C}_i$.

Indeed, any code with these two properties is not box-convex—the first condition prevents it from being realized in $\mathbb{R}^1$, and the second condition means that it cannot be decomposed as an intersection product of interval codes since one of the factors would be the code itself. The code $\mathcal{C}_2$ requires a very slightly different argument, which we outline after our proof that $\mathcal{C}_1$ is not box-convex.

Intuitively, the second condition above alludes to a sense of irreducibility of box codes.

Definition 8. A code $\mathcal{C}$ is *irreducible* if, whenever we can write $\mathcal{C} = \mathcal{D} \boxtimes \mathcal{E}$ for convex codes $\mathcal{D}$ and $\mathcal{E}$, we must have $\mathcal{C} = \mathcal{D}$ or $\mathcal{C} = \mathcal{E}$.

Then, the above two conditions could be generalized into the following lemma:

Lemma 9. *If a code $\mathcal{C}$ is irreducible and is not an interval code, then $\mathcal{C}$ is not box-convex.*

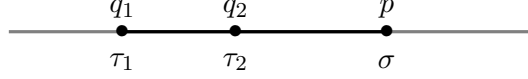
Proof. Assume for the sake of contradiction that $\mathcal{C}$ is box-convex in $\mathbb{R}^d$. Then, by Corollary 5, $\mathcal{C}$ is the intersection product of d -many interval codes. Because $\mathcal{C}$ is irreducible, one of these interval codes must be $\mathcal{C}$ itself. However, this would imply that $\mathcal{C}$ is realizable in $\mathbb{R}^1$, violating our assumption that $\mathcal{C}$ is not an interval code. $\square$

Before moving on to our analysis of $\mathcal{C}_1, \mathcal{C}_2$, and $\mathcal{C}_3$, we establish a simple criteria that obstructs a code from being realized by intervals.

Lemma 10. *Let $\mathcal{C}$ be a code with codewords $\sigma, \tau_1, \tau_2, \tau_3$ such that $\tau_i \subset \sigma$ and $\tau_i \not\subset \tau_j$ for all $i \neq j$, then $\mathcal{C}$ is not an interval code.*

Proof. Suppose for contradiction $\mathcal{C}$ is an interval code, with realization $\mathcal{I}$ in $\mathbb{R}^1$. Let p be a point in the atom $\mathcal{I}^\sigma$. Let q_i be a point in $\mathcal{I}^{\tau_i}$ for $1 \leq i \leq 3$. By pigeonhole principle, two of the q_i 's are on the same side of p . Without loss of generality, let $q_1 < q_2 < p$, so q_1, q_2 are both on the left side of p .

Since $\tau_1 \not\subset \tau_2$, there must be an index $j \in \tau_1$ such that $q_2 \notin U_j$. However, q_1 and p are in I_j , and since I_j is an interval, that means $q_2 \in I_j$ as well, a contradiction. The figure below gives an illustration of this situation. $\square$



Theorem 11. *The code $\mathcal{C}_1 = \{\mathbf{123}, 12, 13, 23, \emptyset\}$ is not box-convex.*

Proof. By Lemma 9, it suffices to establish that $\mathcal{C}_1$ is not an interval code, and that $\mathcal{C}_1$ is irreducible.

$\mathcal{C}_1$ is not an interval code. This can be shown by letting $\sigma = 123, \tau_1 = 12, \tau_2 = 13, \tau_3 = 23$. Note that all $\tau_i \subset \sigma$ and $\tau_i \not\subset \tau_j$ for all $i \neq j$. By Lemma 10, $\mathcal{C}_1$ is not an interval code.

$\mathcal{C}_1$ is irreducible. Suppose that $\mathcal{C}_1 = \mathcal{C} \boxtimes \mathcal{D}$ where $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ and $\mathcal{D} \subseteq 2^{\{1,2,3\}}$ are convex codes. The maximal codeword 123 in $\mathcal{C}_1$ is the intersection of codewords from $\mathcal{C}$ and $\mathcal{D}$ respectively, and the only way this can happen is if $123 \in \mathcal{C}$ and $123 \in \mathcal{D}$. Since every intersection of codewords between $\mathcal{C}$ and $\mathcal{D}$ lies in $\mathcal{C}_1$, we conclude that $\mathcal{C} \subseteq \mathcal{C}_1$ and $\mathcal{D} \subseteq \mathcal{C}_1$. Moreover, each of the codewords $\{12, 13, 23\}$ must appear in $\mathcal{C}$ or $\mathcal{D}$. By pigeonhole principle, two of these codewords appear in one of $\mathcal{C}$ or $\mathcal{D}$, and the third must appear in the same code lest we obtain a singleton codeword in $\mathcal{C} \boxtimes \mathcal{D}$. Hence, without loss of generality we have $\mathcal{C} = \{\mathbf{123}, 12, 13, 23, \emptyset\}$, which shows $\mathcal{C}_1$ that is irreducible. $\square$

A very similar proof shows that $\mathcal{C}_2$ is not box-convex. Again, Lemma 10 shows that $\mathcal{C}_2$ is not an interval code. The code $\mathcal{C}_2$ is not quite irreducible, but one can use a pigeonhole argument as above to show that if $\mathcal{C}_2 = \mathcal{C} \boxtimes \mathcal{D}$ then one of $\mathcal{C}$ and $\mathcal{D}$ is equal to $\mathcal{C}_2$ or $\mathcal{C}_1$. In any case, $\mathcal{C}_2$ cannot be decomposed as an intersection product of interval codes. We now turn to $\mathcal{C}_3$.

Theorem 12. *The code $\mathcal{C}_3 = \{\mathbf{12}, \mathbf{13}, \mathbf{23}, 1, 2, 3, \emptyset\}$ is not box-convex.*

Proof. **$\mathcal{C}_3$ is not an interval code.** Suppose for contradiction that $\mathcal{C}_3$ has a realization $\mathcal{I} = \{I_1, I_2, I_3\}$ by intervals in $\mathbb{R}^1$. The atoms of the maximal codewords $12, 13$, and 23 in this realization are each intervals. By permuting the indices, we can assume that they appear in the order $12, 13, 23$. But then I_2 contains points to the left and to the right of $\mathcal{I}^{13}$, and so I_2 covers $\mathcal{I}^{13}$. Since the codeword 13 does not contain 2 , this is a contradiction.

$\mathcal{C}_3$ is irreducible. Suppose that $\mathcal{C}_3 = \mathcal{C} \boxtimes \mathcal{D}$ where $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ and $\mathcal{D} \subseteq 2^{\{1,2,3\}}$ are convex codes. Since 123 is not a codeword of $\mathcal{C}_3$, at least one of $\mathcal{C}$ or $\mathcal{D}$ does not have 123 as a codeword. Say without loss of generality that $\mathcal{C}$ does not contain 123 . Since $12, 13$, and 23 are codewords in $\mathcal{C}_3$, and each is the intersection of some codeword in $\mathcal{C}$ with a codeword in $\mathcal{D}$, each of $12, 13$, and 23 must appear as a codeword in $\mathcal{C}$. But since $\mathcal{C}$ is convex and does not contain 123 , this means that $\mathcal{C}$ must contain the singleton codeword 1 —indeed, any line segment from the atom of 12 to the atom of 13 in a convex realization of $\mathcal{C}$ must pass through the atom of 1 . Similar reasoning shows that the singletons 2 and 3 are codewords of $\mathcal{C}$, so $\mathcal{C} = \mathcal{C}_3$. $\square$

We conclude with one last example of a convex code that is not box-convex, this time on four indices. Its realizations consist of a central region with three “flaps” attached, and it is not box-convex because one cannot add three flaps to a box to obtain a larger box. Such “flaps” will also play a role in Section 4.

Theorem 13. *The code $\mathcal{C}_4 = \{\mathbf{1234}, 12, 13, 14, \emptyset\}$ is open and closed convex in $\mathbb{R}^2$, but not box-convex in any dimension.*

Proof. The code $\mathcal{C}_4$ is open and closed convex in $\mathbb{R}^2$ since it has a unique maximal codeword (see [4, Lemma 2.5]). To see that $\mathcal{C}_4$ is not box-convex, we will use Lemma 9.

$\mathcal{C}_4$ is not an interval code. Let $\sigma = 1234, \tau_1 = 12, \tau_2 = 13, \tau_3 = 14$. Note that $\tau_i \subset \sigma$ and $\tau_i \not\subset \tau_j$ for all $i \neq j$. By Lemma 10, $\mathcal{C}_4$ is not an interval code.

$\mathcal{C}_4$ is irreducible. Suppose that $\mathcal{C}_4 = \mathcal{C} \boxtimes \mathcal{D}$ where $\mathcal{C} \subseteq 2^{[4]}$ and $\mathcal{D} \subseteq 2^{[4]}$ are convex codes. Since 1234 is the intersection of codewords from $\mathcal{C}$ and $\mathcal{D}$, we must have $1234 \in \mathcal{C}$ and $1234 \in \mathcal{D}$. This implies that $\mathcal{C} \subseteq \mathcal{C}_4$ and $\mathcal{D} \subseteq \mathcal{C}_4$.

Since every pairwise intersection of elements in $\{12, 13, 14\}$ would produce the singleton codeword 1, the only possibility is (up to swapping) $\mathcal{C} = \{\mathbf{1234}, 12, 13, 14, \emptyset\}$ and $\mathcal{D} = \{\mathbf{1234}, \emptyset\}$. Thus $\mathcal{C}_4$ is irreducible. $\square$

3.1 Non-monotonicity of box-convexity

Cruz, Giusti, Itskov and Kronholm [2] showed that open convexity is a monotone property of codes, in the sense that it is preserved by adding new non-maximal codewords. Below, we use the notation $\Delta(\mathcal{C}) \stackrel{\text{def}}{=} \{\sigma \subseteq [n] \mid \sigma \subseteq c \text{ for some } c \in \mathcal{C}\}$ to denote the abstract simplicial complex generated by $\mathcal{C}$.

Theorem 14 (Monotonicity of open convexity, [2]). *Let $\mathcal{C} \subseteq 2^{[n]}$ be an open convex code, and let $\mathcal{D} \subseteq 2^{[n]}$ be a code with $\mathcal{C} \subseteq \mathcal{D} \subseteq \Delta(\mathcal{C})$. Then $\mathcal{D}$ is open convex and $\text{odim}(\mathcal{D}) \leq \text{odim}(\mathcal{C}) + 1$.*

This result fails for closed convex codes, as shown by Gambacini, Jeffs, Macdonald, and Shiu [6]. It is thus natural to ask whether an analogous result holds for box-convex codes. It turns out that it does not. The method of proof used by Cruz, Giusti, Itskov, and Kronholm relies heavily upon the curvature of an open ball cutting through cylinders over a realization of $\mathcal{C}$. Such a technique cannot be applied for boxes since they exhibit no curvature on the boundary.

The codes we constructed earlier in this section provide an explicit counterexample showing that box-convexity is not monotone. Consider the code $\mathcal{C} = \{\mathbf{1234}, 12, 13, \emptyset\}$. This is certainly box-convex, and in fact this code can be realized by intervals in $\mathbb{R}^1$. However, the code $\mathcal{C}_4$ of Theorem 13 has the property that $\mathcal{C} \subseteq \mathcal{C}_4 \subseteq \Delta(\mathcal{C})$, but $\mathcal{C}_4$ is not box-convex.

4 Open embedding dimension versus box embedding dimension

In this section, we explore differences in embedding dimensions between ordinary convex codes and box-convex codes. The open embedding dimension and box embedding dimension of a code may be vastly different. Identifying specific characteristics of these differences allows us to better understand restrictions on box-convexity.

Our examples in this section are based on codes whose realizations are sunflowers. We say that $\mathcal{U} = \{U_1, \dots, U_n\}$ is a *sunflower* if $\text{code}(\mathcal{U}) = \{[n], 1, 2, \dots, n, \emptyset\}$. The sets U_i are called *petals* and $\bigcap_{i \in [n]} U_i$ is called the *center* of the sunflower. Sunflowers have proved to be highly useful in the study of open embedding dimensions of convex codes [8, 9], and they will additionally help us prove that box embedding dimension can be arbitrarily larger than the closed or open embedding dimensions of a code. Specifically, we will prove the following theorem.

Theorem 15. *For even $n \geq 4$, the code $\mathcal{F}_n \stackrel{\text{def}}{=} \{[n], 1, 2, \dots, n, \emptyset\}$ has $\text{bdim}(\mathcal{F}_n) = n/2$, while on the other hand $\text{odim}(\mathcal{F}_n) = \text{cdim}(\mathcal{F}_n) = 2$.*

Proof. Note that $\text{odim}(\mathcal{F}_n)$ and $\text{cdim}(\mathcal{F}_n)$ are both at most two, since $\mathcal{F}_n$ has a unique maximal codeword (see [4, Lemma 2.5]). To prove that $\text{bdim}(\mathcal{F}_n) = n/2$, we work by induction, starting from the slightly smaller base case $n = 2$.

Let $n = 2k$, where $k \geq 1$. We will prove that $\text{bdim}(\mathcal{F}_n) = n/2$ by induction on k . In the base case $k = 1$, observe that $\mathcal{F}_2 = \{\mathbf{12}, 1, 2, \emptyset\}$, which can be realized by intervals in $\mathbb{R}^1$. For the inductive step, it suffices to prove that $\mathcal{F}_n$ is the intersection product of k -many interval codes, and no fewer. We have $\mathcal{F}_n = \mathcal{C}_1 \boxtimes \dots \boxtimes \mathcal{C}_k$, where $\mathcal{C}_i = \{[n], 2i - 1, 2i, \emptyset\}$, which establishes that $\text{bdim}(\mathcal{F}_n) \leq k$.

To establish the matching lower bound, assume for contradiction that $\mathcal{F}_n = \mathcal{C}_1 \boxtimes \dots \boxtimes \mathcal{C}_\ell$ where $\ell < k$ and each $\mathcal{C}_i$ is an interval code. Since $[n]$ is a codeword of $\mathcal{F}_n$, it must be a codeword of each $\mathcal{C}_i$. Consequently, we have $\mathcal{C}_i \subseteq \mathcal{F}_n$ for each $1 \leq i \leq \ell$, and moreover each singleton codeword must appear in some $\mathcal{C}_i$. However, each $\mathcal{C}_i$ can contain at most two singleton codewords, since it is an interval code consisting of the codeword $[n]$ together with singletons and the empty set. There are $2k > 2\ell$ total singletons that must appear among the $\mathcal{C}_i$, a contradiction by pigeonhole principle. We conclude that $\mathcal{F}_n$ is the intersection product of k -many interval codes and no fewer, proving the result. $\square$

5 Integer coordinates and open boxes

In this section we discuss a strong application of the product theorem to box codes. As mentioned in Section 1, the product theorem applies naturally to boxes since each box in $\mathbb{R}^d$ is the Cartesian product of d -many intervals. In particular, Theorem 3 allows us to start with a collection of n boxes in $\mathbb{R}^d$, decompose it into d -many collections of n intervals in $\mathbb{R}^1$, perform some transformations on each collection of intervals in $\mathbb{R}^1$ that do not change their codes, and then return to a collection of boxes in $\mathbb{R}^d$ by taking products of our intervals. In this way, we can obtain nicer realizations in $\mathbb{R}^d$ by performing simple transformations in $\mathbb{R}^1$.

Our main result in this section is Theorem 16, which guarantees realizations that are “nicer” in the sense that all the boxes have corners with bounded integer coordinates, and interiors that realize the same code. Moreover, this theorem guarantees that the family of codes realizable by open boxes (which are products of nonempty open intervals) is identical to the family of box-convex codes, i.e. there is no difference between codes realizable with open boxes and those realizable with closed boxes. As mentioned in Section 1, this contrasts the situation for convex codes, where open convex codes differ greatly from closed convex codes.

Theorem 16. *Given a realization $\mathcal{U} = \{U_1, \dots, U_n\}$ of a code $\mathcal{C} \subseteq 2^{[n]}$ by boxes in $\mathbb{R}^d$, there exists a realization $\mathcal{V} = \{V_1, \dots, V_n\}$ of $\mathcal{C}$ by boxes, with the following properties:*

- (i) *each V_i has corners with integer coordinates between 1 and $2n$; that is, each V_i is the Cartesian product of closed intervals with integer endpoints between 1 and $2n$, and*
- (ii) *the interiors of the V_i also comprise a realization of $\mathcal{C}$.*

Moreover, if $\mathcal{C}$ has a realization by interiors of boxes in $\mathbb{R}^d$, then it has a realization by boxes in $\mathbb{R}^d$.

Proof. First, we work with just one coordinate. We start with a collection $\mathcal{I} = \{I_1, \dots, I_n\}$ of closed intervals in $\mathbb{R}^1$. We may express each interval according to its endpoints, as $I_1 = [a_1, a_2]$, $I_2 = [a_3, a_4]$, and so on. Observe that the relative order of these endpoints completely determines $\text{code}(\mathcal{I})$. Specifically, for every $p \in \mathbb{R}$ we have $i \notin \text{patt}_{\mathcal{I}}(p)$ if and only if $a_{2i} < p$ or $p < a_{2i-1}$. The relative order of the endpoints also determines the code of the interiors of the intervals in $\mathcal{I}$, according to the fact that a point p does not lie in the interior of I_i if and only if $a_{2i} \leq p$ or $p \leq a_{2i-1}$.

To prove item (i) in the theorem, consider the function $f : \mathbb{R} \rightarrow \mathbb{N}$ which sends a_i to k , where a_i is the k -th smallest number out of the set of endpoints $\{a_1, a_2, \dots, a_{2n}\}$. Note that if two endpoints a_i and a_j are equal, then $f(a_i) = f(a_j)$. So, f sorts the endpoints of the intervals from least to greatest, and equal endpoints are sent to the same number. Now, consider the collection of intervals $\mathcal{J} \stackrel{\text{def}}{=} \{J_1, \dots, J_n\}$ where

$$J_i \stackrel{\text{def}}{=} [f(a_{2i-1}), f(a_{2i})] \text{ for } i \in [n].$$

The collection $\mathcal{J}$ realizes the same code as $\mathcal{I}$ since the function f preserves the relative order of the endpoints. Moreover, $\mathcal{J}$ consists of intervals with integer endpoints between 1 and $2n$, since $f(a_i)$ is always an integer between 1 and $2n$. This proves condition (i) for $d = 1$.

To prove (ii), consider the collection $\mathcal{J}' = \{J'_1, \dots, J'_n\}$ where $J'_i \stackrel{\text{def}}{=} [f(a_{2i-1}) - 1/4, f(a_{2i}) + 1/4]$. That is, we extend each interval in $\mathcal{J}$ by $1/4$ to the left and to the right. This operation preserves the relative order of the endpoints of the intervals, except when a left endpoint coincides with a right endpoint, in which case these endpoints become separated with the new left endpoint strictly smaller than the new right endpoint. The latter change merely widens the atom of the codeword that appeared where this left and right endpoint coincided, and so does not remove or introduce any codewords in the realization. In particular, $\text{code}(\mathcal{J}') = \text{code}(\mathcal{J})$. See Figure 4 for an example of this.

We claim moreover that the interiors of the intervals in $\mathcal{J}'$ realize the same code as $\mathcal{J}'$. Let $\mathcal{J}''$ denote the collection of interiors of intervals in $\mathcal{J}'$. To see that $\text{code}(\mathcal{J}') = \text{code}(\mathcal{J}'')$, first note that $\text{patt}_{\mathcal{J}'}(p) \neq \text{patt}_{\mathcal{J}''}(p)$ only when p is a left or right endpoint of an interval in $\mathcal{J}'$. Moreover, there are no points in $\mathbb{R}^1$ that are both a left and right endpoint of some interval in $\mathcal{J}'$. This implies that every intersection pattern that occurs in $\mathcal{J}'$ occurs at a point that is not an endpoint of any interval—if a pattern occurs at a left endpoint, it also occurs slightly to the right of that endpoint, and symmetrically for right endpoints. Hence $\text{code}(\mathcal{J}') \subseteq \text{code}(\mathcal{J}'')$. A symmetric argument shows that the intersection patterns that occur in $\mathcal{J}''$ also occur away from endpoints of the intervals, implying that $\text{code}(\mathcal{J}'') \subseteq \text{code}(\mathcal{J}')$. Thus $\text{code}(\mathcal{J}') = \text{code}(\mathcal{J}'')$ as desired, proving condition (ii).

To obtain conditions (i) and (ii) simultaneously, we may simply apply the sorting argument used earlier to $\mathcal{J}'$, obtaining closed intervals with integer endpoints between 1 and $2n$ and the same relative

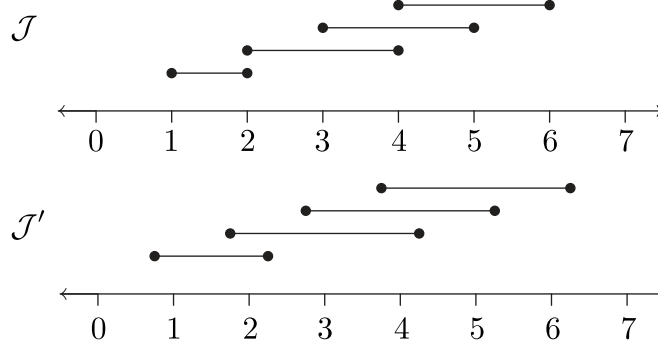


Figure 4: Extending the collection of intervals $\mathcal{J}$ to a collection $\mathcal{J}'$ whose closures and interiors both realize the code of $\mathcal{J}$.

order of endpoints as in $\mathcal{J}'$. Since the relative order of endpoints also determines the code of the interiors, this collection will have interiors that realize $\mathcal{C}$ as well.

A symmetric argument shows that if $\mathcal{C}$ has a realization by open intervals then it has a realization by closed intervals. Again, we can transform our starting realization so that the open intervals have integer endpoints. We then shrink each interval by $1/4$ from the left and right, separating left and right endpoints, and noting that this does not affect the code of the intervals. With left and right endpoints separated, the closures of these intervals will realize the same code as their interiors. This proves the entirety of the theorem for $d = 1$.

For $d > 1$, note that a collection of n -many boxes $\mathcal{U}$ in $\mathbb{R}^d$ can be decomposed into collections $\mathcal{I}_1, \dots, \mathcal{I}_d$ of intervals, where $\mathcal{I}_j$ consists of the n -many intervals obtained by projecting the boxes in $\mathcal{U}$ onto the j -th coordinate. Theorem 3 implies that $\text{code}(\mathcal{U}) = \text{code}(\mathcal{I}_1) \boxtimes \dots \boxtimes \text{code}(\mathcal{I}_d)$. Since we have established Theorem 16 for $d = 1$, we can replace each $\mathcal{I}_j$ by a realization $\mathcal{J}_j$ consisting of intervals with integer endpoints between 1 and $2n$ whose interiors realize the same code as $\mathcal{I}_j$. Then define $\mathcal{V} = \{V_1, \dots, V_n\}$ where V_i is the Cartesian product of the i -th interval from each $\mathcal{J}_j$. Again applying Theorem 3 we have $\text{code}(\mathcal{V}) = \text{code}(\mathcal{J}_1) \boxtimes \dots \boxtimes \text{code}(\mathcal{J}_d)$, and since $\text{code}(\mathcal{J}_j) = \text{code}(\mathcal{I}_j)$ for all $j \in [d]$ this proves (i). The fact that $\text{code}(\mathcal{J}_j)$ is the same as the code of the interiors of the $\mathcal{J}_j$ proves (ii). Finally, if we start with a collection of open boxes in $\mathbb{R}^d$, we may likewise decompose them into collections of open intervals, then apply Theorem 16 for $d = 1$ to obtain realizations of the same codes by closed intervals, and finally apply Theorem 3 to combine these collections of closed intervals to obtain a realization of the original code by closed boxes in $\mathbb{R}^d$. $\square$

6 Weak monotonicity of closed convexity

Our goal in this section is to prove Theorem 17, which provides a weak analogue of monotonicity of open convexity (Theorem 14) for closed convex codes. The result is “weak” in two senses. First, the new non-maximal codewords that we add to our code must comprise a downward-closed set, instead of being arbitrary. This is necessary, as Gambacini, Jeffs, Macdonald, and Shiu [6] gave an example of a closed convex codes with the property that adding a certain non-maximal codeword yields a code that is not closed convex in any dimension. Second, the closed embedding dimension may increase by

two in our theorem, whereas the monotonicity of open convexity theorem established by Cruz, Giusti, Itskov, and Kronholm [2] requires a dimension increase of at most one. It would be interesting to find examples of codes where our “+2” term is necessary, if they exist.

Below, a *downward-closed set* is a set of elements in a poset which contains every element of the poset that is less than some element of the set. Also, recall from Section 3.1 that $\Delta(\mathcal{C})$ is the abstract simplicial complex generated by $\mathcal{C}$.

Theorem 17. *Let $\mathcal{C} \subseteq 2^{[n]}$ be a closed convex code. Let $\mathcal{D}$ be a downward-closed subset of $\Delta(\mathcal{C}) \setminus \mathcal{C}$. Then $\mathcal{C} \cup \mathcal{D}$ is closed convex, and moreover*

$$\text{cdim}(\mathcal{C} \cup \mathcal{D}) \leq \text{cdim}(\mathcal{C}) + 2.$$

Proof. By Corollary 4, it suffices to show that $\mathcal{C} \cup \mathcal{D}$ is equal to $\mathcal{C} \bowtie \mathcal{E}$, where $\mathcal{E}$ is a code with $\text{cdim}(\mathcal{E}) \leq 2$. In fact, the choice $\mathcal{E} \stackrel{\text{def}}{=} \mathcal{D} \cup \{[n], \emptyset\}$ will suffice. Certainly we have $\text{cdim}(\mathcal{E}) \leq 2$ since $\mathcal{E}$ has a unique maximal codeword (see [4, Lemma 2.5]).

To prove that $\mathcal{C} \cup \mathcal{D} = \mathcal{C} \bowtie \mathcal{E}$, we show containment in each direction. We have $\mathcal{C} \subseteq \mathcal{C} \bowtie \mathcal{E}$, since $[n]$ is a codeword of $\mathcal{E}$. Moreover, if $\tau \in \mathcal{D}$ then there exists a maximal codeword σ of $\mathcal{C}$ with $\tau \subseteq \sigma$. We then have $\tau = \sigma \cap \tau$, which proves that $\tau \in \mathcal{C} \bowtie \mathcal{E}$.

For the reverse containment, let $\sigma \in \mathcal{C}$ and $\tau \in \mathcal{E}$. If $\tau = [n]$, then $\sigma \cap \tau = \sigma$, which lies in $\mathcal{C}$ and hence in $\mathcal{C} \cup \mathcal{D}$. If $\tau = \emptyset$, then $\sigma \cap \tau = \emptyset$, which again lies in $\mathcal{C}$. Otherwise, $\tau \in \mathcal{D}$. Note that $\mathcal{C} \cup \mathcal{D}$ contains $\Delta(\mathcal{D})$ since $\mathcal{D}$ is a downward-closed subset of $\Delta(\mathcal{C}) \setminus \mathcal{C}$. Thus $\mathcal{C} \cup \mathcal{D}$ contains $\sigma \cap \tau$, since $\sigma \cap \tau$ is a subset of τ . We conclude that $\mathcal{C} \cup \mathcal{D} = \mathcal{C} \bowtie \mathcal{E}$, which proves the result. $\square$

7 Conclusion and future questions

We have developed a useful “product theorem” (Theorem 3) for realizations of codes, and used it to explain several features of box-convex codes, including points of difference compared to ordinary convex codes. However, a number of fundamental questions remain open, and we describe several below that may be of interest for future research.

1. For each fixed $d \geq 2$, is there a polynomial-time algorithm to recognize whether or not a given code $\mathcal{C} \subseteq 2^{[n]}$ has a realization by boxes in $\mathbb{R}^d$? Every interval code $\mathcal{C} \subseteq 2^{[n]}$ has at most $2n + 1$ codewords (see the proof of [14, Proposition 1.5]), which together with Corollary 5 implies that every box-convex code in $\mathbb{R}^d$ has at most $(2n + 1)^d = O(n^d)$ codewords. Thus it would make sense to measure the complexity of such an algorithm as a function of n . Note that [14] has proved that there is an efficient algorithm for recognizing interval codes.
2. Among all box-convex codes $\mathcal{C} \subseteq 2^{[n]}$, which ones have $\text{bdim}(\mathcal{C})$ as large as possible? What is this largest value in terms of n ?
3. Can we classify “local obstructions” to box-convexity, specializing the notion of local obstructions for convex codes (see [4])? This would mean classifying which simplicial complexes are the nerve of a collection of boxes whose union is a box, analogous to the “convex union representable” complexes of Jeffs and Novik [11].

4. Can we find further applications of Theorem 3 to general open and closed convex codes?
5. Curry, Jeffs, Youngs, and Zhao [3] recently showed that every “inductively pierced” code has a realization consisting of open or closed balls. Does every inductively pierced code also admit a realization by boxes?
6. Work of Goldrup and Phillipson [7] classifies all open and closed convex codes on up to five indices. Our work has only enumerated box-convex codes on up to three indices. This motivates us to ask: which convex codes on up to four or five indices are not box-convex?

References

- [1] Boris Bukh and R. Amzi Jeffs. Planar convex codes are decidable. *SIAM Journal of Discrete Mathematics*, 2022. (To appear) arXiv:2207.06290.
- [2] Joshua Cruz, Chad Giusti, Vladimir Itskov, and Bill Kronholm. On open and closed convex codes. *Discrete & Computational Geometry*, 61:247–270, 2016.
- [3] Ryan Curry, R. Amzi Jeffs, Nora Youngs, and Ziyu Zhao. Recognizing and realizing inductively pierced codes. *Journal of Computational Geometry*, 2022. (To appear) arXiv:2207.06266.
- [4] 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.
- [5] 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.
- [6] Brianna Gambacini, R. Amzi Jeffs, Sam Macdonald, and Anne Shiu. Non-monotonicity of closed convexity in neural codes. *Vietnam Journal of Mathematics*, 50:359–373, 2022.
- [7] Sarah Ayman Goldrup and Kaitlyn Phillipson. Classification of open and closed convex codes on five neurons. *Advances in Applied Mathematics*, 112:101948, 2020.
- [8] R. Amzi Jeffs. Sunflowers of convex open sets. *Advances in Applied Mathematics*, 111:101935, 2019.
- [9] R. Amzi Jeffs. Embedding dimension phenomena in intersection complete codes. *Selecta Mathematica*, 28, 2021.
- [10] R. Amzi Jeffs. Open, closed, and non-degenerate embedding dimensions of neural codes. *Discrete & Computational Geometry*, 2022. (To appear) arXiv:2111.01010.
- [11] R. Amzi Jeffs and Isabella Novik. Convex union representability and convex codes. *International Mathematics Research Notices*, pages 7132–7158, 2019.
- [12] Alexander Kunin, Caitlin Lienkaemper, and Zvi Rosen. Oriented matroids and combinatorial neural codes. *Combinatorial Theory*, 3, 2023.

- [13] Caitlin Lienkaemper, Anne Shiu, and Zev Woodstock. Obstructions to convexity in neural codes. *Advances in Applied Mathematics*, 85:31–59, 2017.
- [14] Zvi Rosen and Yan X. Zhang. Convex neural codes in dimension 1. *arXiv e-prints: 1702.06907*, 2017.
- [15] Martin Tancer. Intersection patterns of convex sets via simplicial complexes: a survey. In *Thirty essays on geometric graph theory*, pages 521–540. Springer, New York, 2013.

Appendix

This appendix supplements Section 3, in which we showed that the codes $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$ are open and closed convex but not box-convex. Here we will argue that, up to permutation of indices, every other open or closed convex code $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ is box-convex. We must first enumerate the open and closed convex codes $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ up to symmetry, which we do in Figure 5, partitioning them according to their maximal codewords. The fact that this is the full list of open and closed convex codes on three indices follows from the fact that such a code must be max-intersection complete (i.e. contain all possible intersections of maximal codewords) and every max-intersection complete code $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ has an open and closed convex realization (see [4, Supplementary Text S1]). In fact, we will do slightly more than show that these codes are box-convex; we will determine their exact box embedding dimensions.

$\{1, \emptyset\}$	$\{1, 2, \emptyset\}$	$\{1, 2, 3, \emptyset\}$
$\{12, \emptyset\}$	$\{12, 1, \emptyset\}$	$\{12, 1, 2, \emptyset\}$
$\{12, 3, \emptyset\}$ $\{12, 3, 1, \emptyset\}$ $\{12, 3, 1, 2, \emptyset\}$	$\{12, 13, 1, \emptyset\}$ $\{12, 13, 1, 2, \emptyset\}$ $\{12, 13, 1, 2, 3, \emptyset\}$	$\star \{12, 13, 23, 1, 2, 3, \emptyset\}$
$\{123, \emptyset\}$ $\dagger \{123, 1, 2, 3, \emptyset\}$ $\{123, 12, 3, \emptyset\}$ $\dagger \{123, 12, 1, 2, 3, \emptyset\}$ $\{123, 12, 13, 2, \emptyset\}$ $\dagger \{123, 12, 13, 1, 2, 3, \emptyset\}$ $\dagger \{123, 12, 13, 23, 1, 2, \emptyset\}$	$\{123, 1, \emptyset\}$ $\{123, 12, \emptyset\}$ $\{123, 12, 1, 2, \emptyset\}$ $\{123, 12, 13, \emptyset\}$ $\{123, 12, 13, 1, 2, \emptyset\}$ $\star \{123, 12, 13, 23, \emptyset\}$ $\dagger \{123, 12, 13, 23, 1, 2, 3, \emptyset\}$	$\{123, 1, 2, \emptyset\}$ $\{123, 12, 1, \emptyset\}$ $\{123, 12, 1, 3, \emptyset\}$ $\{123, 12, 13, 1, \emptyset\}$ $\{123, 12, 13, 2, 3, \emptyset\}$ $\star \{123, 12, 13, 23, 1, \emptyset\}$

Figure 5: All open and closed convex codes on the index set $\{1, 2, 3\}$, up to symmetry. Those marked with $\star$ were shown not to be box-convex in Section 3. Those marked with $\dagger$ have box embedding dimension equal to two. The remainder can be realized by intervals.

We leave it as an exercise to verify that the unmarked codes in Figure 5 can be realized by closed intervals, i.e. have box embedding dimension equal to one. To prove that the $\dagger$ -marked codes have box embedding dimension two, we require the following proposition.

Proposition 18. Any code $\mathcal{C} \subseteq 2^{\{1,2,3\}}$ that contains the codewords 123, 1, 2, and 3 has $\text{bdim}(\mathcal{C}) \geq 2$, i.e. it is not realizable with closed intervals in $\mathbb{R}^1$.

Proof. This follows from Lemma 10, with $\sigma = 123$, $\tau_1 = 1$, $\tau_2 = 2$, and $\tau_3 = 3$. $\square$

This shows that all $\dagger$ -marked codes from Figure 5 have box embedding dimension at least two, except the code $\{\mathbf{123}, 12, 13, 23, 1, 2, \emptyset\}$. However, Lemma 10 likewise takes care of this case, with $\sigma = 123$, $\tau_1 = 12$, $\tau_2 = 13$, and $\tau_3 = 23$. Thus it remains to show that all $\dagger$ -marked codes have realizations by boxes in $\mathbb{R}^2$, which we do below in Figure 6.

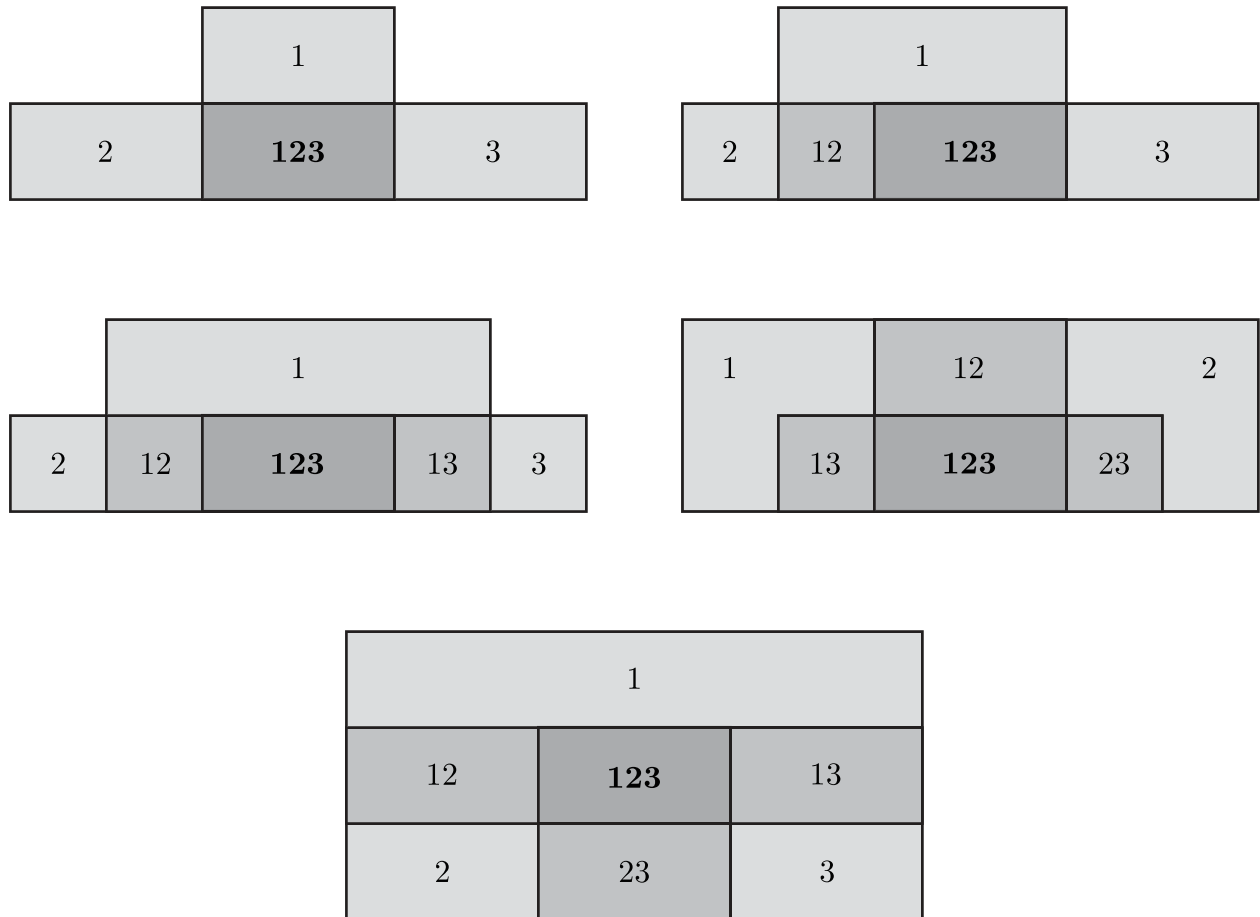


Figure 6: 2-dimensional box realizations of the $\dagger$ -marked codes from Figure 5.