THE DUPLICATE-FREE CONTRADICTION GRAPH DETECTS VC DIMENSION AT ORDERS TWO AND THREE, AND AT NO HIGHER ORDER

KOREA SUPERINTELLIGENCE LABS

Abstract. Campbell, Ibaibarriaga and Reyzin (arXiv:2605.20434) proved that the order-m contradiction graph $G_m(H)$ of a concept class H determines whether $\mathrm{VC}(H) \geq m$: the graph contains a *cube-trace clique* of size 2^m if and only if H shatters an m-point set. The vertices of $G_m(H)$ are labeled samples of length m in which domain points may repeat, and the reverse implication uses the repeated samples $((x, 0), \ldots, (x, 0))$ and $((x, 1), \ldots, (x, 1))$; the closing discussion of that paper notes that the proof "does not automatically apply to a duplicate-free variant in which each sample must use m distinct domain points". We determine exactly when the implication survives in that variant. At $m = 2$ it holds for every domain and every class, by a classification of the four supports and three duplicate-free witness vertices, machine-checked in Lean 4. At $m = 3$ it holds by a computer-assisted argument: two reductions make a counterexample a partition of $\{0, 1\}^n$ into eight subcubes of codimension three with $n \leq 7$, and an exhaustive census of the 16,139 such partitions that use every coordinate, with all realizer assignments, finds none. At every order $m \geq 4$ it fails: for each $m \geq 4$ we exhibit a class of exactly 2^m concepts on exactly $m + 1$ points with VC dimension $m - 1$ whose duplicate-free contradiction graph contains a cube-trace clique of size 2^m — an explicit sixteen-concept class on five points at $m = 4$, carried to every higher order by a padding argument, both machine-checked. The counterexample fails the original trace condition at precisely the repeated vertex the original proof uses, so the theorem for the repeated-sample graph is untouched. An exhaustive search over the five-point domain shows that the 640 counterexamples there are all instances of one construction, and the conjecture of the first version of this paper, that the members of a cube-trace clique share a support, is refuted already at $m = 2$.

1. Introduction

Let X be a set and $H \subseteq \{0, 1\}^X$ a class of binary concepts. Alon, Moran, Schefler and Yehudayoff [2] attached to H a sequence of graphs, the *contradiction graphs* $G_m(H)$, $m \geq 1$, in their graph-theoretic characterization of private learnability, and asked whether two classes, one of finite and one of infinite VC dimension, can have isomorphic contradiction graphs at every order. Campbell, Ibaibarriaga and Reyzin [1] answered this negatively and proved more: the single graph $G_m(H)$ already determines whether $\mathrm{VC}(H) \geq m$ (their Theorem 5.6), so the sequence $(G_m(H))_{m\geq 1}$ determines $\mathrm{VC}(H)$ (Corollary 5.7) and the pair of classes asked for does not exist (Corollary 5.8, answering their Problem 1.1). The mechanism is a graph-intrinsic certificate. A vertex of $G_m(H)$ is an H-realizable labeled sample of length m, two vertices are adjacent when they give some domain point opposite labels, and a *cube-trace clique* (their Definition 5.1) is a clique Q of size 2^m together with a bijection $\phi\colon Q \to \{0, 1\}^m$ under which the set of clique members not adjacent to T is a Boolean subcube for *every* vertex T of the graph. They prove that $G_m(H)$ contains a cube-trace clique of size 2^m if $\mathrm{VC}(H) \geq m$ (Proposition 5.2) and only if $\mathrm{VC}(H) \geq m$ (Theorem 5.4).

Samples are ordered sequences in both papers, and repetitions of domain points are allowed: "Repetitions of domain points are allowed; realizability implies that repeated occurrences of the same point receive the same label" ([1], Definition 3.2), and "the vertices of the contradiction graph $G_m(H)$ are H-realizable sequences of length m" ([2], the sentence following their definition of the contradiction graph). The proof of the reverse implication leans on this. For a domain point x that some clique member labels 0 and another labels 1, the repeated sequences

$$T_x^0 = ((x, 0), \ldots, (x, 0)), \qquad T_x^1 = ((x, 1), \ldots, (x, 1))$$

are vertices, their non-neighbour traces on the clique are two proper Boolean subcubes covering the cube, and such a pair consists of complementary facets (Lemma 5.3); hence every such x lies in every member of the clique, which forces a common m-point support. The closing section of [1] (Section 6, Discussion) says, verbatim:

> The reverse implication uses the vertices $((x,0),\dots,(x,0))$ and $((x,1),\dots,(x,1))$. Therefore the proof applies to the repeated-sample contradiction graph, but it does not automatically apply to a duplicate-free variant in which each sample must use m distinct domain points.

Write $G_m^{\neq}(H)$ for that variant: the subgraph of $G_m(H)$ induced on the vertices whose m domain points are distinct. Passing from $G_m(H)$ to $G_m^{\neq}(H)$ changes the certificate in both directions. The members of the clique must now be duplicate-free, a stronger demand on Q; but the trace condition is imposed only over the smaller vertex set, a weaker demand, and the vertices it no longer sees are exactly the ones the proof uses. The question left open is therefore:

Question 1.1. Let X be any set and $H \subseteq \{0,1\}^X$. If $G_m^{\neq}(H)$ contains a cube-trace clique of size 2^m, is $\mathrm{VC}(H) \geq m$?

The forward direction is free at every m: the clique that Proposition 5.2 of [1] builds from a shattered m-set uses m distinct points (Proposition 3.1). This paper answers Question 1.1 completely: the answer is *yes* at $m = 2$ and $m = 3$, and *no* at every $m \geq 4$.

Theorem ($m = 2$; Theorem 4.3)**.** *For every set X and every $H \subseteq \{0,1\}^X$, if $G_2^{\neq}(H)$ contains a cube-trace clique of size* 4*, then* $\mathrm{VC}(H) \geq 2$.

The proof does not simulate the repeated vertices. Four pairwise contradicting duplicate-free 2-samples have pairwise intersecting two-point supports, and the support pattern is one of two: a single common support, which is then shattered, or a three-point star $\{p,q\},\{p,q\},\{p,r\},\{p,r\}$ with the label at the hub p constant on each pair. In the star case either $\{p,r\}$ is shattered outright, or three duplicate-free vertices, (p,q) labeled by a realizer of one group and (q,r) labeled by the two realizers of the other, have two-element non-neighbour traces that all pass through one clique member; a point of the square $\{0,1\}^2$ has only two neighbours, so no bijection can make all three traces Boolean subcubes. Every step of this argument, for an arbitrary and possibly infinite domain, is formalised and machine-checked in Lean 4 (Section 14).

Theorem ($m = 3$, computer-assisted; Theorem† 8.1)**.** *For every set X and every $H \subseteq \{0,1\}^X$, if $G_3^{\neq}(H)$ contains a cube-trace clique of size* 8*, then* $\mathrm{VC}(H) \geq 3$.

Here the route is different. A counterexample can be shrunk to the 2^m realizers of the clique (Proposition 6.1) and to the union U of the clique's supports; over a finite domain a 2^m-clique of duplicate-free m-samples is a partition of $\{0,1\}^U$ into 2^m subcubes of codimension m, and a halfspace count gives $|U| \leq 7$ at $m = 3$. An exhaustive computation over the resulting finite space — every such partition on 3 to 7 points that uses every coordinate, with every choice of realizers — finds no counterexample. This part is a short C program and is not formalised; Section 8 states what was run and the controls run with it, and records the numbers of partitions of $\{0,1\}^n$ into 2^m subcubes of codimension m for $m = 2$, $n \leq 6$ and $m = 3$, $n \leq 7$ (Table 1).

Theorem ($m \geq 4$; Theorem 10.3)**.** *For every $m \geq 4$ there are a set X with $|X| = m+1$ and a class $H \subseteq \{0,1\}^X$ with $|H| = 2^m$ and* $\mathrm{VC}(H) = m - 1$ *such that $G_m^{\neq}(H)$ contains a cube-trace clique of size* 2^m.

The witness at $m = 4$ is explicit (Section 9): sixteen concepts on the five points y_1, y_2, y_3, a, b. The clique is indexed by $(v,\beta) \in \{0,1\}^3 \times \{0,1\}$; its member $S_{v,\beta}$ labels y_1, y_2, y_3 by v and labels *one* of the two bottom points — a if v has Hamming weight at most one, b otherwise — by β. The realizer of $S_{v,\beta}$ must also label the other bottom point, and it does so by β except at the two nodes 000 and 111, whose neighbours in $\{0,1\}^3$ all carry the same colour as they do; there it labels the other bottom point by 0. Those two *interior* nodes are what lowers the VC dimension: the realizers of $S_{000,0}$ and $S_{000,1}$ agree everywhere except at a, those of $S_{111,0}$ and $S_{111,1}$ agree except at b, and no four-point set is shattered;

and they are invisible to every duplicate-free 4-sample, which can drop only one of the five points and so never sees a node together with a differently coloured neighbour. Interior nodes of both colours need two nodes of $\{0,1\}^{m-1}$ with disjoint closed neighbourhoods, i.e. at Hamming distance at least 3, and those exist exactly when $m-1 \geq 3$: this is why $m = 3$ is the threshold. The five finite parts of the statement — injectivity, duplicate-freeness, realizability, pairwise contradiction, and $\mathrm{VC}(H_4) < 4$ — are decided by the Lean kernel over the definitions; the trace condition is reduced by hand to a finite lemma with 1,280 instances, also decided by the kernel. The passage to every $m \geq 4$ is a padding lift (Section 10): adjoin one point z, double the class by letting each concept take both values at z, and append (z, β) to each clique member. The trace condition for a sample that avoids z is an intersection of $m+1$ subcubes, and an intersection of subcubes can be empty; the side condition that makes it nonempty — every concept of the class realizes some clique member — is carried along the induction and holds for the five-point witness. The lift transports all three extremal parameters.

Two further things are settled. An exhaustive search over the five-point domain at $m = 4$ (Section 11) finds exactly 640 counterexamples among the 589,185 perfect matchings of the 5-cube with all realizer assignments, and every one of them is an instance of the construction just described with the colouring of $\{0,1\}^3$ varied. And the first version of this paper conjectured that the members of a cube-trace clique of $G_m^{\neq}(H)$ always share a support, which would have given Question 1.1 at every m through Proposition 4.4 of [1]; that conjecture is false at the first non-trivial order (Section 12): the *twin star* $\{000, 011, 100, 111\}$ on three points carries a cube-trace 4-clique in $G_2^{\neq}$ with two different supports (and $\mathrm{VC} = 2$, so it is no counterexample to the theorem).

The counterexamples contradict nothing in [1]. The very clique that refutes Question 1.1 at $m = 4$ fails the trace condition of $G_4(H_4)$ at the repeated vertex $((a,0),(a,0),(a,0),(a,0))$, whose trace has twelve members (Proposition 9.3); this is exactly the vertex Theorem 5.4 of [1] uses and the duplicate-free graph deletes.

We are not aware of any treatment of the duplicate-free variant: neither [1] nor [2] considers it beyond the sentence quoted above (the word "duplicate-free" occurs once in the whole of [1]), and the searches run for this paper found nothing. On 2026-09-07 OpenAlex listed no work citing either of its two records for arXiv:2605.20434 (W7161982800, W7162149937: `cited_by_count` 0, and a `cites:` filter returning no works for either) and none citing its record for arXiv:2304.03996 (W4364382962); Semantic Scholar listed no citer of [1] and five citers of [2] — [1] itself and four papers on private and multiclass learnability (arXiv:2310.18428, 2407.07765, 2506.12856, 2511.12659), none of which concerns contradiction-graph certificates for the VC dimension; a full-text search of a local arXiv text corpus (376 shards, to the 2609 announcement window) for "cube-trace" returned only [1] itself and an unrelated equation label in arXiv:2402.06401, and for "duplicate-free variant" and "duplicate-free sample" only uses in program analysis and knapsack enumeration; the arXiv API returned no work for "cube-trace" and three for "contradiction graph" ([1], [2], and one on hallucination detection); and the OEIS had none of the sequences of Table 1 beyond the identity of Remark 8.2. The same searches on 2026-09-03 gave the same answers, and added that a full-text search of the corpus for "contradiction graph" returns, besides [1], only papers whose "contradiction graph" is a different object (retrieval-augmented generation, support testing, belief modelling).

Changes since version 1. Version 1 of this paper (2026-09-03) proved the $m = 2$ and $m = 3$ cases, priced the census route at $m = 4$ as out of reach, and proposed the common-support conjecture as the way to all m. This version answers Question 1.1 negatively for every $m \geq 4$ (Sections 9 and 10), records the closed-form construction, the exhaustive search on five points and the order-3 degeneration of the construction (Section 11), refutes the conjecture of version 1 (Conjecture 9.1 there) by the twin star (Section 12), and replaces the open problems (Section 13). Sections 2–8 are unchanged apart from a few sentences pointing forward to the new material (the last sentence of Section 2.1, the end of Section 3, Remark 4.4, and the positive control of Section 8); every statement proved there stands, and the machine-checked statements of version 1 are pinned to the same formal declarations. The title and abstract now state the sharp threshold.

2. Preliminaries

Throughout, X is a set, $H \subseteq \{0,1\}^X$, and $m \geq 1$. We follow the definitions of [1], Section 3, and keep their numbering when we refer to them.

2.1. **Samples, realizability, the two graphs.** A *length-m labeled sequence* (or *m-sample*) is a tuple $S = ((x_1, y_1), \dots, (x_m, y_m)) \in (X \times \{0,1\})^m$; we write $(x, b) \in S$ if (x, b) is an entry of S and $\operatorname{supp}(S) = \{x_1, \dots, x_m\}$. It is *$H$-realizable* if some $h \in H$ has $h(x_i) = y_i$ for all i (Definition 3.2 of [1]); such an h is a *realizer* of S. It is *duplicate-free* if $x_i \neq x_j$ whenever $i \neq j$, i.e. $|\operatorname{supp}(S)| = m$. Two sequences S, T are *adjacent*, $S \sim T$, if there are $x \in X$ and $b \in \{0,1\}$ with $(x, b) \in S$ and $(x, 1-b) \in T$ (Definition 3.3). Adjacency is symmetric, a realizable sequence is never adjacent to itself, and two sequences realized by the same concept are never adjacent.

Definition 2.1. The *contradiction graph* $G_m(H)$ has the H-realizable m-samples as vertices and adjacency as its edge relation. The *duplicate-free contradiction graph* $G_m^{\neq}(H)$ is the subgraph of $G_m(H)$ induced on the duplicate-free vertices.

Adjacency and, below, the non-neighbour trace of a vertex depend only on the set of entries of a sequence. For duplicate-free sequences that set determines the sequence up to the order of its entries, so nothing in this paper changes if the vertices of $G_m^{\neq}(H)$ are taken to be m-element sets of labeled points instead of sequences.

2.2. **Traces, subcubes, cube-trace cliques.** For Q a set of vertices and T a vertex, the *non-neighbour trace* of T on Q is $M_Q(T) = \{S \in Q : S \not\sim T\}$ (Definition 3.4). A *Boolean subcube* of $\{0,1\}^m$ is a nonempty set of the form $\{\sigma \in \{0,1\}^m : \sigma_i = \tau_i \text{ for all } i \in I\}$ with $I \subseteq [m]$ and $\tau \in \{0,1\}^I$ (Definition 3.5); it has $2^{m-|I|}$ elements. The two-element subcubes of $\{0,1\}^2$ are its four edges, so a point of $\{0,1\}^2$ lies in exactly two of them. Shattering and $\mathrm{VC}(H)$ are as usual: $R \subseteq X$ is shattered if every $g \colon R \to \{0,1\}$ is the restriction of some $h \in H$, and $\mathrm{VC}(H) \geq d$ means that some d-element set is shattered.

Definition 2.2 (cube-trace clique; Definition 5.1 of [1], in either graph)**.** Let G be $G_m(H)$ or $G_m^{\neq}(H)$. A clique Q of G with $|Q| = 2^m$ is a *cube-trace clique* of G if there is a bijection $\phi \colon Q \to \{0,1\}^m$ such that $\phi(M_Q(T))$ is a Boolean subcube of $\{0,1\}^m$ for every vertex T of G.

Since subcubes are nonempty, a cube-trace clique has no vertex of G adjacent to all its members. It is convenient to carry the bijection as an indexing: a cube-trace clique of $G_m^{\neq}(H)$ is the same thing as a family $(S_\sigma)_{\sigma \in \{0,1\}^m}$ of pairwise distinct, duplicate-free, H-realizable m-samples, pairwise adjacent, such that for every duplicate-free H-realizable T the set $\{\sigma : S_\sigma \not\sim T\}$ is a Boolean subcube; ϕ is the inverse of $\sigma \mapsto S_\sigma$. For $G_m(H)$ the same holds with "duplicate-free" dropped from both places. This is how the formal statements of Appendix A are phrased (`CubeTraceRep` for G_m, `CubeTraceDF` for $G_m^{\neq}$).

3. The forward direction

Proposition 5.2 of [1] builds, from a shattered set $R = \{x_1, \dots, x_m\}$, the clique of the $S_\sigma = ((x_1, \sigma_1), \dots, (x_m, \sigma_m))$, $\sigma \in \{0,1\}^m$. Its members use m distinct points, so the same clique lives in $G_m^{\neq}(H)$.

Proposition 3.1 (Proposition 5.2 of [1], in both graphs)**.** *If $\mathrm{VC}(H) \geq m$, then $G_m(H)$ and $G_m^{\neq}(H)$ each contain a cube-trace clique of size 2^m.*

Proof. Let R be shattered, $|R| = m$, and list it as $x_1, \dots, x_m$. Each S_σ is realizable because R is shattered, duplicate-free because the x_i are distinct, and $S_\sigma \sim S_\tau$ for $\sigma \neq \tau$ at any x_i with $\sigma_i \neq \tau_i$. Let T be any realizable m-sample, with realizer h, and let $I = \{i : x_i \in \operatorname{supp}(T)\}$. Then $S_\sigma \not\sim T$ if and only if $\sigma_i = h(x_i)$ for all $i \in I$, so $\{\sigma : S_\sigma \not\sim T\}$ is the subcube with fixed set I and values $h(x_i)$. This holds for every realizable T, duplicate-free or not. $\square$

So only the reverse implication is at issue in Question 1.1. We record it as a statement about two-point supports at $m = 2$ (Section 4) and as a finite computation at $m = 3$ (Section 8), and refute it from $m = 4$ on (Sections 9 and 10).

4. Order two

4.1. **The finite obstruction.** The only fact about the geometry of $\{0,1\}^2$ that the proof needs is that its two-element subcubes are the four edges of a 4-cycle.

Lemma 4.1. *Let a, b, c, d be the four points of $\{0,1\}^2$. The three sets $\{c, d\}$, $\{a, c\}$, $\{b, c\}$ are not all Boolean subcubes of $\{0,1\}^2$.*

Proof. A two-element Boolean subcube is an edge of the square, and the vertex c has two edges. Formally this is a finite check: the statement quantifies over the four points and is decided by evaluation over $\{0,1\}^2$ (declaration `no_three_edges`, proved by `decide`). □

4.2. **The configuration theorem.** A cube-trace 4-clique of $G_2^{\neq}(H)$ provides, for each $\sigma \in \{0,1\}^2$, a two-point support and a realizer, subject to pairwise contradiction and to the subcube condition at the duplicate-free vertices that label two distinct points by one of the realizers.

Theorem 4.2. *Let $H \subseteq \{0,1\}^X$. Suppose that for each $\sigma \in \{0,1\}^2$ we are given a concept $h_\sigma \in H$ and a two-element set $R_\sigma \subseteq X$ such that*

(C1) *for all $\sigma \neq \tau$ there is $z \in R_\sigma \cap R_\tau$ with $h_\sigma(z) \neq h_\tau(z)$;*
(C2) *for all distinct $u, v \in X$ and every $\sigma_0 \in \{0,1\}^2$, the set*

$$\{\rho \in \{0,1\}^2 : h_\rho(z) = h_{\sigma_0}(z) \text{ for all } z \in R_\rho \cap \{u, v\}\}$$

is a Boolean subcube of $\{0,1\}^2$.

Then $\mathrm{VC}(H) \geq 2$.

Proof. Case 1: all R_σ coincide, say $R_\sigma = \{p, q\}$ for all σ. The map $\sigma \mapsto (h_\sigma(p), h_\sigma(q))$ is injective by (C1), hence a bijection onto $\{0,1\}^2$, so $\{p, q\}$ is shattered by $\{h_\sigma\} \subseteq H$. (This is Proposition 4.4 of [1] in the case at hand.)

Case 2: some $R_{s_0} \neq R_{t_0}$. By (C1) there is $p \in R_{s_0} \cap R_{t_0}$ with $h_{s_0}(p) \neq h_{t_0}(p)$; write $R_{s_0} = \{p, q\}$ and $R_{t_0} = \{p, r\}$. Then p, q, r are distinct, since the two supports are two-element sets sharing p and differing. Every $\rho \in \{0,1\}^2$ falls into exactly one of three classes:

(I) $p \in R_\rho$, $h_\rho(p) = h_{s_0}(p)$, and $R_\rho = \{p, q\}$;
(II) $p \in R_\rho$, $h_\rho(p) = h_{t_0}(p)$, and $R_\rho = \{p, r\}$;
(III) $p \notin R_\rho$ and $R_\rho = \{q, r\}$.

Indeed, if $p \in R_\rho$ and $h_\rho(p) = h_{s_0}(p)$, then either $\rho = s_0$ or (C1) gives a contradiction between ρ and s_0 at a point of $\{p, q\}$ other than p, so $q \in R_\rho$ and $R_\rho = \{p, q\}$; if $p \in R_\rho$ and $h_\rho(p) \neq h_{s_0}(p)$, then $h_\rho(p) = h_{t_0}(p)$ and the same argument against t_0 gives $R_\rho = \{p, r\}$; and if $p \notin R_\rho$, then $\rho \neq s_0, t_0$ and (C1) forces $q \in R_\rho$ and $r \in R_\rho$.

Class (III) is empty. Suppose w is in class (III). A member ρ' of class (I) contradicts w only at q, so $h_{\rho'}(q) \neq h_w(q)$: all class-(I) members carry the label $1 - h_w(q)$ at q. Two of them would then agree at p and at q and could not contradict, so class (I) has at most one member. Likewise class (II) has at most one member. Two members of class (III) both differ from h_{s_0} at q and from h_{t_0} at r, hence agree at both points and cannot contradict; so class (III) has at most one member. That accounts for at most three of the four ρ.

The star. Every ρ is therefore in class (I) or (II), and each class has at most two members (its members agree at p and must contradict at the other point, where only two labels exist). So each class has exactly two: s_0, s_2 on $\{p, q\}$ with $h_{s_0}(p) = h_{s_2}(p)$ and $h_{s_0}(q) \neq h_{s_2}(q)$, and t_0, t_2 on $\{p, r\}$ with $h_{t_0}(p) = h_{t_2}(p) \neq h_{s_0}(p)$ and $h_{t_0}(r) \neq h_{t_2}(r)$.

Case 2a: $h_{s_0}(r) \neq h_{s_2}(r)$. Then $\{p, r\}$ is shattered: with $\varepsilon = h_{s_0}(p)$, the concepts h_{s_0}, h_{s_2} realize $(\varepsilon, 0)$ and $(\varepsilon, 1)$ on (p, r) in some order, and h_{t_0}, h_{t_2} realize $(1-\varepsilon, 0)$ and $(1-\varepsilon, 1)$.

Case 2b: $h_{s_0}(r) = h_{s_2}(r) =: \beta$. Let $t_a \in \{t_0, t_2\}$ be the one with $h_{t_a}(r) = \beta$. Apply (C2) three times. With $(u, v, \sigma_0) = (p, q, t_0)$ the set is $\{t_0, t_2\}$: s_0, s_2 disagree with h_{t_0} at p, while t_2 agrees at p and $R_{t_2} \cap \{p, q\} = \{p\}$. With (q, r, s_0) the set is $\{s_0, t_a\}$: s_2 disagrees with h_{s_0} at q, and for t_0, t_2 only the point r is tested. With (q, r, s_2) the set is $\{s_2, t_a\}$ by the same computation. These are three distinct two-element subsets of $\{0,1\}^2$ through the common point t_a, and (C2) says all three are Boolean subcubes, contradicting Lemma 4.1. $\square$

4.3. **The theorem.**

Theorem 4.3. *Let X be any set and $H \subseteq \{0,1\}^X$. If $G_2^{\neq}(H)$ contains a cube-trace clique of size 4, then* $\mathrm{VC}(H) \geq 2$.

Proof. Let $(S_\sigma)_{\sigma\in\{0,1\}^2}$ be the clique, choose a realizer h_σ of each S_σ, and put $R_\sigma = \mathrm{supp}(S_\sigma)$, a two-element set since S_σ is duplicate-free. Condition (C1) is the clique property: $S_\sigma \sim S_\tau$ gives a common point with opposite labels, and the realizers reproduce those labels. For (C2), fix distinct u, v and σ_0, and let $T = ((u, h_{\sigma_0}(u)), (v, h_{\sigma_0}(v)))$. It is duplicate-free and realized by h_{σ_0}, hence a vertex of $G_2^{\neq}(H)$, and $S_\rho \not\sim T$ exactly when h_ρ agrees with h_{σ_0} on $R_\rho \cap \{u, v\}$. So the set in (C2) is $\{\rho : S_\rho \not\sim T\}$, a Boolean subcube by hypothesis. Theorem 4.2 applies. $\square$

Remark 4.4. The three vertices used in Case 2b are the duplicate-free replacement for T_x^0, T_x^1. They are not padded constant sequences, and the useful ones are those whose support avoids the hub p and meets both petals. Nothing in the argument uses finiteness of X or of H, and the trace condition is invoked at only three vertices. Case 2a is not vacuous: Section 12 exhibits a cube-trace 4-clique with the star pattern, in which $\{p, r\}$ is shattered and the supports are not all equal.

5. Controls at order two

Three finite checks calibrate Theorem 4.3: the hypothesis cannot be weakened to "a 4-clique of duplicate-free vertices", the vertex the source's proof uses is exactly what $G_2^{\neq}$ removes, and the hypothesis is satisfiable with the conclusion sharp.

Let $H^* = \{000, 010, 100, 101\} \subseteq \{0,1\}^{\{p,q,r\}}$, concepts written as their values at p, q, r, and let

$$s_A = ((p,0),(q,0)), \quad s_B = ((p,0),(q,1)), \quad s_C = ((p,1),(r,0)), \quad s_D = ((p,1),(r,1)).$$

This is the star pattern of Theorem 4.2 on the smallest possible domain. By the census of Section 8 it is, up to renaming the three points, the only full-support pattern at $m = 2$: there are six partitions of $\{0,1\}^3$ into four subcubes of codimension two that use every coordinate, and none on four or more points (Table 1).

Proposition 5.1. *The samples s_A, s_B, s_C, s_D are duplicate-free, H^*-realizable and pairwise adjacent, so they form a clique of size $4 = 2^2$ in $G_2^{\neq}(H^*)$; and $\mathrm{VC}(H^*) = 1$. Consequently $G_2^{\neq}(H^*)$ contains no cube-trace clique of size* 4.

Proof. The first two assertions are finite checks over the eight concepts on three points (declarations `star_clique`, `star_vc`, proved by `decide`): the realizers are $000, 010, 100, 101$ in order, and no two-point set is shattered. The last assertion is Theorem 4.3 applied to H^*: a cube-trace 4-clique would give $\mathrm{VC}(H^*) \geq 2$. $\square$

The phenomenon "a 2^m-clique without shattering" is Proposition 4.6 and Theorem 4.7 of [1] (their tree class P has $\mathrm{VC}(P) = 1$, and $G_m(P)$ contains a 2^m-clique for every m); what Proposition 5.1 adds is that the smallest instance is duplicate-free and lives in $G_2^{\neq}$, so Theorem 4.3 has content: it is the trace condition, not the clique, that it rules out here.

Proposition 5.2. *The repeated sample $T_q = ((q,0),(q,0))$ is H^*-realizable, hence a vertex of $G_2(H^*)$, and is not duplicate-free. Its non-neighbour trace on the clique is $\{s_A, s_C, s_D\}$: T_q is adjacent to s_B and to no other member. A three-element subset of $\{0,1\}^2$ is never a Boolean subcube.*

Proof. Finite checks (`repeated_witness`, `not_subcube_three`, both by `decide`): 000 realizes T_q; s_B contains $(q,1)$; s_A contains $(q,0)$ and s_C, s_D do not mention q; and a Boolean subcube of $\{0,1\}^2$ has 1, 2 or 4 elements. $\square$

So in $G_2(H^*)$ the vertex T_q already witnesses that the clique is not cube-trace, as in the proof of Theorem 5.4 of [1]; $G_2^{\neq}(H^*)$ deletes it, and Theorem 4.3 finds the obstruction elsewhere: here the witnesses of Case 2b are $((p,1),(q,0))$, $((q,0),(r,0))$, $((q,1),(r,0))$ with traces $\{s_C, s_D\}$, $\{s_A, s_C\}$, $\{s_B, s_C\}$, three edges at s_C (an illustration by hand, not a formal check).

Proposition 5.3. *Let $H_{\mathrm{full}} = \{0,1\}^{\{p,q\}}$ be the full class on two points. Then* $\mathrm{VC}(H_{\mathrm{full}}) = 2$ *— some two-point set is shattered and no three-point set is — and $G_2^{\neq}(H_{\mathrm{full}})$ contains a cube-trace clique of size* 4.

Proof. The VC statement is a finite check (`full_vc`, by `decide`); the clique is Proposition 3.1. So the hypothesis of Theorem 4.3 is satisfiable and its conclusion cannot be strengthened to $\mathrm{VC}(H) \geq 3$. □

6. The realizer reduction

Shrinking the class makes the cube-trace hypothesis easier and the VC conclusion harder, because the trace condition is a universal statement over the vertex set and VC is monotone. Hence a counterexample to Question 1.1 can always be shrunk to the realizers of its clique.

Proposition 6.1. *Let $(S_\sigma)_{\sigma \in \{0,1\}^m}$ be a cube-trace clique of $G_m^{\neq}(H)$ and suppose* $\mathrm{VC}(H) < m$. *Then there is a map $f \colon \{0,1\}^m \to \{0,1\}^X$ with $f(\{0,1\}^m) \subseteq H$ and* $\mathrm{VC}(f(\{0,1\}^m)) < m$ *such that the same family (S_σ) is a cube-trace clique of $G_m^{\neq}(f(\{0,1\}^m))$. In particular, if Question 1.1 has a negative answer at order m, it has one with a class of at most 2^m concepts.*

Proof. Let $f(\sigma)$ be a realizer of S_σ and $H' = f(\{0,1\}^m) \subseteq H$. Each S_σ is H'-realizable by construction, and the clique and duplicate-free conditions do not mention H. Every duplicate-free H'-realizable T is H-realizable, so the subcube condition it must satisfy in $G_m^{\neq}(H')$ is one it already satisfies in $G_m^{\neq}(H)$. Finally a set shattered by H' is shattered by H, so $\mathrm{VC}(H') \leq \mathrm{VC}(H) < m$. □

7. Cliques over finite domains are subcube partitions

Nothing in this section or the next is part of the formal development. The statements here are proved in ordinary mathematics, those of Section 8 rest on a computation, and both carry a dagger. The second reduction — restricting the domain to the union of the supports — is stated here; the order-2 theorem does not need it.

Let U be a finite set, identified with $[n]$, and identify $\{0,1\}^U$ with the cube $\{0,1\}^n$. A duplicate-free m-sample S over U determines the set

$$C(S) = \{h \in \{0,1\}^U : h(x) = y \text{ for all } (x,y) \in S\},$$

a subcube of $\{0,1\}^n$ of codimension exactly m (it fixes the m coordinates in $\mathrm{supp}(S)$), and S is H-realizable if and only if $C(S) \cap H \neq \emptyset$. Conversely a subcube of codimension m determines a duplicate-free m-sample up to the order of its entries.

Proposition† 7.1. *Two duplicate-free m-samples S, T over U are adjacent if and only if $C(S) \cap C(T) = \emptyset$. Consequently, a set of 2^m pairwise adjacent duplicate-free m-samples over U is, up to the order of entries within each sample, the same thing as a partition of $\{0,1\}^n$ into 2^m subcubes of codimension m.*

Proof. If $S \sim T$ at x, no h can carry both labels of x. If $S \not\sim T$, the two partial labelings agree on $\mathrm{supp}(S) \cap \mathrm{supp}(T)$ and have a common extension. For the consequence: 2^m pairwise disjoint subsets of $\{0,1\}^n$ of size 2^{n-m} each have total size 2^n, so they cover. □

Note that H plays no role in Proposition 7.1: the partition property follows from pairwise adjacency alone, and realizability then says that each piece meets H.

Proposition† 7.2. *Let $\mathcal{Q}$ be a partition of $\{0,1\}^n$ into 2^m subcubes of codimension m, and for a coordinate $x \in [n]$ let Q_x^0, Q_x^1, Q_x^* be the pieces fixing x to* 0*, fixing x to* 1*, and not fixing x. Then $|Q_x^0| = |Q_x^1| =: a_x$ and $2a_x + |Q_x^*| = 2^m$. Consequently*

$$\sum_{x \in [n]} a_x = m\, 2^{m-1}, \qquad a_x \leq 2^{m-1}, \qquad \sum_{x \in [n]} a_x^2 \geq \binom{2^m}{2}.$$

Proof. A piece fixing x to 0 lies inside the halfspace $\{h : h(x) = 0\}$, which has 2^{n-1} points, and contributes 2^{n-m} of them; a piece not fixing x contributes half of its 2^{n-m} points; a piece fixing x to 1 contributes none. Hence $2^{n-m}|Q^0_x| + 2^{n-m-1}|Q^*_x| = 2^{n-1}$, i.e. $2|Q^0_x| + |Q^*_x| = 2^m$, and the same for the other halfspace. Summing $2a_x$ over x counts the pairs (piece, fixed coordinate), of which there are $m\,2^m$. Two pieces are disjoint exactly when some coordinate is fixed to opposite values in both, and coordinate x separates a_x^2 pairs; every one of the $\binom{2^m}{2}$ pairs is separated at least once. □

Corollary† 7.3. *If every coordinate of $[n]$ is fixed by some piece of $\mathcal{Q}$, then $n \le 3$ when $m = 2$ and $n \le 7$ when $m = 3$.*

Proof. Every a_x is then at least 1. At $m = 2$: $\sum a_x = 4$, $a_x \le 2$, $\sum a_x^2 \ge 6$; four coordinates would force $a_x = 1$ throughout and $\sum a_x^2 = 4$, so $n \le 3$, and $(2, 1, 1)$ is the only profile at $n = 3$. At $m = 3$: $\sum a_x = 12$, $a_x \le 4$, $\sum a_x^2 \ge 28$; with $n = 8$ the sum of squares is maximised by the profile $(4, 2, 1, 1, 1, 1, 1, 1)$ at $26 < 28$, so $n \le 7$ (where $(4, 3, 1, 1, 1, 1, 1)$ gives 30). □

Lemma† 7.4 (domain restriction)**.** *Let (S_σ) be a cube-trace clique of $G^{\neq}_m(H)$, let $U = \bigcup_\sigma \operatorname{supp}(S_\sigma)$, and let $H|_U = \{h|_U : h \in H\} \subseteq \{0,1\}^U$. Then (S_σ), read over the domain U, is a cube-trace clique of $G^{\neq}_m(H|_U)$, and $\mathrm{VC}(H|_U) \le \mathrm{VC}(H)$.*

Proof. Each S_σ has its points in U and is realized by the restriction of its realizer. A duplicate-free $H|_U$-realizable m-sample over U is a duplicate-free H-realizable m-sample over X, and its adjacencies with the S_σ are the same in both readings, so its trace is a subcube by hypothesis. A subset of U shattered by $H|_U$ is shattered by H. □

Together with Proposition 6.1, this makes Question 1.1 a finite problem at each m: a negative answer at order m would be one with $|H| \le 2^m$, with $X = U$ the union of the supports, hence (Proposition 7.1) a partition of $\{0,1\}^U$ into 2^m subcubes of codimension m using every coordinate of U, with one realizer chosen inside each piece, $|U| \le m\,2^m$ trivially and $|U| \le 7$ at $m = 3$ by Corollary 7.3.

Remark 7.5 (the two graphs). The vertex set of $G_m(H)$ is that of $G^{\neq}_m(H)$ plus the realizable sequences of support size less than m, and traces depend only on the set of entries. So the cube-trace condition of [1] asks for subcube traces at every realizable partial labeling with support of size *at most* m, the duplicate-free one at those of size *exactly* m; the proof of their Theorem 5.4 uses support size 1.

8. Order three: a computer-assisted answer

Theorem† 8.1 (computer-assisted)**.** *Let X be any set and $H \subseteq \{0,1\}^X$. If $G^{\neq}_3(H)$ contains a cube-trace clique of size 8, then $\mathrm{VC}(H) \ge 3$.*

Proof. Suppose not. By Proposition 6.1 and Lemma 7.4 there is a counterexample with $H = \{h_\sigma : \sigma \in \{0,1\}^3\}$ and $X = U$ the union of the eight supports, so $3 \le |U| \le 7$ by Corollary 7.3, and by Proposition 7.1 the clique is a partition $\mathcal{Q}$ of $\{0,1\}^U$ into eight subcubes of codimension three that uses every coordinate, with $h_\sigma \in C(S_\sigma)$, $\mathrm{VC}(H) < 3$, and every duplicate-free H-realizable 3-sample having a subcube trace under some bijection ϕ. Every duplicate-free H-realizable 3-sample over U is a three-element subset $R \subseteq U$ labeled by some h_σ. The census described below enumerates every such $(\mathcal{Q}, (h_\sigma))$ on 3 to 7 points, and for each one with $\mathrm{VC}(H) < 3$ tries every bijection ϕ; none makes all traces subcubes. □

8.1. **The census.** The computation is a C program of about 150 lines, run with parameters (n, m) for $m = 2$, $n = 2, \dots, 6$ and $m = 3$, $n = 3, \dots, 7$. For each (n, m) it does the following.

1. *Partitions.* It enumerates the partitions of $\{0,1\}^n$ into 2^m subcubes of codimension m by the lowest-uncovered-point recursion: given the set of points covered so far, take the least uncovered point p and, for each m-subset R of the coordinates, try the piece consisting of the points agreeing with p on R, provided it is disjoint from the covered set. Since the piece containing p is determined by the partition, each partition is generated exactly once.
2. *Full support.* It keeps the partitions in which every coordinate is fixed by some piece (by Lemma 7.4 the domain may be taken to be the union of the supports).

(3) *Realizers.* For each piece it lists the concepts h in that piece all of whose $\binom{n}{m}$ traces — the trace of the sample "R labeled by h" is the set of pieces whose fixed values agree with h on their fixed coordinates inside R — have a power of two as cardinality. This pre-filter is sound: in a counterexample every trace is a Boolean subcube, and a Boolean subcube has 2^k elements, so no realizer of a counterexample is discarded.

(4) *Leaves.* For each choice of one listed realizer per piece it tests whether some m-subset R is shattered (with 2^m concepts, R is shattered exactly when their restrictions to R are pairwise distinct); if so the class is not a counterexample. Otherwise it collects the distinct traces and searches the $(2^m)!$ bijections ϕ for one under which every trace is a Boolean subcube, pruning a partial bijection as soon as a trace inside its domain fails. A success is reported as a hit.

Table 1 gives the counts. There were no hits. The whole census takes under four seconds on one core, and it was re-run end to end for this paper.

m	n	partitions of $\{0,1\}^n$ into 2^m codim-m subcubes	full support	hits
2	2	1	1	0
2	3	9	6	0
2	4	30	0	0
2	5	70	0	0
2	6	135	0	0
3	3	1	1	0
3	4	272	268	0
3	5	4,380	3,030	0
3	6	30,020	7,800	0
3	7	132,685	5,040	0

TABLE 1. The census. "Full support" counts the partitions in which every coordinate is fixed by some piece; "hits" counts the pairs (partition, realizer assignment) with VC $< m$ and a bijection making every trace a Boolean subcube. Computed, not proved.

8.2. **Controls on the census.** A search that reports no counterexample is evidence about the search unless it also finds what it should. Three checks were run.

(1) *Positive control.* With the VC test inverted — accepting only classes that *do* shatter an m-set — the same code reports 1, 12, 1 and 456 hits at $(m,n) = (2,2)$, $(2,3)$, $(3,3)$, $(3,4)$: it finds cube-trace cliques where Proposition 3.1 says they exist, through the same realizer filter and the same bijection search. (Section 12 looks at what these hits are.)

(2) *The filter rejects nothing at* $(3,4)$. The leaf count there is $68{,}608 = 268 \cdot 2^8$, every assignment of realizers to every full-support partition, so that cell is a complete check with the pre-filter inert; a copy of the program with the filter removed from the code, run at $(3,4)$ for this paper, gives the same 68,608 leaves and no hit. At $m = 3$, $n \geq 5$ the filter empties every leaf set; its soundness is a proof (step 3 of Section 8.1), not a measurement. In the founding computation a run with the filter disabled at $(3,5)$, about $3{,}030 \cdot 4^8 \approx 2 \cdot 10^8$ leaves, did not finish in twenty minutes and was stopped; it was not repeated.

(3) *Cross-checks on the counts.* A subcube of codimension $n-1$ in $\{0,1\}^n$ is an edge of the n-cube, so the entries at $(m,n) = (2,3)$ and $(3,4)$ are the numbers of perfect matchings of Q_3 and Q_4, namely 9 and 272 (OEIS A005271 [3]: $1, 2, 9, 272, 589185, \ldots$). The $m = 2$ totals agree with the closed form of Remark 8.2. The full-support column at $m = 2$ is 0 from $n = 4$ on, as Corollary 7.3 predicts (the census did not go past $n = 7$ at $m = 3$, where the corollary puts the first 0 at $n = 8$).

Remark 8.2 (the $m = 2$ column). The classification in the proof of Theorem 4.2 uses only pairwise contradiction, so over $[n]$ a partition of $\{0,1\}^n$ into four subcubes of codimension two is either the four labelings of one pair, $\binom{n}{2}$ ways, or a star: a hub p, a petal q paired with the label 0 at p and a petal r with the label 1, $n(n-1)(n-2)$ ordered choices. So the count is $\binom{n}{2}+n(n-1)(n-2) = \frac{1}{2}n(n-1)(2n-3)$, which is $1, 9, 30, 70, 135$ at $n = 2, \dots, 6$. An OEIS search for these five terms returns only A002414, the octagonal pyramidal numbers $k(k+1)(2k-1)/2$ [3]; at $k = n-1$ that is the same polynomial, so the agreement is an identity of formulas for a different object, not a coincidence of five terms. The $m = 3$ rows of Table 1, and the full-support columns, are not in the OEIS (searched 2026-09-03).

Remark 8.3. The partitions counted here are the subcube partitions of the n-cube whose pieces all have the same codimension m. Partitions of $\{0,1\}^n$ into subcubes of arbitrary sizes are counted by OEIS A018926 [3] $(1, 2, 8, 154, 89512, 71319425714$ for $n = 0, \dots, 5)$; the equal-codimension refinement is a proper sub-family that A018926 does not separate.

9. Order four: a counterexample

From here on the answer to Question 1.1 is negative. This section gives the explicit witness at $m = 4$ and its controls; all statements in it are formalised and machine-checked (Section 14).

9.1. **The class H_4.** Let $X_4 = \{y_1, y_2, y_3, a, b\}$, five distinct points. For $v \in \{0,1\}^3$ write $\mathrm{wt}(v) = v_1 + v_2 + v_3$ for its Hamming weight, and colour v by a bottom point:

$$c(v) = \begin{cases} a & \mathrm{wt}(v) \le 1, \\ b & \mathrm{wt}(v) \ge 2, \end{cases} \qquad \bar{c}(v) \in \{a, b\} \setminus \{c(v)\}.$$

The nodes 000 and 111 are *interior*: each has the same colour as all three of its neighbours in $\{0,1\}^3$ (000 and its neighbours have weight ≤ 1; 111 and its neighbours have weight ≥ 2). Every other node has a neighbour of the other colour. The clique is indexed by $\sigma = (v, \beta) \in \{0,1\}^3 \times \{0,1\}$, identified with $\{0,1\}^4$:

$$S_{v,\beta} = \big((y_1, v_1), (y_2, v_2), (y_3, v_3), (c(v), \beta)\big),$$

a duplicate-free 4-sample. Its realizer $h_{v,\beta} \in \{0,1\}^{X_4}$ is

$$h_{v,\beta}(y_i) = v_i, \qquad h_{v,\beta}(c(v)) = \beta, \qquad h_{v,\beta}(\bar{c}(v)) = \gamma(v, \beta) := \begin{cases} 0 & v \in \{000, 111\}, \\ \beta & \text{otherwise}, \end{cases}$$

and $H_4 = \{h_{v,\beta} : (v, \beta) \in \{0,1\}^4\}$. Written as strings of values at (y_1, y_2, y_3, a, b), the sixteen concepts are

$$\begin{array}{llll} h_{000,0} = 00000 & h_{000,1} = 00010 & h_{001,0} = 00100 & h_{001,1} = 00111 \\ h_{010,0} = 01000 & h_{010,1} = 01011 & h_{011,0} = 01100 & h_{011,1} = 01111 \\ h_{100,0} = 10000 & h_{100,1} = 10011 & h_{101,0} = 10100 & h_{101,1} = 10111 \\ h_{110,0} = 11000 & h_{110,1} = 11011 & h_{111,0} = 11100 & h_{111,1} = 11101. \end{array}$$

They are pairwise distinct (v is read off the y_i and β off $c(v)$), so $|H_4| = 16$.

Theorem 9.1. *The family $(S_{v,\beta})_{(v,\beta)\in\{0,1\}^4}$ is a cube-trace clique of size $16 = 2^4$ in $G_4^{\neq}(H_4)$, and $\mathrm{VC}(H_4) < 4$. Hence Question 1.1 has a negative answer at $m = 4$.*

Proof. There are six things to check. Five of them are finite statements over the definitions and are decided by the Lean kernel (declarations `S_inj`, `S_dup`, `S_real`, `S_clique`, `not_vc4`, all by `decide`); we indicate why they hold. *Injective*: $S_{v,\beta}$ determines v and β. *Duplicate-free*: $y_1, y_2, y_3, c(v)$ are four distinct points. *Realizable*: $h_{v,\beta}$ realizes $S_{v,\beta}$ by construction. *Clique*: if $v \ne v'$ the two samples carry opposite labels at some y_i; if $v = v'$ and $\beta \ne \beta'$ they carry opposite labels at $c(v)$. $\mathrm{VC}(H_4) < 4$: a four-point subset of X_4 omits one point z, and with sixteen concepts it is shattered exactly when restriction to it is injective on H_4; for each z two concepts agree off z: $h_{000,0}, h_{000,1}$ agree off a (both label b by $\gamma = 0$, because 000 is interior), $h_{111,0}, h_{111,1}$ agree off b (likewise at 111), and $h_{v,0}, h_{v\oplus e_t,0}$ agree off y_t (every $h_{v,0}$ labels both a and b by 0).

The sixth is the trace condition, which is not a finite statement as it stands (the subcube predicate is an existential over subsets of $\{0,1\}^4$). Supplying a decision procedure makes it decidable, but evaluating that procedure over all 10^4 length-4 sequences inside the kernel reached 30.6 GB of memory and was still growing when it was stopped after 230 s of CPU: this is run C1 of the timing table in the research record of 2026-09-04, quoted here and not repeated. The condition is therefore reduced by hand to a finite lemma.

(K1) A duplicate-free 4-sample T over X_4 misses exactly one point z: the image of its injective point map has four elements, so its complement in X_4 has one (`exists_missing`).

(K2) If T misses exactly z and is realized by h_ρ, then for every σ: $S_\sigma \not\sim T$ if and only if every entry (x, ℓ) of S_σ with $x \neq z$ has $\ell = h_\rho(x)$. Indeed every point of S_σ other than z occurs in T with the label h_ρ gives it, and a contradiction point lies in $\mathrm{supp}(T)$, hence differs from z (`trace_eq`).

(K3) For each of the $5 \cdot 16$ pairs (z, ρ), the set of σ satisfying the condition in (K2) is $\{\sigma : \sigma_i = \rho_i \text{ for all } i \in I(z, \rho)\}$, where, writing $\rho = (v, \beta)$,

$$I(z,\rho) = \begin{cases} [4] \setminus \{t\} & z = y_t, \\ \{1,2,3\} & z = c(v), \\ [4] & z = \bar{c}(v). \end{cases}$$

This is a statement about $5 \cdot 16 \cdot 16 = 1{,}280$ triples (z, ρ, σ) and is decided by the kernel (`K3f`, by `decide +kernel`). In words: dropping y_t frees direction t — the trace is $\{(v, \beta), (v \oplus e_t, \beta)\}$, where the neighbour's label at its own bottom point is read from $h_{v,\beta}$, and equals β because either the neighbour has the colour $c(v)$ or v is not interior, so $\gamma(v, \beta) = \beta$; dropping $c(v)$ frees β — the trace is $\{(v, 0), (v, 1)\}$; dropping $\bar{c}(v)$ leaves the single point ρ.

Given a duplicate-free H_4-realizable T, (K1) gives z, its realizer is some h_ρ, and (K2)+(K3) exhibit $\{\sigma : S_\sigma \not\sim T\}$ as the Boolean subcube with fixed set $I(z, \rho)$ and values ρ (`S_cube`). The six fields assemble into the formal structure (`counter4`). □

9.2. **Controls.**

Proposition 9.2 (extremality). $\mathrm{VC}(H_4) = 3$ *exactly: the set* $\{y_1, y_2, y_3\}$ *is shattered and no four-point set is. And* $|H_4| = 16 = 2^4$. *Consequently the conclusion* $\mathrm{VC}(H_4) < 4$ *of Theorem 9.1 cannot be strengthened to* $\mathrm{VC}(H_4) \leq 2$*, and the class has the least number of concepts a counterexample at order* 4 *can have (Proposition 6.1).*

Proof. $\{y_1, y_2, y_3\}$ is shattered because $h_{v,0}$ restricts to v there; no four-point set is shattered by Theorem 9.1; $|H_4| = 16$ because $(v, \beta) \mapsto h_{v,\beta}$ is injective (declarations `vc3`, `H4_vc_exact`, `H4_card`). □

The domain, too, is as small as it can be: over a domain of m points every duplicate-free m-sample has the whole domain as its support, so a 2^m-clique has a common support and Proposition 4.4 of [1] makes it shattered; a counterexample needs at least $m + 1$ points, and H_4 lives on $4 + 1 = 5$. (This sentence is ordinary mathematics, not part of the formal development.)

Proposition 9.3 (the repeated vertex). *The repeated sample* $T_a = ((a, 0), (a, 0), (a, 0), (a, 0))$ *is* H_4*-realizable, hence a vertex of* $G_4(H_4)$*, and is not duplicate-free. Its non-neighbour trace on the clique of Theorem 9.1 is*

$$\{S_{v,\beta} : c(v) = b\} \cup \{S_{v,0} : c(v) = a\},$$

a set of 12 *members, which is not a Boolean subcube under any bijection. Consequently no bijection makes* $(S_{v,\beta})$ *a cube-trace clique of* $G_4(H_4)$*: the same sixteen samples are a cube-trace clique of the duplicate-free graph and of the repeated-sample graph they are not.*

Proof. $h_{000,0}$ realizes T_a; T_a repeats a; a clique member is adjacent to T_a exactly when it contains $(a, 1)$, i.e. when $c(v) = a$ and $\beta = 1$, which leaves the $8 + 4 = 12$ members displayed; and 12 is not a power of two, so the trace is not a subcube. The first four facts are finite checks (`rep_realizable`, `rep_not_dupFree`, `rep_trace_card`, `rep_not_subcube`, all by `decide`), and the last sentence follows because the cube-trace condition of $G_4(H_4)$ is imposed at T_a: for the indexing of Theorem 9.1 that is the formal statement `S_not_cubeTraceRep`, and for an arbitrary bijection it follows from the cardinality 12 (`rep_trace_card`), since a Boolean subcube of $\{0, 1\}^4$ has 1, 2, 4, 8 or 16 elements. □

This is exactly the vertex the proof of Theorem 5.4 of [1] uses (a is an informative point: $(a,0) \in S_{000,0}$ and $(a,1) \in S_{000,1}$), and the vertex the duplicate-free graph deletes. Theorem 9.1 therefore contradicts nothing in [1]; it settles precisely the variant its Discussion says the proof does not reach. Proposition 5.2 is the order-2 analogue of this control, on a class that has no cube-trace clique at all.

10. Every order $m \geq 4$: the padding lift

A single counterexample propagates upward. Adjoin a fresh point z to the domain, let every concept take both values at z, and append (z,β) to every clique member. The lift is coordinate-free and transports the three extremal parameters, so the general-m statement is proved once, from H_4, by induction; all statements of this section are formalised and machine-checked.

10.1. **Set-up and the side condition.** Let $X' = X \sqcup \{z\}$. For $h \in \{0,1\}^X$ and $\beta \in \{0,1\}$ let $h^\beta \in \{0,1\}^{X'}$ extend h by $h^\beta(z) = \beta$, and let $H' = \{h^\beta : h \in H,\ \beta \in \{0,1\}\}$. For a sample S over X let $S \mathbin{+\!\!+} (z,\beta)$ be S with the entry (z,β) appended. Given a family $(S_\sigma)_{\sigma\in\{0,1\}^m}$ over X, put

$$S'_{\sigma,\beta} = S_\sigma \mathbin{+\!\!+} (z,\beta), \qquad (\sigma,\beta) \in \{0,1\}^m \times \{0,1\} \cong \{0,1\}^{m+1}.$$

The trace condition for a sample that avoids z will be an intersection of subcubes, which may be empty; what makes it nonempty is

(RC) every $h \in H$ realizes some member S_σ of the clique,

a condition that the realizer reduction (Proposition 6.1) produces automatically and that H_4 satisfies by construction ($h_{v,\beta}$ realizes $S_{v,\beta}$).

Lemma 10.1. (a) *A nonempty intersection of finitely many Boolean subcubes of $\{0,1\}^m$ is a Boolean subcube.* (b) *If $A \subseteq \{0,1\}^m$ is a Boolean subcube, then $\{(\sigma,\beta) : \sigma \in A,\ \beta = \beta_0\}$ and $\{(\sigma,\beta) : \sigma \in A\}$ are Boolean subcubes of $\{0,1\}^{m+1}$.*

Proof. (a) If ρ lies in every $A_j = \{\sigma : \sigma_i = \tau^j_i\ (i \in I_j)\}$, then $\bigcap_j A_j = \{\sigma : \sigma_i = \rho_i\ (i \in \bigcup_j I_j)\}$. (b) Add the new coordinate to the fixed set with value β_0, or leave it free (`isSubcube_iInter`, `isSubcube_snoc_fix`, `isSubcube_snoc_free`). □

10.2. **The lift.**

Theorem 10.2 (padding lift)**.** *Let $m \geq 1$ and let $(S_\sigma)_{\sigma\in\{0,1\}^m}$ be a cube-trace clique of $G^{\neq}_m(H)$ satisfying* (RC)*. Then $(S'_{\sigma,\beta})_{(\sigma,\beta)\in\{0,1\}^{m+1}}$ is a cube-trace clique of $G^{\neq}_{m+1}(H')$ satisfying* (RC)*. Moreover* $\mathrm{VC}(H) \geq d$ *implies* $\mathrm{VC}(H') \geq d+1$*;* $\mathrm{VC}(H) < m$ *implies* $\mathrm{VC}(H') < m+1$*;* $|H'| = 2|H|$ *and* $|X'| = |X| + 1$ *when H and X are finite.*

Proof. Clique fields. $S'_{\sigma,\beta}$ is duplicate-free because z is new, realized by h^β_σ where h_σ realizes S_σ, and $S'_{\sigma,\beta} \sim S'_{\sigma',\beta'}$ for $(\sigma,\beta) \neq (\sigma',\beta')$ inside X if $\sigma \neq \sigma'$ and at z otherwise; injectivity is clear. (RC) for H': h^β realizes $S'_{\sigma,\beta}$ when h realizes S_σ (`liftS_rc`).

Trace field. Let T' be a duplicate-free $(m+1)$-sample over X' realized by h^{β_0}, $h \in H$.

Case $z \in \mathrm{supp}(T')$. Delete the entry at z: $T' = T \mathbin{+\!\!+} (z,\beta_0)$ up to order, with T a duplicate-free m-sample over X realized by h. A lifted member $S'_{\sigma,\beta}$ contradicts T' if and only if $S_\sigma \sim T$ or $\beta \neq \beta_0$ (`padSeq_adj_iff_mem`), so the trace of T' is $\{\sigma : S_\sigma \not\sim T\} \times \{\beta_0\}$: the trace of T, a subcube by hypothesis, with one more coordinate fixed (Lemma 10.1(b)).

Case $z \notin \mathrm{supp}(T')$. Then T' is a duplicate-free $(m+1)$-sample U over X realized by h, and $S'_{\sigma,\beta} \sim T'$ if and only if $S_\sigma \sim U$ (no contradiction can occur at z). Let $U \setminus i$ be U with its i-th entry deleted, a duplicate-free m-sample realized by h. A contradiction between S_σ and U at the entry j of U survives every deletion $i \neq j$, and there is such an i because $m+1 \geq 2$; so $S_\sigma \sim U$ if and only if $S_\sigma \sim U \setminus i$ for some i (`padSeq_adj_iff_notMem`). Hence the trace of T' is

$$\Big(\bigcap_{i\in[m+1]} \{\sigma : S_\sigma \not\sim U \setminus i\}\Big) \times \{0,1\},$$

an intersection of $m+1$ subcubes (each $U \setminus i$ is a vertex of $G^{\neq}_m(H)$) times a free coordinate. By (RC) h realizes some S_{σ_0}, and two samples realized by one concept are never adjacent, so σ_0 lies in every one of the $m+1$ sets; the intersection is nonempty, hence a subcube by Lemma 10.1(a), and so is its product with $\{0,1\}$ by Lemma 10.1(b) (`lift_cube`).

VC dimension. If $R \subseteq X$ with $|R| = d$ is shattered by H, then $R \cup \{z\}$ is shattered by H': extend a realizer of a labeling on R by the required value at z (`lift_vcge`). If $R' \subseteq X'$ with $|R'| = m+1$ were shattered by H', then $R' \setminus \{z\}$ contains an m-set $R \subseteq X$, and R is shattered by H: a labeling g of R extended by 0 at z is realized by some h^β, whose restriction h realizes g (`lift_not_vcge`).

Sizes. $(h,\beta) \mapsto h^\beta$ is a bijection from $H \times \{0,1\}$ onto H' (`Hpad_card`), and X' has one more point than X. □

Theorem 10.3. *For every $m \geq 4$ there are a set X with $|X| = m+1$ and a class $H \subseteq \{0,1\}^X$ with $|H| = 2^m$ and $\mathrm{VC}(H) = m-1$ — that is, $\mathrm{VC}(H) \geq m-1$ and $\mathrm{VC}(H) < m$ — such that $G^{\neq}_m(H)$ contains a cube-trace clique of size 2^m.*

Proof. Induction on m, carrying the invariant "a cube-trace clique of $G^{\neq}_m(H)$ satisfying (RC), $|X| = m+1$, $|H| = 2^m$, $\mathrm{VC}(H) \geq m-1$, $\mathrm{VC}(H) < m$" (`Lift.Witness`, `counter_step`). The base $m = 4$ is H_4: Theorem 9.1, Proposition 9.2, and (RC) by construction (`H4_realizer_clique`). The step is Theorem 10.2 with $d = m-1$ (`lift_witness`, `lift_vcge`, `lift_not_vcge`, `Hpad_card`). The formal statement is `counter_all`, quoted in Appendix A. □

Corollary† 10.4 (the sharp threshold)**.** *Let $m \geq 2$. The implication "if $G^{\neq}_m(H)$ contains a cube-trace clique of size 2^m then $\mathrm{VC}(H) \geq m$" holds for every set X and every $H \subseteq \{0,1\}^X$ when $m = 2$ (Theorem 4.3, formal) and when $m = 3$ (Theorem† 8.1, computer-assisted), and fails for every $m \geq 4$ (Theorem 10.3, formal).*

The corollary carries a dagger only because its $m = 3$ case does. (At $m = 1$ the implication is immediate: two adjacent realizable 1-samples are $(x,0)$ and $(x,1)$ for one point x, which is then shattered.)

Remark 10.5. The lift is stated for a general class and needs (RC); without it the intersection in the second case can be empty, and the empty set is not a Boolean subcube (Definition 3.5 of [1] demands nonemptiness). An earlier draft of the argument asserted that the intersection contains the index of "the clique member realized by h", which need not exist for an arbitrary H; the repair is to carry (RC) as part of the induction hypothesis. For the classes produced by Proposition 6.1 the condition is automatic.

11. The closed form, the five-point census, and the order-3 degeneration

Except for Proposition 11.2, this section is outside the formal development: the closed-form family is proved in ordinary mathematics and the census is a computation, and both carry a dagger. Theorem 10.3 does not depend on anything here.

11.1. **The family $K(m,c,\gamma)$.** Fix $m \geq 2$. Let $X = \{y_1, \ldots, y_{m-1}, a, b\}$, $n = m+1$, and $V = \{0,1\}^{m-1}$, with $N[v]$ the closed neighbourhood of v in the cube V (v and its $m-1$ Hamming neighbours). A *colouring* is a map $c\colon V \to \{a,b\}$; $\bar{c}(v)$ is the other bottom point; v is *interior* for c if c is constant on $N[v]$. A *twist* is a map $\gamma\colon V \times \{0,1\} \to \{0,1\}$ with $\gamma(v,\beta) = \beta$ at every non-interior v (and arbitrary at interior v). Put, for $(v,\beta) \in V \times \{0,1\} \cong \{0,1\}^m$,

$$S_{v,\beta} = \big((y_1,v_1), \ldots, (y_{m-1},v_{m-1}), (c(v),\beta)\big), \qquad h_{v,\beta}(y_i) = v_i,\ h_{v,\beta}(c(v)) = \beta,\ h_{v,\beta}(\bar{c}(v)) = \gamma(v,\beta),$$

and $K(m,c,\gamma) = \{h_{v,\beta}\}$, a class of 2^m distinct concepts. The class H_4 of Section 9 is $K(4,c_0,\gamma_0)$ with $c_0 = a$ on $N[000]$ and b elsewhere, and $\gamma_0(v,\cdot) \equiv 0$ at the two interior nodes.

Proposition† 11.1. *Let $m \geq 2$, c a colouring and γ a twist.*

(i) *$(S_{v,\beta})$ is a cube-trace clique of $G^{\neq}_m(K(m,c,\gamma))$ under the indexing (v,β), for* every *c and γ.*
(ii) *If c has an interior node of each colour and $\gamma(v,\cdot)$ is constant at every interior node v, then $K(m,c,\gamma)$ shatters no m-subset of X, so* $\mathrm{VC}(K(m,c,\gamma)) = m-1$.

(iii) *Within the family, interior nodes of both colours are necessary for* VC $< m$*: if* $K(m, c, \gamma)$ *shatters no* m*-set, then* c *has an interior node of each colour, and* γ *is constant at one interior node of each colour. Two interior nodes of different colours have disjoint closed neighbourhoods, i.e. are at Hamming distance at least* 3 *in* $V = \{0,1\}^{m-1}$*, and such a pair exists if and only if* $m - 1 \geq 3$*. So the family contains a counterexample to Question 1.1 exactly when* $m \geq 4$*.*

Proof. (i) Distinctness, duplicate-freeness and realizability are as in Theorem 9.1, and two members contradict at some y_i if their top parts differ and at $c(v)$ otherwise. For the trace condition let T be a duplicate-free m-sample realized by $h_{v,\beta}$; since $|X| = m+1$, T omits exactly one point z and its labels are those of $h_{v,\beta}$ off z, so $S_{v',\beta'} \not\sim T$ if and only if its labels agree with $h_{v,\beta}$ off z. If $z = \bar{c}(v)$: agreement on the y_i forces $v' = v$, and then $c(v') = c(v) \in \mathrm{supp}(T)$ forces $\beta' = \beta$; the trace is the point $\{(v,\beta)\}$. If $z = c(v)$: $v' = v$ and β' is free; the trace is $\{(v,0),(v,1)\}$. If $z = y_t$: agreement off y_t gives $v' \in \{v, v \oplus e_t\}$; for $v' = v$, $\beta' = \beta$; for $v' = v \oplus e_t$, $\beta' = h_{v,\beta}(c(v'))$, which is β if $c(v') = c(v)$ and $\gamma(v,\beta)$ if $c(v') = \bar{c}(v)$ — and in the latter case v has a neighbour of the other colour, is not interior, and $\gamma(v,\beta) = \beta$. Either way the trace is the edge $\{(v,\beta),(v \oplus e_t, \beta)\}$. All three are Boolean subcubes.

(ii) An m-subset of X omits one point z, and with 2^m concepts it is shattered exactly when restriction to it is injective; we exhibit two concepts agreeing off z. If $z = a$, let u be an interior node of colour a: $h_{u,0}, h_{u,1}$ agree on the y_i and at $b = \bar{c}(u)$ ($\gamma(u,\cdot)$ is constant) and differ only at a. If $z = b$, the same at an interior node of colour b. If $z = y_t$: consider a pair $v, v' = v \oplus e_t$. If $c(v) \neq c(v')$, both nodes have a neighbour of the other colour, so neither is interior, and $h_{v,\beta}, h_{v',\beta}$ label both bottom points by β: they differ only at y_t. If $c(v) = c(v')$ and the two are not both interior with different constants, choose β to be the constant of an interior member of the pair — the two constants coincide when both members are interior — and any β when neither is interior; again $h_{v,\beta}$ and $h_{v',\beta}$ agree at $c(v)$ and at $\bar{c}(v)$. The remaining possibility — both interior, of the same colour, with different constants — cannot occur for every pair in direction t: then every node would be interior, c would be constant on the connected cube V, and c would not have both colours. Finally $\{y_1, \ldots, y_{m-1}\}$ is shattered, so VC $= m - 1$.

(iii) The set $X \setminus \{a\}$ is not shattered, so two concepts agree off a. Two concepts agreeing off a must agree on the y_i, so have the same top v; if $c(v) = b$ they agree at b too and are equal; so $c(v) = a$, they are $h_{v,0}, h_{v,1}$, and agreement at b means $\gamma(v,0) = \gamma(v,1)$, which forces v interior. The same at b. If u, w are interior of different colours then $N[u] \cap N[w] = \emptyset$, since a common neighbour would have both colours; disjoint closed neighbourhoods in a cube means distance ≥ 3, which $\{0,1\}^1$ and $\{0,1\}^2$ do not afford and $\{0,1\}^{m-1}$ does for $m - 1 \geq 3$ (antipodal nodes). Conversely for $m \geq 4$ the colouring c_0 above has the interior nodes 0 and 1 of colours a and b. □

Machine checks of the closed form, from the definitions and independent of the formal development: for $m = 2, \ldots, 7$ the family $K(m, c_0, \gamma_0)$ is cube-trace under the natural indexing, shatters an m-set at $m = 2, 3$ (the sets $\{y_1, a\}$ and $\{y_1, y_2, a\}$) and shatters none at $m = 4, 5, 6, 7$ (re-run for this paper, 11 s in Python). The threshold, stated in the kernel, is the next proposition.

11.2. **The order-3 instance is a positive instance.**

Proposition 11.2. *Let* $X_3 = \{y_1, y_2, a, b\}$ *and let* $H_3 = K(3, c_0, \gamma_0)$ *with* $c_0 = a$ *on* $\{00, 01, 10\}$ *and* b *on* $\{11\}$*: eight concepts on four points. Then* $G_3^{\neq}(H_3)$ *contains a cube-trace clique of size* 8*, and* $\mathrm{VC}(H_3) = 3$*.*

Proof. Only 00 is interior (11 has the a-coloured neighbours 01 and 10), so $\gamma_0(11, \beta) = \beta$, no two concepts agree off b, and $\{y_1, y_2, a\}$ is shattered; no four-point set is shattered because $|X_3| = 4$ and $|H_3| = 8 < 16$. The clique fields and the trace condition are finite checks: here the domain is small enough for the kernel to evaluate a decision procedure for the subcube predicate at all 8^3 length-3 sequences (`Three.S_cube`, by `decide +kernel`; `three_cubeTraceDF`, `three_vc`). □

So the construction run one order down still produces a cube-trace clique — the trace condition never sees the colouring — but not a counterexample, exactly as Theorem† 8.1 demands; the negative answer begins at $m = 4$.

11.3. **The census over five points at order four.** By Proposition 6.1 and Lemma 7.4, a counterexample at order 4 whose supports span five points is a partition of $\{0,1\}^5$ into sixteen subcubes of codimension 4 with one realizer per piece. A subcube of codimension 4 in $\{0,1\}^5$ is an edge of the 5-cube, so the partition is a perfect matching of Q_5; a realizable duplicate-free 4-sample is an edge $\{h, h \oplus e\}$ through a realizer h (its labels are those of h off the coordinate e), and a piece is non-adjacent to it exactly when the piece's edge meets $\{h, h \oplus e\}$. So every trace has one or two members, and the cube-trace condition says that some bijection to $\{0,1\}^4$ sends every two-element trace to an edge of Q_4. Each piece σ with realizer h_σ and direction f_σ has the four two-element traces $\{\sigma, \pi(h_\sigma \oplus e)\}$, $e \neq f_\sigma$, where π is the piece containing a point, and the four pieces $\pi(h_\sigma \oplus e)$ are distinct (the points are pairwise at distance 2, so no edge holds two of them); hence the graph Γ on the pieces with these edges is 4-regular and the bijection is an isomorphism $\Gamma \cong Q_4$. In particular Γ must be undirected.

Proposition† 11.3 (computed)**.** *Over a five-point domain: there are* 589,185 *perfect matchings of* Q_5*;* 3,025 *of them admit at least one assignment of realizers making the clique cube-trace; there are* 439,000 *cube-trace pairs (matching, realizer assignment) in all, of which* 640 *have* $\mathrm{VC} < 4$ *and* 438,360 *have* $\mathrm{VC} = 4$*. The* 640 *counterexamples consist of* 640 *distinct classes on* 80 *distinct cliques, and they are exactly the classes* $K(4, c, \gamma)$ *with* $\mathrm{VC} < 4$ *— the three top points chosen among the five, c a colouring of* $\{0,1\}^3$ *with an interior node of each colour, γ constant at the two interior nodes, and, at the non-interior nodes, either $\gamma(v, \beta) = \beta$, which is a twist in the sense of Section 11.1, or uniformly $\gamma(v, \beta) = 1 - \beta$, which is not a twist and whose clique is cube-trace under a re-indexed bijection but never under the natural indexing.*

What was computed. A C program enumerates the perfect matchings by the lowest-uncovered-vertex recursion, and for each matching every choice of one endpoint per edge as its realizer, pruned only by the necessary condition that Γ be undirected (checked pairwise as realizers are assigned); at each leaf it tests $\Gamma \cong Q_4$ by backtracking and tests $\mathrm{VC} \geq 4$ (with sixteen concepts, a four-point set is shattered exactly when deleting the fifth coordinate keeps the realizers distinct). The program was written for this paper from the definitions above and runs in 10 s; an earlier program in the research record, which factors out the pieces whose four neighbouring pieces are all parallel to them, reports the same matching counts and the same 640 and 438,360; in place of the 439,000 it reports 6,205 "cube-trace pairs after factoring", the sum over matchings of the number of cube-trace assignments divided by $2^{\text{free pieces}}$, which our count reproduces exactly, the division being exact for every matching. The number of perfect matchings agrees with OEIS A005271 [3] ($a(5) = 589185$). The identification of the 640 with the family $K(4, c, \gamma)$ was checked by generating that family from its definition ($\binom{5}{3}$ choices of the top points, 2^8 colourings, the constants at the interior nodes, the two conventions at the non-interior nodes), discarding the members with $\mathrm{VC} = 4$, and comparing the two sets of (clique, class) pairs: they are equal. Each of the 80 cliques has three hub coordinates, two interior nodes and both interior colours; $80 = \binom{5}{3} \cdot 8$, the 8 being the colourings of $\{0,1\}^3$ with an interior node of each colour (an antipodal pair $\{u, \bar{u}\}$, four choices, and which of the two is coloured a). □

The census says that Proposition 11.1 is not one family of counterexamples among several on five points: it is all of them. Whether every counterexample at every $m \geq 4$ is a $K(m, c, \gamma)$, and whether one can live on more than $m + 1$ points, is open (Section 13).

12. The supports of a cube-trace clique

Version 1 of this paper conjectured (its Conjecture 9.1) that for every $m \geq 1$, every set X and every $H \subseteq \{0,1\}^X$ the members of a cube-trace clique of size 2^m in $G_m^{\neq}(H)$ all have the same support, and observed that this would answer Question 1.1 affirmatively at once through Proposition 4.4 of [1]. The conjecture is false at the first non-trivial order.

Let $H_{\mathrm{tw}} = \{000, 011, 100, 111\} \subseteq \{0,1\}^{\{p,q,r\}}$, concepts written as their values at p, q, r — so r carries the same label as q on every concept: q and r are *twins* — and let

$$S_{00} = ((p,0),(q,0)), \quad S_{01} = ((p,0),(q,1)), \quad S_{10} = ((p,1),(r,0)), \quad S_{11} = ((p,1),(r,1)).$$

Proposition 12.1 (the twin star)**.** *$(S_\sigma)_{\sigma\in\{0,1\}^2}$ is a cube-trace clique of size 4 in $G_2^{\neq}(H_{\mathrm{tw}})$; its supports are $\{p,q\}$ and $\{p,r\}$, not all equal; and $\mathrm{VC}(H_{\mathrm{tw}}) = 2$ (some two-point set is shattered, no three-point set is).*

Proof. All three statements are finite checks over the sixteen concepts on three points; the trace condition is evaluated by the kernel at all 6^2 length-2 sequences through a decision procedure for the subcube predicate (`St_cube`, by `decide +kernel`; `twin_cubeTraceDF`, `twin_supports_differ`, `twin_vc`). By hand: the ten realizable duplicate-free 2-samples and their traces are

$(p,0),(q,0)\colon \{00\}$	$(p,0),(q,1)\colon \{01\}$	$(p,1),(r,0)\colon \{10\}$	$(p,1),(r,1)\colon \{11\}$
$(p,1),(q,0)\colon \{10,11\}$	$(p,1),(q,1)\colon \{10,11\}$	$(p,0),(r,0)\colon \{00,01\}$	$(p,0),(r,1)\colon \{00,01\}$
$(q,0),(r,0)\colon \{00,10\}$	$(q,1),(r,1)\colon \{01,11\}$		

— points and edges of the square. □

The twin star is the star pattern of Theorem 4.2 in its Case 2a: the two members on $\{p,q\}$ disagree at r, and $\{p,r\}$ is shattered. Theorem 4.3 and its proof are unaffected, since Case 2a was settled by shattering and only Case 2b used the three witnesses; what fails is the stronger statement that the pattern cannot occur. A control keeps this from being misread: $\mathrm{VC}(H_{\mathrm{tw}}) = 2$, so the twin star refutes the conjecture of version 1 and not the theorem.

Remark 12.2 (measured)*.* The positive controls of Section 8 say more. At $m = 2$ the 12 cube-trace cliques with realizers found on three points lie on the 6 star partitions (a hub and an ordered pair of petals), each with two realizer assignments, $r \equiv q$ or $r \equiv 1 - q$ on the class; all have two distinct supports. At $m = 3$ the 456 hits on four points lie on 84 distinct partitions and every one of them has exactly two distinct supports, and there are no hits on 5, 6 or 7 points; by the two reductions — the proof of Proposition 6.1 passes a cube-trace clique to the class of its own realizers without using that proposition's hypothesis $\mathrm{VC}(H) < m$, and Lemma 7.4 then restricts the domain — every cube-trace 8-clique of $G_3^{\neq}(H)$, for any H, appears among these, so at $m = 3$ the supports of a cube-trace clique span at most four points and take at most two values. These are census outputs, re-computed for this paper with a program written from the definitions (see Section 14), not theorems.

13. What remains open

Question 1.1 is closed, and the closure raises four questions.

Is $\mathrm{VC} \geq m - 1$ the true statement? Every counterexample found — the 640 on five points at $m = 4$, and every lifted one — has $\mathrm{VC} = m - 1$ exactly. Does a cube-trace 2^m-clique in $G_m^{\neq}(H)$ force $\mathrm{VC}(H) \geq m - 1$? At $m = 2$ Theorem 4.3 gives more; nothing in this paper settles the question at any $m \geq 4$.

Are all counterexamples of the form $K(m, c, \gamma)$? True on five points at $m = 4$ (Proposition 11.3). The next cell is $(m, n) = (4, 6)$ — partitions of $\{0,1\}^6$ into sixteen subcubes of codimension 4 — whose size is unknown to us; the perfect-matching route that made $(4, 5)$ a ten-second computation does not apply there.

Is $|U| \leq m + 1$ always? The supports of a cube-trace duplicate-free clique span at most $m + 1$ points at $m = 2$ (Corollary 7.3) and at $m = 3$ (Remark 12.2), and the counterexamples at $(4, 5)$ are consistent with it. The halfspace bound of Proposition 7.2 gives only $|U| \leq 19$ at $m = 4$ (profile $(8, 7, 1^{17})$; on 20 coordinates the sum of squares is at most $118 < 120$), a true bound that is no longer needed. A proof of $|U| \leq m + 1$ would make Question 1.1 at each m a computation over a tiny domain and would settle the previous question at $m = 4$ outright.

Formalising order three. Theorem† 8.1 is now the other side of the threshold. It would become a machine-checked theorem given a verified enumerator for the subcube partitions of $\{0,1\}^n$ into 2^m pieces of codimension m (a proof that the lowest-uncovered-point recursion is complete) together with Corollary 7.3; the computation itself is seconds.

The unordered variant. A variant that [1] does not raise takes sets rather than sequences as vertices. For duplicate-free samples the two graphs coincide (Section 2), so every statement of this paper holds for it verbatim.

14. Verification

Propositions 3.1, 5.1, 5.2, 5.3, 6.1, 9.2, 9.3, 11.2 and 12.1, Lemmas 4.1 and 10.1, and Theorems 4.2, 4.3, 9.1, 10.2 and 10.3 are formalised and machine-checked in Lean 4 (version 4.32.0) [4] over Mathlib [5], in nine modules of 1,887 lines whose headline statements are quoted verbatim in Appendix A; the domain is an arbitrary type and the class an arbitrary set of Boolean-valued functions on it. The finite parts are decided by the kernel's `decide`: Lemma 4.1 over the four points of $\{0,1\}^2$; the controls over the eight concepts on three points and the four on two; the five clique and VC fields of Theorem 9.1 and the 1,280 instances of its finite lemma (K3) over the definitions of Section 9.1; the twin star and the order-3 instance, where a decision procedure for the subcube predicate (module `SubcubeDec`) is evaluated at all 6^2, respectively 8^3, sequences; and the twelve-element trace of the repeated vertex. There is no `native_decide`, no `sorry` and no added axiom, and every one of the forty declarations bound to the results of this paper — thirty in the five new modules, ten in the four modules of version 1 — depends only on `propext`, `Classical.choice` and `Quot.sound` (`#print axioms`, re-run for this draft on 2026-09-07; one of them, `rep_not_dupFree`, uses only `propext` and `Quot.sound`). The five new modules were timed once, in the research record of 2026-09-07, at 63 s of user CPU in total and a peak of 8.61 GB (`SubcubeDec` 4.60 s, `Counter4` 16.85 s, `Lift` 8.09 s, `TwinStar` 7.61 s, `Threshold` 25.81 s); those figures are quoted from that record and the modules were not re-elaborated for this draft. The memory is dominated by the Mathlib import, whose baseline alone is about 4.5 s and 6.9 GB. Sections 7, 8 and 11 (except Proposition 11.2) and Remark 12.2 are outside the formal development: the former is ordinary mathematics, the census of Section 8.1 rests on the C program described there, and the computations of Sections 11.3 and 12 were re-run for this version with two programs written from the definitions (10 s for the five-point census, under 7 s for the $m \le 3$ cells, one core), which agree with the research record's programs on every number; all programs are distributed with the formalisation. Repository reference to be added at submission.

Appendix A. The formal statements

The following are quoted verbatim from the Lean sources; `Fin m → Bool` is $\{0,1\}^m$, `X → Bool` is $\{0,1\}^X$, and the indexing `S : (Fin m → Bool) → Seq X m` carries the bijection ϕ of Definition 2.2 as explained in Section 2. A declaration ending in `:= by decide` is a finite check. The modules `Defs`, `DupFreeTwo`, `Reduction` and `Controls` are those of version 1; `SubcubeDec`, `Counter4`, `Lift`, `TwinStar` and `Threshold` are new.

Definitions (module `Defs`). The objects of Section 2.

```
abbrev Seq (X : Type*) (m : ℕ) := Fin m → X × Bool

def Realizable (H : Set (X → Bool)) (s : Seq X m) : Prop :=
  ∃ h ∈ H, ∀ i, h (s i).1 = (s i).2

def DupFree (s : Seq X m) : Prop :=
  ∀ i j, (s i).1 = (s j).1 → i = j

def Adj (s t : Seq X m) : Prop :=
  ∃ i j, (s i).1 = (t j).1 ∧ (s i).2 ≠ (t j).2

def IsSubcube (A : Set (Fin m → Bool)) : Prop :=
  ∃ (I : Finset (Fin m)) (τ : Fin m → Bool), ∀ σ, σ ∈ A ↔ ∀ i ∈ I, σ i = τ i

structure CubeTraceRep (H : Set (X → Bool)) (m : ℕ) where
  /-- the clique, indexed by the cube; the indexing IS the bijection `phi` of Definition 5.1 -/
  S : (Fin m → Bool) → Seq X m
  inj : Function.Injective S
```

```
  real : ∀ σ, Realizable H (S σ)
  clique : ∀ σ τ, σ ≠ τ → Adj (S σ) (S τ)
  cube : ∀ T : Seq X m, Realizable H T → IsSubcube {σ | ¬ Adj (S σ) T}

structure CubeTraceDF (H : Set (X → Bool)) (m : ℕ) where
  S : (Fin m → Bool) → Seq X m
  inj : Function.Injective S
  dup : ∀ σ, DupFree (S σ)
  real : ∀ σ, Realizable H (S σ)
  clique : ∀ σ τ, σ ≠ τ → Adj (S σ) (S τ)
  cube : ∀ T : Seq X m, Realizable H T → DupFree T → IsSubcube {σ | ¬ Adj (S σ) T}

def Shatters (H : Set (X → Bool)) (R : Finset X) : Prop :=
  ∀ g : X → Bool, ∃ h ∈ H, ∀ x ∈ R, h x = g x

def VCge (H : Set (X → Bool)) (d : ℕ) : Prop :=
  ∃ R : Finset X, R.card = d ∧ Shatters H R
```

The forward direction (module `Defs`; Proposition 3.1).

```
theorem cubeTrace_of_vcge (H : Set (X → Bool)) (m : ℕ) (h : VCge H m) :
    Nonempty (CubeTraceRep H m) ∧ Nonempty (CubeTraceDF H m)
```

Order two (module `DupFreeTwo`; Lemma 4.1, Theorems 4.2 and 4.3). `subcubeB` is a Boolean-valued form of the subcube predicate on an explicitly given subset of the finite cube, used so that `decide` can evaluate it; `Config` packages conditions (C1) and (C2) of Theorem 4.2, with `supp` the support predicate and `hr` the realizer.

```
def subcubeB {m : ℕ} (A : (Fin m → Bool) → Bool) : Bool :=
  decide (∃ (I : Finset (Fin m)) (τ : Fin m → Bool), ∀ σ, (A σ = true) ↔ ∀ i ∈ I, σ i = τ i)

theorem no_three_edges : ∀ a b c d : Fin 2 → Bool, a ≠ b → a ≠ c → a ≠ d → b ≠ c → b ≠ d →
    c ≠ d → ¬ (subcubeB (fun σ => σ == c || σ == d) = true ∧
      subcubeB (fun σ => σ == a || σ == c) = true ∧
      subcubeB (fun σ => σ == b || σ == c) = true) := by decide

structure Config (H : Set (X → Bool)) where
  supp : (Fin 2 → Bool) → X → Prop
  hr : (Fin 2 → Bool) → (X → Bool)
  mem : ∀ σ, hr σ ∈ H
  two : ∀ σ, ∃ u v : X, u ≠ v ∧ ∀ z, supp σ z ↔ (z = u ∨ z = v)
  contra : ∀ σ τ, σ ≠ τ → ∃ z, supp σ z ∧ supp τ z ∧ hr σ z ≠ hr τ z
  trace : ∀ (u v : X), u ≠ v → ∀ σ₀ : Fin 2 → Bool,
      IsSubcube {ρ | ∀ z, supp ρ z → (z = u ∨ z = v) → hr ρ z = hr σ₀ z}

theorem vcge_two_of_config {H : Set (X → Bool)} (cfg : Config H) : VCge H 2

theorem vcge_two_of_cubeTraceDF {H : Set (X → Bool)} (C : CubeTraceDF H 2) : VCge H 2
```

The realizer reduction (module `Reduction`; Proposition 6.1).

```
theorem exists_small_counterexample {H : Set (X → Bool)} (C : CubeTraceDF H m)
    (hvc : ¬ VCge H m) :
    ∃ (f : (Fin m → Bool) → (X → Bool)) (C' : CubeTraceDF (Set.range f) m),
      Set.range f ⊆ H ∧ ¬ VCge (Set.range f) m ∧ C'.S = C.S
```

Controls at order two (module `Controls`; Propositions 5.1, 5.2 and 5.3). The domain $\{p, q, r\}$ is `Fin 3` with p, q, r the coordinates $0, 1, 2$; `Hstar` is H^*, `sA`–`sD` are $s_A, \ldots, s_D$, `Tq` is T_q, and `Hfull` is the full class on `Fin 2`.

```
def Hstar : Set (Fin 3 → Bool) :=
  {h | h = ![false, false, false] ∨ h = ![false, true, false] ∨
       h = ![true, false, false] ∨ h = ![true, false, true]}

def sA : Seq (Fin 3) 2 := ![(0, false), (1, false)]
def sB : Seq (Fin 3) 2 := ![(0, false), (1, true)]
```

```
def sC : Seq (Fin 3) 2 := ![(0, true), (2, false)]
def sD : Seq (Fin 3) 2 := ![(0, true), (2, true)]

def Tq : Seq (Fin 3) 2 := ![(1, false), (1, false)]

theorem star_clique :
    (DupFree sA ∧ DupFree sB ∧ DupFree sC ∧ DupFree sD) ∧
    (Realizable Hstar sA ∧ Realizable Hstar sB ∧ Realizable Hstar sC ∧ Realizable Hstar sD) ∧
    (Adj sA sB ∧ Adj sA sC ∧ Adj sA sD ∧ Adj sB sC ∧ Adj sB sD ∧ Adj sC sD) := by decide

theorem star_vc : VCge Hstar 1 ∧ ¬ VCge Hstar 2 := by decide

theorem star_no_cubeTrace : IsEmpty (CubeTraceDF Hstar 2)

theorem not_subcube_three : ∀ a b c : Fin 2 → Bool, a ≠ b → a ≠ c → b ≠ c →
    subcubeB (fun σ => σ == a || σ == b || σ == c) = false := by decide

theorem repeated_witness :
    Realizable Hstar Tq ∧ ¬ DupFree Tq ∧
      (¬ Adj sA Tq ∧ Adj sB Tq ∧ ¬ Adj sC Tq ∧ ¬ Adj sD Tq) := by decide

def Hfull : Set (Fin 2 → Bool) := fun _ => True

theorem full_vc : VCge Hfull 2 ∧ ¬ VCge Hfull 3 := by decide

theorem full_cubeTraceDF : Nonempty (CubeTraceDF Hfull 2)
```

Order four (module `Counter4`; Theorem 9.1 and Proposition 9.2). The domain X_4 is `Fin 5` with y_1, y_2, y_3, a, b the coordinates $0, 1, 2, 3, 4$; `wt`, `col`, `S`, `other` and `real` are wt, c, $S_{v,\beta}$, γ and $h_{v,\beta}$ of Section 9.1, with $\sigma = (v, \beta)$ read as `σ 0, σ 1, σ 2` and `σ 3`; `trB z ρ` is the condition of (K2) and `Iof` the set $I(z, \rho)$ of (K3).

```
def wt (σ : Fin 4 → Bool) : ℕ :=
  (if σ 0 then 1 else 0) + (if σ 1 then 1 else 0) + (if σ 2 then 1 else 0)

def col (σ : Fin 4 → Bool) : Fin 5 := if wt σ ≤ 1 then 3 else 4

def S (σ : Fin 4 → Bool) : Seq (Fin 5) 4 := ![(0, σ 0), (1, σ 1), (2, σ 2), (col σ, σ 3)]

def other (σ : Fin 4 → Bool) : Bool := if wt σ = 0 ∨ wt σ = 3 then false else σ 3

def real (σ : Fin 4 → Bool) : Fin 5 → Bool :=
  ![σ 0, σ 1, σ 2, if col σ = 3 then σ 3 else other σ, if col σ = 4 then σ 3 else other σ]

def H4 : Set (Fin 5 → Bool) := {h | ∃ σ, real σ = h}

theorem S_inj : Function.Injective S := by decide

theorem S_dup : ∀ σ, DupFree (S σ) := by decide

theorem S_real : ∀ σ, Realizable H4 (S σ) := by decide

theorem S_clique : ∀ σ τ, σ ≠ τ → Adj (S σ) (S τ) := by decide

theorem not_vc4 : ¬ VCge H4 4 := by decide

theorem vc3 : VCge H4 3

theorem H4_realizer_clique : ∀ h ∈ H4, ∃ σ, ∀ i, h (S σ i).1 = (S σ i).2

theorem H4_card : Nat.card H4 = 16

theorem exists_missing (T : Seq (Fin 5) 4) (hT : DupFree T) :
```

```
    ∃ z : Fin 5, (∀ i, (T i).1 ≠ z) ∧ ∀ x : Fin 5, x ≠ z → ∃ i, (T i).1 = x

theorem trace_eq {T : Seq (Fin 5) 4} {ρ : Fin 4 → Bool} {z : Fin 5}
    (hh : ∀ i, real ρ (T i).1 = (T i).2) (hz : ∀ i, (T i).1 ≠ z)
    (hs : ∀ x : Fin 5, x ≠ z → ∃ i, (T i).1 = x) (σ : Fin 4 → Bool) :
    ¬ Adj (S σ) T ↔ ∀ i, (S σ i).1 ≠ z → (S σ i).2 = real ρ (S σ i).1

def trB (z : Fin 5) (ρ : Fin 4 → Bool) : (Fin 4 → Bool) → Bool :=
  fun σ => decide (∀ i, (S σ i).1 ≠ z → (S σ i).2 = real ρ (S σ i).1)

def Iof (z : Fin 5) (ρ : Fin 4 → Bool) : Finset (Fin 4) :=
  if z = 0 then Finset.univ.erase 0 else if z = 1 then Finset.univ.erase 1
  else if z = 2 then Finset.univ.erase 2 else if z = col ρ then {0, 1, 2} else Finset.univ

theorem K3f : ∀ z : Fin 5, ∀ ρ σ : Fin 4 → Bool, trB z ρ σ = true ↔ ∀ i ∈ Iof z ρ, σ i = ρ i := by
  decide +kernel

theorem S_cube : ∀ T : Seq (Fin 5) 4, Realizable H4 T → DupFree T →
    IsSubcube {σ | ¬ Adj (S σ) T}

theorem counter4 : Nonempty (CubeTraceDF H4 4) ∧ ¬ VCge H4 4
```

The padding lift (module `Lift`; Lemma 10.1, Theorems 10.2 and 10.3). The enlarged domain is `Option X` with $z =$ `none`; `pad β h` is h^β, `Hpad H` is H', `liftS S` is $(\sigma, \beta) \mapsto S'_{\sigma,\beta}$ with the index $\{0,1\}^{m+1}$ split by `Fin.init` and `Fin.last`; `Witness` is a cube-trace clique of $G^{\neq}_m(H)$ together with (RC) (its field `rc`).

```
theorem isSubcube_iInter {k : ℕ} {A : Fin k → Set (Fin m → Bool)} (hA : ∀ j, IsSubcube (A j))
    {ρ : Fin m → Bool} (hρ : ∀ j, ρ ∈ A j) : IsSubcube {σ | ∀ j, σ ∈ A j}

def pad (β : Bool) (h : X → Bool) : Option X → Bool := fun x => x.elim β h

def Hpad (H : Set (X → Bool)) : Set (Option X → Bool) := {h' | ∃ h ∈ H, ∃ β, h' = pad β h}

structure Witness (X : Type) (H : Set (X → Bool)) (m : ℕ) where
  S : (Fin m → Bool) → Seq X m
  inj : Function.Injective S
  dup : ∀ σ, DupFree (S σ)
  real : ∀ σ, Realizable H (S σ)
  clique : ∀ σ τ, σ ≠ τ → Adj (S σ) (S τ)
  cube : ∀ T : Seq X m, Realizable H T → DupFree T → IsSubcube {σ | ¬ Adj (S σ) T}
  rc : ∀ h ∈ H, ∃ σ, ∀ i, h (S σ i).1 = (S σ i).2

theorem lift_cube (hm : 0 < m) (W : Witness X H m) (T' : Seq (Option X) (m + 1))
    (hT : Realizable (Hpad H) T') (hd : DupFree T') :
    IsSubcube {σ' | ¬ Adj (liftS W.S σ') T'}

def lift_witness (hm : 0 < m) (W : Witness X H m) : Witness (Option X) (Hpad H) (m + 1)

theorem lift_vcge {d : ℕ} (h : VCge H d) : VCge (Hpad H) (d + 1)

theorem lift_not_vcge (h : ¬ VCge H m) : ¬ VCge (Hpad H) (m + 1)

theorem Hpad_card {H : Set (X → Bool)} : Nat.card (Hpad H) = 2 * Nat.card H

theorem counter_step (k : ℕ) : ∃ (X : Type) (H : Set (X → Bool)),
    Nat.card X = k + 5 ∧ Nat.card H = 2 ^ (k + 4) ∧ Nonempty (Witness X H (k + 4)) ∧
      VCge H (k + 3) ∧ ¬ VCge H (k + 4)

theorem counter_all (m : ℕ) (hm : 4 ≤ m) :
    ∃ (X : Type) (H : Set (X → Bool)),
      Nat.card X = m + 1 ∧ Nat.card H = 2 ^ m ∧ Nonempty (CubeTraceDF H m) ∧
        VCge H (m - 1) ∧ ¬ VCge H m
```

The twin star (module `TwinStar`; Proposition 12.1). The domain is `Fin 3` with p, q, r the coordinates $0, 1, 2$; `St` is $\sigma \mapsto S_\sigma$ and `Ht` is H_{tw}.

```
def St (σ : Fin 2 → Bool) : Seq (Fin 3) 2 := ![(0, σ 0), (if σ 0 then 2 else 1, σ 1)]

def Ht : Set (Fin 3 → Bool) := {h | ∃ σ : Fin 2 → Bool, ![σ 0, σ 1, σ 1] = h}

theorem St_cube : ∀ T : Seq (Fin 3) 2, Realizable Ht T → DupFree T →
    IsSubcube {σ | ¬ Adj (St σ) T} := by decide +kernel

theorem twin_cubeTraceDF : Nonempty (CubeTraceDF Ht 2)

theorem twin_supports_differ : ∃ σ τ : Fin 2 → Bool, ∃ i, ∀ j, (St σ i).1 ≠ (St τ j).1

theorem twin_vc : VCge Ht 2 ∧ ¬ VCge Ht 3 := by decide
```

The threshold and the repeated vertex (module `Threshold`; Propositions 11.2, 9.2 and 9.3). `Three.H3` is H_3 on `Fin 4` $(y_1, y_2, a, b$ the coordinates $0, 1, 2, 3)$, built as in `Counter4` with `col σ = if wt σ ≤ 1 then 2 else 3` and `other σ = if wt σ = 0 then false else σ 2`; `Trep` is T_a.

```
theorem three_cubeTraceDF : Nonempty (CubeTraceDF Three.H3 3)

theorem three_vc : VCge Three.H3 3 ∧ ¬ VCge Three.H3 4 := by decide

theorem H4_vc_exact : VCge Counter4.H4 3 ∧ ¬ VCge Counter4.H4 4

def Trep : Seq (Fin 5) 4 := ![(3, false), (3, false), (3, false), (3, false)]

theorem rep_realizable : Realizable Counter4.H4 Trep := by decide

theorem rep_not_dupFree : ¬ DupFree Trep := by decide

theorem rep_trace_card :
    (Finset.univ.filter (fun σ : Fin 4 → Bool => ¬ Adj (Counter4.S σ) Trep)).card = 12

theorem rep_not_subcube : ¬ IsSubcube {σ | ¬ Adj (Counter4.S σ) Trep} := by decide

theorem S_not_cubeTraceRep (C : CubeTraceRep Counter4.H4 4) : C.S ≠ Counter4.S
```

References

[1] J. Campbell, D. Ibaibarriaga and L. Reyzin, *Contradiction graphs determine VC dimension*, arXiv:2605.20434 (v1, 19 May 2026; still the only version on 2026-09-07). Statement numbers refer to the compiled v1 e-print.

[2] N. Alon, S. Moran, H. Schefler and A. Yehudayoff, *A unified characterization of private learnability via graph theory*, arXiv:2304.03996 (v4, 12 June 2024); in: S. Agrawal and A. Roth (eds.), Proceedings of Thirty Seventh Conference on Learning Theory (COLT 2024), Proceedings of Machine Learning Research **247**, PMLR, 2024, pp. 94–129.

[3] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`, entries A005271 (perfect matchings of the n-cube), A002414 (octagonal pyramidal numbers) and A018926 (partitions of the n-cube into subcubes); accessed 2026-09-03 and 2026-09-07.

[4] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: A. Platzer and G. Sutcliffe (eds.), Automated Deduction — CADE 28, Lecture Notes in Computer Science **12699**, Springer, Cham, 2021, pp. 625–635; `doi:10.1007/978-3-030-79876-5_37`.

[5] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381, `doi:10.1145/3372885.3373824`. The version used here is the `v4.32.0` tag.