NON-CLASHING TEACHING ON SMALL DOMAINS: A COMPLETE FIVE-POINT CENSUS, A PLANAR CLOSED-NEIGHBOURHOOD CENSUS, AND THE CLASS WHERE A WITHDRAWN PROOF FAILS

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. For a finite concept class $\mathcal{C}$ over a finite domain, Kirkpatrick, Simon and Zilles and Fallat, Kirkpatrick, Simon, Soltani and Zilles ask whether $\mathrm{NCTD}(\mathcal{C}) > \mathrm{VCD}(\mathcal{C})$ is possible; the only search they report is over "those classes for which PBTD > VCD is known from the literature". We report an exhaustive computation over all 4 294 967 295 concept classes on a five-point domain — the joint distributions of (VCD, NCTD) and of (VCD, NCTD^+) — which finds NCTD $\leq$ VCD throughout, so on domains of at most five points the question is settled, by enumeration. We then show that Warmuth's class — the class both of the sources above print when they pose the question — has VC dimension 2 and no two-element fragment of frequency 1, so it is a counterexample to Lemma 2 of a preprint that claimed NCTD $\leq$ VCD in general and was withdrawn by its authors with the comment "The proof of Lemma 2 is wrong"; the remaining step of that argument is correct, so the whole claim rested on that lemma. For the closed neighbourhoods of a planar graph Bhore, Khazaliya and Mc Inerney prove NCTD ≤ 5 and ask whether the bound is tight: we exhibit the smallest graph on which the value 2 is attained and report an exhaustive census showing that 2 is the true maximum up to ten vertices. Around these we give a formal development, in Lean 4 against *Mathlib*, of the small-domain values $\mathrm{NCTD}(2^X) = \lceil n/2 \rceil$ for $n \leq 5$ with explicit witnesses, and of the hypercube case of the degree lower bound; both are theorems of Kirkpatrick, Simon and Zilles for every n, and are recorded here as verified, not as new. The census computations are external to the formal development and are labelled as such wherever they are used.

1. Introduction

1.1. **Two dimensions of a concept class.** Let X be a finite set, the *domain.* A *concept* is a subset $C \subseteq X$ and a *concept class* is a nonempty finite family $\mathcal{C} \subseteq 2^X$. Two combinatorial parameters of $\mathcal{C}$ meet in this paper.

The first is the Vapnik–Chervonenkis dimension. A set $Y \subseteq X$ is *shattered* by $\mathcal{C}$ when every $p \subseteq Y$ is $C \cap Y$ for some $C \in \mathcal{C}$, and $\mathrm{VCD}(\mathcal{C})$ is the largest size of a shattered set.

The second comes from machine teaching. A teacher hands the learner a set of labelled examples; since the labels of a sample are forced by the concept it is drawn from, a *teaching map* is just a map $T\colon \mathcal{C} \to 2^X$. Kirkpatrick, Simon and Zilles [1] call T *non-clashing* — following Kuzmin and Warmuth [7] — when no two distinct concepts are each consistent with the other's teaching set. Unfolded, that is the pairwise condition

$$\big(T(C) \cup T(C')\big) \cap \big(C \triangle C'\big) \neq \emptyset \qquad \text{for all distinct } C, C' \in \mathcal{C}, \tag{1}$$

which is the form used verbatim in [3]: some point of $T(C) \cup T(C')$ lies in one of the two concepts and not the other. With $\mathrm{ord}(T) = \max_{C\in\mathcal{C}} |T(C)|$, the *non-clashing teaching dimension* is

$$\mathrm{NCTD}(\mathcal{C}) = \min\{\mathrm{ord}(T) : T \text{ is a non-clashing teaching map for } \mathcal{C}\}. \tag{2}$$

The variant NCTD^+ requires in addition $T(C) \subseteq C$, i.e. only positive examples.

1.2. **The question.** Both papers that introduce and develop NCTD — [1] and its journal successor [2] — ask whether NCTD can exceed VCD. The question is still called open in a paper revised in March 2026, which states it as follows [3]:

> "A key open question of theirs is whether there is a concept class $\mathcal{C}$ such that $\mathrm{NCTD}(\mathcal{C}) > \mathrm{VCD}(\mathcal{C})$."

What is known in the positive direction is that $\mathrm{NCTD}(\mathcal{C}) \le \mathrm{VCD}(\mathcal{C})$ for every finite *maximum* class ([2], §4.2, by way of the representation maps of Chalopin, Chepoi, Moran and Warmuth), and a counting bound: if $\mathrm{NCTD}(\mathcal{C}) = d$ then $|\mathcal{C}| \le 2^d\binom{X(\mathcal{C})}{d}$, where $X(\mathcal{C}) \le n$ is the least number of domain points an optimal non-clashing map need use ([1], Theorem 13; [2], Theorem 19; the same statement in the cruder form $|\mathcal{C}| \le 2^d|X_d|$ opens [1] §4 and [2] §3.2.1, X_d being the d-element subsets of the domain). Simon [6] sharpens the bound $2^d\binom{n}{d}$ — which he attributes to [1] — to

$$|\mathcal{C}| \;\le\; \left(2\sqrt{\tfrac{2}{d+1}} - \tfrac{2}{d+1}\right)\cdot 2^d\binom{n}{d},$$

"a factor of order $\sqrt{d}$", by way of Johnson graphs. Everywhere below we use the bound only in the weaker shape $|\mathcal{C}| \le 2^d\binom{n}{d}$.

Also known, and central to what this paper does *not* claim, is the exact value of NCTD on the full power set. Write $P_m = 2^{[m]}$. Then

$$\mathrm{NCTD}(P_m) = \left\lceil \tfrac{m}{2} \right\rceil \quad \text{and} \quad \mathrm{NCTD}^+(P_m) = m \qquad \text{for every } m \tag{3}$$

([1], Theorem 19; [2], Theorem 29 with Remark 23), and both papers draw the consequence explicitly. In the words of [1]: "Since the NCTD of any concept class over a domain $\mathcal{X}$ is trivially upper bounded by the NCTD of the power set over $\mathcal{X}$, this result in particular implies that $\lceil|\mathcal{X}|/2\rceil$ is an upper bound on the NCTD of any concept class over a domain $\mathcal{X}$." ([2] repeats the sentence with "powerset" for "power set".) The lower half of (3) is their average-degree bound $\mathrm{NCTD}(\mathcal{C}) \ge \lceil\frac{1}{2}\deg_{\mathrm{avg}}(\mathcal{C})\rceil$ ([1], Theorem 14; [2], Theorem 21) applied to the hypercube. Section 3 below adds nothing to (3); it records a machine-checked proof of the cases $n \le 5$, with the witnesses written out.

What is known in the negative direction is one sentence. In [2], §4.2 ("VCD"), it reads:

> "So far, there is no concept class for which NCTD is known to exceed VCD. Note that any such concept class would have to fulfill PBTD > VCD as well. We tested those classes for which PBTD > VCD is known from the literature, but found that all of them satisfy NCTD ≤ VCD."

The same sentence stands in [1], §6, with the two dimensions of its first clause transposed ("no concept class for which VCD is known to exceed NCTD") — a slip the journal version repairs. Here PBTD is the *preference-based teaching dimension* of [1, 2], after Gao, Ries, Simon and Zilles. Neither paper prints a domain size, a class count or an enumeration: the search was over a handful of classes named in the literature, not over a domain. That is the gap this paper fills at the small end.

1.3. **What is settled here.** Because both dimensions are finite and elementary, small domains are decidable, and the question above has an exact answer there.

The five-point census. Section 4 reports an exhaustive computation over all 4 294 967 295 concept classes on a five-point domain — 1 228 157 orbit representatives under the hyperoctahedral group of order 3840 — with the joint distribution of (VCD, NCTD) and of $(\mathrm{VCD}, \mathrm{NCTD}^+)$. Every class satisfies NCTD ≤ VCD. This is the main content of the paper. It is external to the formal development and is labelled as such throughout, and it is what settles the question of [1, 2] on domains of at most five points. The published bound (3) does not settle it: it gives $\mathrm{NCTD} \le 2$ on four points and $\mathrm{NCTD} \le 3$ on five, so it disposes only of the classes whose VC dimension is at least $\lceil n/2\rceil$. On five points that leaves the 166 075 480 classes of VC dimension at most 2, and on three and four points the 2068 classes of VC dimension at most 1; for those the table below, and nothing else here, is the evidence.

Warmuth's class and a withdrawn proof. A March 2026 preprint [4] announced a positive answer for all finite classes by an ordered compression scheme. It was withdrawn by its authors on 3 August 2026, the comment field reading, in full, "The proof of Lemma 2 is wrong". Section 5 adds what that comment does not: the engine of the argument, its Lemma 2, is false at $\mathcal{C}_W$, *the* class both [1] and [2] print when they pose the open question (Theorem 5.2). We also isolate the failure: the step the preprint's abstract calls one that "one can easily see" is in fact correct (Lemma 5.1), so the entire claim rested on Lemma 2. A companion enumeration (Remark 5.5) locates every class on which that scheme is stuck at its first round: there is none on a domain of at most four points, and exactly 960 on five, of which Warmuth's class is the smallest.

Closed neighbourhoods of planar graphs. For G planar and $\mathcal{B} = \{N[v] : v \in V(G)\}$, [3] proves $\mathrm{NCTD}(\mathcal{B}) \leq 5$ and asks whether that is tight, noting that balls in planar graphs have been proposed [5] as a candidate class with NCTD > VCD. Section 6 exhibits a six-vertex graph — two disjoint triangles joined by two edges — with $\mathrm{NCTD}(\mathcal{B}) = \mathrm{VCD}(\mathcal{B}) = 2$ (Theorem 6.1). An exhaustive census over the 1 052 805 connected planar graphs on at most ten vertices, again external, shows that 2 is the maximum in that range, that this graph is the unique smallest one attaining it, and that there is no instance of NCTD > VCD.

A verified small-domain development. Section 3 carries a Lean 4 proof of the hypercube case of the average-degree bound of [1, 2] — $n \leq 2\,\mathrm{NCTD}(2^{[n]})$ for every n (Theorem 3.1) — and of the matching upper bounds for $n \leq 5$ from explicit witnesses, hence of (3) in the range $n \leq 5$ (Theorems 3.3, 3.4 and 3.5). None of this is new mathematics: (3) is [1] Theorem 19 for every m. What Section 3 supplies is a machine-checked proof and printed witnesses, and the statements the census section leans on.

1.4. **What is not claimed.** We do not claim the value $\mathrm{NCTD}(2^{[n]}) = \lceil n/2 \rceil$, at any n: it is [1] Theorem 19 and [2] Theorem 29, and so is the bound $\mathrm{NCTD}(\mathcal{C}) \leq \lceil |X|/2 \rceil$ for every class on X. Nor do we claim the double count of Theorem 3.1, which is [1] Theorem 14 specialised to the cube. We do not claim that NCTD ≤ VCD holds in general, nor that no class on a six-point domain is a counterexample, nor that the planar bound 5 of [3] fails to be tight for large n. We make no claim about *why* the authors of [4] withdrew their preprint beyond quoting the comment they recorded; what is proved here is that the statement of its Lemma 2 is false, at a specific class.

1.5. **Where this was looked for.** Everything reported here was searched for before being claimed, and one search came back positive, which is why Section 3 attributes rather than claims: the exact values $\mathrm{NCTD}(P_m) = \lceil m/2 \rceil$ and $\mathrm{NCTD}^+(P_m) = m$ are Theorem 19 of [1] and Theorem 29 of [2], for every m, and the double count behind them is Theorem 14 of [1]. For the censuses the record is as follows. All eleven forward citations of [2] recorded by Semantic Scholar were listed and read by title on 3 September 2026; the only one that could plausibly have contained an enumeration is Simon [6], whose contribution is the characterisation of maximum classes of non-clashing teaching dimension 1 by tournaments and the $\sqrt{d}$ improvement of the counting bound — a structural theorem, not a census. A full-text corpus of arXiv through August 2026 was searched for "non-clashing teaching" together with enumeration vocabulary, with no hit reporting a census; the OEIS returns no result for the census sequences $8, 116, 98, 32, 1$ and $1928, 14508, 33088$, and the planar sequence $0, 1, 1, 1, 2, 2, 2, 2, 2$ matches only generic short sequences. OpenAlex records no citation of either [3] or [4]. Read those negatives as "not found", not as "new"; every computation reported here is an afternoon of processor time for anyone who tries it, which is precisely why its absence from the record is worth correcting rather than celebrating.

2. Preliminaries

Throughout, X is a finite domain, $n = |X|$, and $[n] = \{1, \ldots, n\}$; we identify X with $[n]$ and write a concept $C \subseteq [n]$ as the bit string whose i-th symbol is 1 iff $i \in C$. A concept class is a nonempty $\mathcal{C} \subseteq 2^{[n]}$, and $2^{[n]}$ itself denotes the class of *all* 2^n concepts.

Definition 2.1 (shattering, VCD)**.** $\mathcal{C}$ shatters $Y \subseteq [n]$ if for every $p \subseteq Y$ there is $C \in \mathcal{C}$ with $C \cap Y = p$. $\mathrm{VCD}(\mathcal{C})$ is the largest $|Y|$ over shattered Y.

Definition 2.2 (non-clashing, NCTD)**.** A teaching map $T \colon \mathcal{C} \to 2^{[n]}$ is non-clashing if (1) holds, and $\mathrm{NCTD}(\mathcal{C})$ is given by (2). We write $\mathrm{NCTD}(\mathcal{C}) \leq k$ to mean that some non-clashing T has $|T(C)| \leq k$ for all $C \in \mathcal{C}$. Upper bounds below are proved in that form; a lower bound $\mathrm{NCTD}(\mathcal{C}) \geq k+1$ is proved by refuting it.

Two consequences of the pairwise shape of (1) are used constantly and recorded once.

Lemma 2.3 (monotonicity). *If $\mathcal{C}' \subseteq \mathcal{C}$ then* $\mathrm{NCTD}(\mathcal{C}') \le \mathrm{NCTD}(\mathcal{C})$*: the restriction of a non-clashing map to a subclass is non-clashing, and its order does not increase. In particular* $\mathrm{NCTD}(\mathcal{C}) \le \mathrm{NCTD}(2^{[n]})$ *for every class $\mathcal{C}$ on $[n]$, so* $\mathrm{NCTD}(2^{[n]})$ *is the largest value* NCTD *takes on an n-point domain, and a single witness for the power set settles every class on that domain at once.*

Lemma 2.4 (the counting bound at $d = 1$; not new). *If a concept class $\mathcal{C}$ on $[n]$ admits a non-clashing teaching map of order at most 1, then $|\mathcal{C}| \le 2n + 1$.*

Proof. Send C to the labelled single example it is taught by: to the pair $(x, \mathbf{1}[x \in C])$ if $T(C) = \{x\}$, and to a symbol $*$ if $T(C) = \emptyset$. If two distinct concepts C, C' received the same value then no point of $T(C) \cup T(C')$ separates them — either both teaching sets are empty, or they are the same singleton $\{x\}$ with x on the same side of both concepts — contradicting (1). So the map is injective and $|\mathcal{C}| \le 2n + 1$. □

Lemma 2.4 is the case $d = 1$ of the counting bound of [1], Theorem 13 (= [2], Theorem 19), and of [6]. Those sources state it for a class *of* non-clashing teaching dimension d and normalise the witnessing map so that every teaching set has exactly d elements ([1], Proposition 10; [2], Proposition 6), which at $d = 1$ makes their bound $2X(\mathcal{C}) \le 2n$; the extra $+1$ here is the price of stating the hypothesis as "order at most 1", which admits the at most one concept taught by the empty set.

3. The largest value of NCTD on a small domain, verified

Everything in this section is a case of (3), i.e. of [1] Theorem 19 and [2] Theorem 29, and none of it is new. It is set out in full because the census of Section 4 is read against it, because the witnesses are worth printing, and because these are the statements the formal development of Section 8 carries; a reader who wants only the new material may go straight to Section 4.

3.1. A lower bound for the cube, for every n.

The full power set is the unique largest class on its domain, and its teaching maps are constrained by the edges of the hypercube: two concepts differing in exactly one point can be separated only at that point, so that point must be taught to one of them. Counting incidences gives the following, which is the average-degree bound $\mathrm{NCTD}(\mathcal{C}) \ge \lceil \frac{1}{2} \deg_{\mathrm{avg}}(\mathcal{C}) \rceil$ of [1] Theorem 14 (= [2] Theorem 21) specialised to $\mathcal{C} = 2^{[n]}$, where every concept has degree n; the proof below is theirs, written out for the cube.

Theorem 3.1 (the hypercube edge bound; [1] Theorem 14 at $2^{[n]}$). *Let $n \in \mathbb{N}$ and let T be a non-clashing teaching map for $2^{[n]}$ with $|T(C)| \le k$ for every $C \subseteq [n]$. Then $n \le 2k$. Equivalently,* $\mathrm{NCTD}(2^{[n]}) \ge \lceil n/2 \rceil$.

Proof. For $C \subseteq [n]$ and $x \in [n]$ write $C^x = C \triangle \{x\}$. The concepts C and C^x are distinct and $C \triangle C^x = \{x\}$, so (1) forces $x \in T(C) \cup T(C^x)$, that is

$$1 \;\le\; \mathbf{1}[x \in T(C)] + \mathbf{1}[x \in T(C^x)] \qquad \text{for all } C \subseteq [n],\ x \in [n].$$

Sum over all $2^n n$ pairs (C, x). The first term sums to $\sum_C |T(C)|$. For fixed x the involution $C \mapsto C^x$ is a bijection of $2^{[n]}$, so the second term sums to the same quantity. Hence

$$2^n n \;\le\; 2 \sum_{C \subseteq [n]} |T(C)| \;\le\; 2 \cdot 2^n k,$$

and dividing by $2^n > 0$ gives $n \le 2k$. □

Nothing in the proof rests on a computation. Note also that Theorem 3.1 is a statement about the *full* power set and does not descend to subclasses: on the same five-point domain the class $\{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$ has $\mathrm{NCTD} \le 1$, taught by $T(C) = C$, well below $\lceil 5/2 \rceil$.

Remark 3.2 (comparison with the counting bound). For the power set the counting bound $|\mathcal{C}| \le 2^d \binom{n}{d}$ of [2, 6] reads $2^n \le 2^d \binom{n}{d}$ and so gives its own lower bound on $d = \mathrm{NCTD}(2^{[n]})$. Comparing the two for $n \le 30$: the two bounds agree for $n \le 4$ and at $n = 6$, and the edge bound of Theorem 3.1 is strictly stronger at $n = 5$ and at every n with $7 \le n \le 30$. The decisive case for this paper is $n = 5$: at $d = 2$

the counting bound reads $32 \le 40$ and is satisfied, so it permits $\mathrm{NCTD}(2^{[5]}) = 2$, whereas the edge count reads $5 \le 4$ and rules it out. Asymptotically the counting bound gives only $d \gtrsim 0.227\,n$ — the root of $\alpha + H(\alpha) = 1$ — against $n/2$ here. (Both are lower bounds on d; neither is tight at $n = 6$, where each forces only $d \ge 3$, and (3) gives $d = 3$.)

3.2. **Matching upper bounds on three, four and five points.** The upper bounds of (3) are proved in [1, 2] by sub-additivity: for even m the power set P_m is the free combination of $m/2$ copies of P_2, and $\mathrm{NCTD}(P_2) = 1$; for odd m one passes through P_{m+1}. That argument produces a witness too, and one could unfold it. We use instead witnesses found directly, by an exhaustive search over teaching maps of the given order implemented in C, because a single flat table is what the kernel can check without any of the sub-additivity machinery: each is a finite object whose non-clashing property is a check of all $\binom{2^n}{2}$ pairs, and it is that check — not the search — that carries the proof. We print the witnesses so that the reader can repeat the check. Concepts and teaching sets are written as bit strings of length n, position i standing for the point i.

U_4, an order-2 non-clashing map for $2^{[4]}$:

0000 → 1100	1000 → 0101	0100 → 1010	1100 → 0011
0010 → 1010	1010 → 0011	0110 → 1100	1110 → 0101
0001 → 0101	1001 → 1100	0101 → 0011	1101 → 1010
0011 → 0011	1011 → 1010	0111 → 0101	1111 → 1100

U_3, an order-2 non-clashing map for $2^{[3]}$:

000 → 110	100 → 110	010 → 101	110 → 101
001 → 101	101 → 101	011 → 110	111 → 110

U_5, an order-3 non-clashing map for $2^{[5]}$:

00000 → 11100	10000 → 11100	01000 → 10110	11000 → 10011
00100 → 11010	10100 → 11010	01100 → 11001	11100 → 10101
00010 → 11010	10010 → 10110	01010 → 10101	11010 → 11100
00110 → 11100	10110 → 11001	01110 → 10011	11110 → 10011
00001 → 11001	10001 → 10101	01001 → 10011	11001 → 11010
00101 → 10101	10101 → 10011	01101 → 11100	11101 → 11100
00011 → 10011	10011 → 10011	01011 → 11100	11011 → 11001
00111 → 00111	10111 → 11100	01111 → 11010	11111 → 10110

Theorem 3.3 ([1] Theorem 19 at $m = 3, 4$)**.** *Every concept class on a domain of three points, and every concept class on a domain of four points, has* $\mathrm{NCTD} \le 2$*. Consequently, if a concept class $\mathcal{C}$ on a four-point domain shatters a set of size $d \ge 2$ then* $\mathrm{NCTD}(\mathcal{C}) \le d$*; in particular* $\mathrm{NCTD}(\mathcal{C}) \le \mathrm{VCD}(\mathcal{C})$ *for every concept class on a three- or four-point domain whose VC dimension is at least* 2*.*

Proof. The tables U_3 and U_4 are non-clashing of order 2 for $2^{[3]}$ and $2^{[4]}$: this is the verification of $\binom{8}{2} = 28$ and $\binom{16}{2} = 120$ pair conditions from (1), an exhaustive finite check. Lemma 2.3 transfers the bound to every subclass. The second sentence is Lemma 2.3 together with monotonicity of the bound in d, and the third is the second applied to a shattered set of maximum size. □

For a domain of at most two points the bound $\mathrm{NCTD} \le 2$ is immediate, since $T(C) = X$ has order $|X| \le 2$ and separates every pair. The value 2 in Theorem 3.3 is best possible on both domains, and the gap to VCD is already visible at $n = 4$.

Theorem 3.4 ([1] Theorem 19 at $m = 4$)**.** $\mathrm{NCTD}(2^{[4]}) = 2$ *while* $\mathrm{VCD}(2^{[4]}) = 4$.

Proof. Upper bound: the witness U_4. Lower bound: $|2^{[4]}| = 16 > 9 = 2 \cdot 4 + 1$, so Lemma 2.4 forbids order 1; alternatively Theorem 3.1 gives $4 \le 2k$. That $\mathrm{VCD}(2^{[4]}) = 4$ is the definition, the whole domain being shattered. □

Theorem 3.5 ([1] Theorem 19 at $m = 5$)**.** $\mathrm{NCTD}(2^{[5]}) = 3$. *Consequently every concept class on a domain of at most five points has* $\mathrm{NCTD} \le 3$.

Proof. Upper bound: the witness U_5 is non-clashing of order 3 for $2^{[5]}$, a check of the $\binom{32}{2} = 496$ pair conditions. Lower bound: Theorem 3.1 with $n = 5$ gives $5 \le 2k$, so $k \ge 3$; equivalently, no order-2 map for $2^{[5]}$ exists, and we note that U_5 indeed has a concept with a three-element teaching set, so its order is exactly 3. The consequence is Lemma 2.3. □

Corollary 3.6. *For* $3 \le n \le 5$, $\mathrm{NCTD}(2^{[n]}) = \lceil n/2 \rceil$*, and this is the largest value* NCTD *takes on an* n*-point domain. Since* $\mathrm{VCD}(2^{[n]}) = n$*, the gap* $\mathrm{VCD} - \mathrm{NCTD}$ *on the full power set is* $\lfloor n/2 \rfloor$ *there.*

This is (3) in the range $n \le 5$; [1] Theorem 19 gives it for every n, and gives $\mathrm{NCTD}(2^{[1]}) = \mathrm{NCTD}(2^{[2]}) = 1$ at the bottom, which is outside the formal development here. What Theorems 3.1, 3.3, 3.4 and 3.5 add is that in this range the value is kernel-checked, from printed witnesses, against a transcription of the definitions of [1, 3]; Section 8 says what that means.

4. The five-point census

This section reports computations that are not part of the formal development. They are labelled as such, they support no theorem above, and the theorems above do not depend on them.

An exhaustive scan over all $2^{32} - 1 = 4\,294\,967\,295$ nonempty classes on a five-point domain was run in C, keeping the lexicographically minimal representative of each orbit under the hyperoctahedral group $\mathbb{Z}_2^5 \rtimes S_5$ of order 3840 acting on the 32 concepts and weighting each representative by its orbit size. There are 1 228 157 orbit representatives, and their weights sum to exactly $2^{32} - 1$, which is the arithmetic check on the scan. On each representative an exact solver computed VCD, NCTD and NCTD^+. The joint distribution of $(\mathrm{VCD}, \mathrm{NCTD})$, as weighted counts over all classes, is

	NCTD 0	1	2	3
VCD 0	32			
1		44 880		
2		5 914 552	160 116 016	
3		95 880	3 958 066 436	643 424
4			162 997 070	7 089 004
5				1

and that of $(\mathrm{VCD}, \mathrm{NCTD}^+)$ is

	NCTD^+ 0	1	2	3	4	5
VCD 0	32					
1		44 880				
2		186 640	130 944 312	34 420 256	479 360	
3			294 079 564	3 192 590 192	472 135 984	
4				51 324 168	118 761 906	
5						1

Three readings. First, *every* class on a five-point domain satisfies $\mathrm{NCTD} \le \mathrm{VCD}$, and therefore so does every class on a domain of at most five points; on that range the question of [1, 2] is settled, by enumeration. This is the one statement here that the known bound (3) does not give. That bound yields $\mathrm{NCTD} \le \lceil n/2 \rceil$ and so covers only the classes with $\mathrm{VCD} \ge \lceil n/2 \rceil$; on five points the 166 075 480 classes of VC dimension at most 2, and on three and four points the $124 + 1944 = 2068$ classes of VC dimension at most 1, are covered by these tables and by nothing else in this paper. Second, the largest value of NCTD is 3 on five points, 2 on four and on three, matching Corollary 3.6 and hence (3), which is one check on the solver. Third, the positive variant behaves differently, as expected: $\mathrm{NCTD}^+ > \mathrm{VCD}$ occurs in bulk, and the largest gap $\mathrm{NCTD}^+ - \mathrm{VCD}$ already on a four-point domain is 3, attained by the single class $\{S \subseteq [4] : |S| \geq 3\}$ with VCD = 1 and $\mathrm{NCTD}^+ = 4$; the analogue on three points is $\{S \subseteq [3] : |S| \geq 2\}$ with gap 2. The value $\mathrm{NCTD}^+ = 5$ at VCD = 5 in the second table is $\mathrm{NCTD}^+(P_5) = 5$, again (3).

The domain-3 and domain-4 tables underlying the same computation were reproduced cell for cell by an independent implementation; the cell $(\mathrm{VCD}, \mathrm{NCTD}) = (3, 1)$ on four points has count 2, and its two classes are exactly the even- and the odd-weight halves of the 4-cube. The published values that the solver can be tested against were reproduced exactly: $\mathrm{NCTD}(\mathcal{C}_W) = \mathrm{NCTD}^+(\mathcal{C}_W) = 2$ with $\mathrm{VCD}(\mathcal{C}_W) = 2$ from [1, 2], the bound $\mathrm{NCTD} \leq 2 < 3 = \mathrm{VCD}$ for the class of all balls of a cycle of length $n \geq 6$ from [5] — the class $\mathcal{B}^*$ of Section 6 — checked, with equality, for $n = 5, \ldots, 9$, and $\mathrm{NCTD}(P_n) = \lceil n/2 \rceil$, $\mathrm{NCTD}^+(P_n) = n$ for $n \leq 5$ from (3). Total cost of the five-point scan: about 0.14 processor-hours.

5. Warmuth's class, and Lemma 2 of a withdrawn preprint

5.1. **The class.** Warmuth's class $\mathcal{C}_W$ is a class of ten concepts on a five-point domain, communicated by Manfred Warmuth and, on the authority of [1, 2], shown by Darnstädt et al. [8] to be the smallest concept class with PBTD > VCD; it is printed as Table 1 of [1] and Table 2 of [2], in both cases in the discussion of the question of Section 1. Its ten concepts are

$$\begin{array}{lllll} C_1 = 10001, & C_2 = 11000, & C_3 = 01100, & C_4 = 00110, & C_5 = 00011, \\ C_1' = 10101, & C_2' = 11010, & C_3' = 01101, & C_4' = 10110, & C_5' = 01011, \end{array}$$

with $\mathrm{VCD}(\mathcal{C}_W) = 2$ and $\mathrm{PBTD}(\mathcal{C}_W) = 3$. Both sources introduce it in the same words: "This concept class ... was communicated by Manfred Warmuth and proven by Darnstädt et al. (2016) to be the smallest concept class for which PBTD exceeds VCD. In particular, $\mathrm{VCD}(\mathcal{C}_W) = 2$ while $\mathrm{PBTD}(\mathcal{C}_W) = 3$." The minimality is used here only as context; the 2016 paper [8] was not read for this work, and the attribution above is [1, 2]'s.

5.2. **The claim that was withdrawn.** The preprint [4] announced a proof that $\mathrm{NCTD}(\mathcal{C}) \leq \mathrm{VCD}(\mathcal{C})$ for every finite class. Its abstract reads, in part:

> "... for any finite concept class, we construct fragments of size equals to its Vapnik-Chervonenkis dimension which identify concepts through an ordered compression scheme. Naturally, these fragments are used as teaching sets, one can easily see that they satisfy the non-clashing condition, i.e., this open question is resolved for finite concept classes."

Write $d = \mathrm{VCD}(\mathcal{C})$. In the terminology of [4] a *fragment* is a pattern $f_Y \in \mathcal{C}_Y = \{C \cap Y : C \in \mathcal{C}\}$ for a d-element $Y \subseteq X$, and its *frequency* is the number of concepts of $\mathcal{C}$ restricting to it. The scheme's first step is, verbatim: "Consider each f_Y whose frequency equals to 1, assigning f_Y to the unique function as its compression set"; the assigned concept is then removed from the pool and the count repeated. The engine is its Lemma 2:

> "There exists at least one fragment with frequency 1 at each time of counting."

justified by "For the first time of compression, there exists at least one f_X whose frequency equals to 1, otherwise, $\mathrm{VCdim}(\mathcal{C}) \geq d + 1$ due to the definition of VC dimension" — where the source's X is the d-element set we write Y. The preprint was withdrawn by its authors on 3 August 2026 with the comment "The proof of Lemma 2 is wrong". The numbering "Lemma 2" and "Theorem 3" used here, and every quotation in this section, are from the v3 e-print of [4], the last version distributed with source; the withdrawn v4 carries no source, so no numbering there can be checked, and the authors' own comment names Lemma 2 as well.

5.3. **The rest of the argument is sound.** Before exhibiting the counterexample it is worth recording that nothing else in that argument is at fault, because it localises what has to be repaired. The step the abstract calls one that "one can easily see" — that a teaching map produced by an ordered removal is non-clashing — is correct, and correct for a reason worth stating abstractly: the non-clashing condition (1) is symmetric in the two concepts and is satisfied as soon as *one* of the two teaching sets separates the pair.

Lemma 5.1 (ordered removal implies non-clashing; the content of [4], Theorem 3, given its Lemma 2)**.** *Let $\mathcal{C}$ be a concept class, $T\colon \mathcal{C} \to 2^X$ a teaching map and $r\colon \mathcal{C} \to \mathbb{N}$ a rank function such that for all distinct $C, C' \in \mathcal{C}$ with $r(C) \le r(C')$ there is $x \in T(C)$ with $x \in C \triangle C'$. Then T is non-clashing.*

Proof. Let $C \ne C'$. If $r(C) \le r(C')$ the hypothesis gives $x \in T(C) \subseteq T(C) \cup T(C')$ with $x \in C \triangle C'$; otherwise $r(C') \le r(C)$ and the hypothesis applied to the pair in the other order gives $x \in T(C')$ with $x \in C' \triangle C = C \triangle C'$. Either way (1) holds. □

A concept assigned in some round of the scheme of [4] has a teaching set whose fragment had frequency 1 in the pool at that time, so it is separated from every concept still in the pool, i.e. from every concept of equal or later rank; Lemma 5.1 then applies with r the round number. So the claim of [4] rests entirely on its Lemma 2: on the scheme being able to make an assignment at all.

5.4. **Where Lemma 2 fails.** It fails at $\mathcal{C}_W$, and it fails at the very first round.

Theorem 5.2. $\mathrm{VCD}(\mathcal{C}_W) = 2$*, and for every two-element $Y \subseteq X$ and every $p \subseteq Y$ the number of concepts $C \in \mathcal{C}_W$ with $C \cap Y = p$ is never* 1*. Hence $\mathcal{C}_W$ has no fragment of size* $\mathrm{VCD}(\mathcal{C}_W)$ *and frequency* 1*, and the ordered compression scheme of* [4] *cannot make its first assignment on $\mathcal{C}_W$; in particular the statement of Lemma 2 of* [4] *is false.*

Proof. Both halves are finite checks; both have short reasons as well.

For the VC dimension, the pair $\{1,2\}$ is shattered: $C_4 \cap \{1,2\} = \emptyset$, $C_1 \cap \{1,2\} = \{1\}$, $C_3 \cap \{1,2\} = \{2\}$, $C_2 \cap \{1,2\} = \{1,2\}$. No triple is shattered. Indeed the concepts of $\mathcal{C}_W$ are five of size 2, namely $\{1,5\}$, $\{1,2\}$, $\{2,3\}$, $\{3,4\}$, $\{4,5\}$ — the edges of a 5-cycle — and five of size 3, namely $\{1,3,5\}$, $\{1,2,4\}$, $\{2,3,5\}$, $\{1,3,4\}$, $\{2,4,5\}$. If Y with $|Y| = 3$ were shattered, some concept would meet Y in $\emptyset$, hence be contained in the two-point set $X \setminus Y$ and so equal it, and some concept would contain Y, hence equal Y. So Y would be a three-element concept whose complement is a two-element concept; but the complements of the five two-element concepts are $\{2,3,4\}, \{3,4,5\}, \{1,4,5\}, \{1,2,5\}, \{1,2,3\}$, and none of those is a concept.

For the frequencies, each of the $\binom{5}{2} = 10$ pairs $Y = \{i,j\}$ splits $\mathcal{C}_W$ into the four blocks $\{C : C \cap Y = \emptyset\}$, $\{C : C \cap Y = \{i\}\}$, $\{C : C \cap Y = \{j\}\}$, $\{C : C \cap Y = Y\}$ of sizes exactly $2, 3, 3, 2$ — the same profile for all ten pairs — so no block has size 1. Both statements are verified exhaustively over the finite index sets named, in the formal development of Section 8. □

Remark 5.3*.* The failure is not delicate. Every pair splits $\mathcal{C}_W$ into blocks $2, 3, 3, 2$; there is no fragment of frequency 1 to be found by any tie-breaking, any ordering of the domain, or any reformulation of the scheme that still starts from a frequency-1 fragment of size VCD.

Warmuth's class is not exotic and it is not far from the question: it is the class both sources of the open question print in the paragraph where they pose it, and, by [1, 2], the smallest class with PBTD $>$ VCD — the property their own quoted search sentence uses to narrow the field. It also satisfies the inequality it is used here to test:

Proposition 5.4 (the upper half of [2] Proposition 39 and [1] Proposition 27; not new)**.** $\mathrm{NCTD}(\mathcal{C}_W) \le 2 = \mathrm{VCD}(\mathcal{C}_W)$*, witnessed by the order-*2 *map*

$$\begin{aligned}
&C_1 = 10001 \to 01100, \quad C_2 = 11000 \to 11000, \quad C_3 = 01100 \to 01100, \quad C_4 = 00110 \to 00110,\\
&C_5 = 00011 \to 11000, \quad C_1' = 10101 \to 01010, \quad C_2' = 11010 \to 10010, \quad C_3' = 01101 \to 01001,\\
&C_4' = 10110 \to 10100, \quad C_5' = 01011 \to 10100.
\end{aligned}$$

Proof. A check of the $\binom{10}{2} = 45$ pair conditions of (1), plus the observation that each teaching set has two elements. The matching lower bound $\mathrm{NCTD}(\mathcal{C}_W) \ge 2$ is from [1, 2] and is not reproved here. □

Remark 5.5 (how rare fragment-freeness is; computation)*.* An exhaustive scan over all classes on domains of size at most five, external to the formal development, locates every class on which the scheme of [4] is stuck at round 0 — call such a class *fragment-free*: it has no fragment of size VCD and frequency 1. On a domain of three or four points there is no fragment-free class at all, among the $255 + 65\,535 = 65\,790$ of them; so there the scheme runs to completion on every class and, with

Lemma 5.1, yields a second and independent proof that NCTD $\leq$ VCD there. On a five-point domain there are exactly 960 fragment-free classes, in four orbits under the group of order 3840, of sizes $192, 192, 384, 192$; their representatives have $(|\mathcal{C}|, \mathrm{VCD}, \mathrm{NCTD}) = (10,2,2)$, $(20,3,2)$, $(21,3,2)$ and $(22,3,3)$. The smallest of them, and the only one of VC dimension 2, is Warmuth's class: the orbit of $\mathcal{C}_W$ under the group has 192 members and is the first of the four. The class exhibited in Theorem 5.2 is thus not one counterexample among many but, in the smallest domain where any exists, the smallest one.

6. Closed neighbourhoods of planar graphs

For a graph G let $N[v]$ be the closed neighbourhood of v and $\mathcal{B}(G) = \{N[v] : v \in V(G)\}$, a concept class on $V(G)$. For planar G, [3] proves $\mathrm{NCTD}(\mathcal{B}(G)) \leq 5$ (its Theorem 13, in v2) — against $\mathrm{VCD}(\mathcal{B}(G)) \leq 4$, which it attributes to [9] — and closes with

> "It is also important to determine whether Theorem 13 is tight, since the balls of a planar graph has been mentioned [CCMR24] as a potential concept class $\mathcal{C}$ negatively answering the question of [KSZ19, FKS+23], i.e., such that $\mathrm{NCTD}(\mathcal{C}) > \mathrm{VCD}(\mathcal{C})$."

(Its [CCMR24] is [5] and its [KSZ19, FKS+23] are [1, 2]. Statement numbers of [3] cited here are those of its v2; "Theorem 13" is verbatim in the sentence above.)

Computing $\mathrm{NCTD}(\mathcal{B}(G))$ for the small planar graphs answers the first half of that question in the small range, and the answer is that the bound 5 is far from attained: the value 2 is first reached at six vertices, and there at a single graph.

Theorem 6.1. *Let G be the graph on $\{1, \dots, 6\}$ consisting of the two disjoint triangles $\{1,3,5\}$ and $\{2,4,6\}$ together with the two edges $\{1,6\}$ and $\{2,5\}$, so that*

$$\begin{aligned} &N[1] = \{1,3,5,6\}, \quad N[2] = \{2,4,5,6\}, \quad N[3] = \{1,3,5\}, \\ &N[4] = \{2,4,6\}, \qquad N[5] = \{1,2,3,5\}, \quad N[6] = \{1,2,4,6\}. \end{aligned}$$

Then $\mathcal{B}(G)$ has six distinct members and $\mathrm{NCTD}(\mathcal{B}(G)) = \mathrm{VCD}(\mathcal{B}(G)) = 2$.

Proof. G has eight edges, so it is planar: a subdivision of K_5 has at least ten edges and a subdivision of $K_{3,3}$ at least nine, so by Kuratowski's theorem every graph with at most eight edges is planar. (Drawing the two triangles side by side and the two joining edges outside them is just as quick.)

The six closed neighbourhoods listed are pairwise distinct, so $|\mathcal{B}(G)| = 6$. The pair $\{1,5\}$ is shattered: $N[4] \cap \{1,5\} = \emptyset$, $N[6] \cap \{1,5\} = \{1\}$, $N[2] \cap \{1,5\} = \{5\}$ and $N[1] \cap \{1,5\} = \{1,5\}$. No triple is shattered, for the simplest of reasons: shattering a three-element set needs eight concepts with distinct traces, and $|\mathcal{B}(G)| = 6$. So $\mathrm{VCD}(\mathcal{B}(G)) = 2$.

For NCTD ≤ 2, the map

$$N[3] \mapsto \{1,2\}, \quad N[5] \mapsto \{1,2\}, \quad N[4] \mapsto \{1,2\}, \quad N[6] \mapsto \{1,3\}, \quad N[1] \mapsto \{1,6\}, \quad N[2] \mapsto \{1,5\}$$

is non-clashing — a check of the $\binom{6}{2} = 15$ pair conditions — and has order 2.

For NCTD ≥ 2, an order-1 map assigns to each of the six concepts either one point of the domain or nothing, so there are exactly $7^6 = 117\,649$ candidates; the assertion is that none of them is non-clashing, and it is proved by evaluating the non-clashing condition on all 117 649 of them. This exhaustion is the one claim in this paper obtained by compiled evaluation rather than by kernel reduction; see Section 8. ☐

Remark 6.2 (the planar census; computation). All connected planar graphs on $n \leq 10$ vertices were generated with `nauty` (`geng -c` piped through `planarg`) and each was evaluated with the same exact solver. The counts $1, 2, 6, 20, 99, 646, 5974, 71\,885, 1\,052\,805$ for $n = 2, \dots, 10$ agree with OEIS A003094, which is the arithmetic check. The results:

n	2	3	4	5	6	7	8	9	10
$\max \mathrm{VCD}(\mathcal{B}(G))$	0	1	2	2	2	2	3	3	3
$\max \mathrm{NCTD}(\mathcal{B}(G))$	0	1	1	1	2	2	2	2	2
$\#\{\mathrm{NCTD} > \mathrm{VCD}\}$	0	0	0	0	0	0	0	0	0

So over the whole range in which the bound 5 of [3] could have been tight, the true maximum is 2; and the value 2 is attained at $n = 6$ by exactly one of the 99 connected planar graphs, the graph of Theorem 6.1. The same computation, run for the larger class $\mathcal{B}^*(G)$ of *all* balls of all radii — the class of [5], in whose terms that paper's own closing question is posed — gives, for connected planar graphs on $n \le 8$ vertices, $\max \mathrm{VCD} = 1, 2, 3, 3, 3, 3, 4$ against $\max \mathrm{NCTD} = 1, 2, 2, 2, 2, 2, 2$ for $n = 2, \ldots, 8$, again with no instance of NCTD $>$ VCD. These censuses are external to the formal development.

7. What remains

Six points. A full census of the 2^{64} classes on six points is out of reach — some $4 \cdot 10^{14}$ orbit representatives under a group of order 46 080 — so the question of [1, 2] will not be settled at $n = 6$ the way it is settled here at $n = 5$. The value $\mathrm{NCTD}(2^{[6]}) = 3$ is known, by (3), and a direct search confirms the lower half (no order-2 map for $2^{[6]}$ exists, 1 410 154 search nodes); a search for an explicit order-3 map by the same method was inconclusive after $2 \cdot 10^8$ nodes, so no printable witness for $n = 6$ is offered here, and the formal development stops at $n = 5$. The natural instrument for that witness is a satisfiability encoding, with 64 concepts and the $\binom{6}{3} = 20$ candidate teaching sets of size 3; the sub-additive construction of [1] unfolded at $m = 6$ would also produce one.

VC dimension one. On every domain of size at most five the censuses show $\mathrm{NCTD}(\mathcal{C}) \le 1$ whenever $\mathrm{VCD}(\mathcal{C}) \le 1$. A proof of that for all n would close the only gap between the known bound (3) and the full statement NCTD $\le$ VCD on domains of at most four points; the natural route is the unique-fragment property at $d = 1$ together with Lemma 5.1.

Planar graphs. Extending the census of Remark 6.2 to $n = 11$ costs about 4.6 processor-hours by generate-and-filter and would be cheap with a direct planar generator. Nothing here bears on whether the bound 5 of [3] is tight for large n, which is the question actually asked; what the census settles is that no small graph attains it. The all-balls census stops at $n = 8$ for want of a better solver, not for want of processor time.

The question itself. On domains of at most five points the answer is that NCTD $\le$ VCD, without exception. That is a small domain, and the known bounds leave room for a counterexample from $n = 6$ on: (3) gives only $\mathrm{NCTD} \le \lceil n/2 \rceil$, which says nothing about a class of VC dimension below $\lceil n/2 \rceil$, and the counting bound of Remark 3.2 is weaker still; but the classes on which the ordered compression scheme of [4] stalls — the fragment-free classes of Remark 5.5 — are now enumerated on five points and all of them satisfy NCTD $\le$ VCD as well, so a counterexample, if there is one, will not be found among the smallest obstructions to that particular argument.

8. Verification

Every theorem, proposition and lemma stated above — Lemmas 2.3, 2.4 and 5.1, Theorems 3.1, 3.3, 3.4, 3.5, 5.2 and 6.1, and Proposition 5.4 — has been formally verified in Lean 4 against *Mathlib* [11, 12], the definitions being transcriptions of the ones quoted in Section 1. Four of these (Theorems 3.1, 3.3, 3.4 and 3.5) restate known results of [1, 2] and are formalised, not discovered, here; the new statements carried by the development are Theorem 5.2 and Theorem 6.1. Corollary 3.6 only combines them, with the identity $\mathrm{VCD}(2^{[n]}) = n$, which is immediate from the definition and is itself verified in the development at $n = 4$. The development is five modules and about 740 lines; it contains no `sorry` and declares no axiom by hand. A print of the axiom set of every statement above returns `propext`, `Classical.choice` and `Quot.sound` and nothing else, with one exception: the order-1 exhaustion inside Theorem 6.1 — the $7^6 = 117\,649$ candidate teaching data for six concepts — is discharged by compiled evaluation rather than by kernel reduction, and so additionally depends on the axiom asserting that the compiler's evaluation of that closed Boolean expression is faithful. Every other finite check quoted above, including all of Theorem 5.2, every witness table, and the shattering statements, is reduced by the kernel itself. Alongside the positive statements the development carries negative controls, because a mis-stated finite predicate is the way an exhaustive verification fails silently: that $\mathrm{VCD}(\mathcal{C}_W) \le 1$ is *false*, that the map teaching nothing is *not* non-clashing on $\mathcal{C}_W$, on $2^{[4]}$ and on $\mathcal{B}(G)$, that the witness maps really need their stated order and not less, and that the lower bound of Theorem 3.1 genuinely concerns the full power set — the class of all singletons and $\emptyset$ on five points has NCTD ≤ 1.

Outside the formal development, and labelled as such where they occur, are the census computations: the five-point class census of Section 4, the fragment-free enumeration of Remark 5.5, the planar and all-balls censuses of Remark 6.2, and the search reports of Section 7. These were written in C, they support no theorem above, and each carries its own arithmetic check: the orbit weights of the five-point scan sum to exactly $2^{32} - 1$, the fragment-free orbit sizes sum to the 960 the census reports independently, and the planar graph counts reproduce OEIS A003094. Where a computed value overlaps a published one — $\mathrm{NCTD}(\mathcal{C}_W) = \mathrm{NCTD}^+(\mathcal{C}_W) = 2$ from [1, 2], and "any cycle C_n with $n \geq 6$ admits an NCTM of size $2 < \mathrm{VCD}(\mathcal{B}^*(C_n)) = 3$" from [5] — it reproduces it.

The graph enumerations use `nauty` [10] release 2.7r3 (`geng` and `planarg`); the census programs themselves are plain C. The source of the development and of the census programs is currently held in a private repository: repository reference to be added at submission.

On the reference list. Each entry records how far the reading went. Statement numbers attributed to a preprint are the printed numbering of the version named there, and are version-sensitive: those attributed to [4] are the numbering of its v3 e-print, the version that was distributed with source, and those attributed to [3] are the numbering of its v2, obtained by compiling the v2 e-print and confirmed against the rendering arXiv itself publishes. Statement numbers attributed to [1, 2] are those of the published versions, and the two number differently: what is Theorem 13, Theorem 14, Theorem 19 and Proposition 27 in [1] is Theorem 19, Theorem 21, Theorem 29 and Proposition 39 in [2]. In particular the counting bound is Theorem 13 of [1] but Theorem 19 of [2], while Theorem 19 of [1] is the power-set theorem and Lemma 20 of [1] is its additivity lemma for NCTD^+.

References

[1] D. Kirkpatrick, H. U. Simon and S. Zilles, *Optimal collusion-free teaching*, Proceedings of the 30th International Conference on Algorithmic Learning Theory (ALT 2019), Proceedings of Machine Learning Research **98** (2019), 506–528. Read in full from the publisher's PDF on 3 September 2026: the definitions of a teacher mapping, of the non-clashing property (its Definition 7, after [7]) and of NCTD (its Definition 9); the counting bound (its Theorem 13, with $X(\mathcal{C})$ its Definition 12); the average-degree lower bound $\mathrm{NCTD}(\mathcal{C}) \geq \lceil \frac{1}{2} \deg_{\mathrm{avg}}(\mathcal{C}) \rceil$ (its Theorem 14); the exact value $\mathrm{NCTD}(P_m) = \lceil m/2 \rceil$, $\mathrm{NCTD}^+(P_m) = m$ for the power set on m points (its Theorem 19) and the sentence drawing the consequence for every class over a domain; the sentence about the search quoted in Section 1 (its §6, printed there with VCD and NCTD transposed in the first clause); Warmuth's class as its Table 1; and the value $\mathrm{NCTD}(\mathcal{C}_W) = \mathrm{NCTD}^+(\mathcal{C}_W) = 2$ as its Proposition 27. Volume and page range confirmed against the publisher's landing page and against the bibliography of [2].

[2] S. Fallat, D. Kirkpatrick, H. U. Simon, A. Soltani and S. Zilles, *On batch teaching without collusion*, Journal of Machine Learning Research **24** (2023), paper 40, 1–33. Read in full from the journal's PDF on 3 September 2026: the same search sentence as [1], there without the transposition; the result that NCTD $\leq$ VCD for finite maximum classes, by way of the representation maps of Chalopin, Chepoi, Moran and Warmuth (2022); the counting bound $|\mathcal{C}| \leq 2^d \binom{X(\mathcal{C})}{d}$ (its Theorem 19, with $X(\mathcal{C})$ its Definition 18, and in the form $|\mathcal{C}| \leq 2^d |X_d|$ opening its §3.2.1); the average-degree bound (its Theorem 21); $\mathrm{NCTD}(P_m) = \lceil m/2 \rceil$ (its Theorem 29, with Remark 23 for the lower half); Warmuth's class as its Table 2; and the value $\mathrm{NCTD}(\mathcal{C}_W) = \mathrm{NCTD}^+(\mathcal{C}_W) = 2$ as its Proposition 39.

[3] S. Bhore, L. Khazaliya and F. Mc Inerney, *Non-clashing teaching in graphs: algorithms, complexity, and bounds*, arXiv:2602.00657v2 [cs.CC], v1 31 January 2026, v2 24 March 2026. The abstract page, fetched 3 September 2026, records in its comment field that "An extended abstract will appear in ICLR 2026 proceedings"; that extended abstract was not obtained, and nothing here rests on it. The v2 e-print source was read in full and compiled, on 3 September 2026. Source of the statement of the open question quoted in Section 1, of the pairwise form (1), of the planar bound $\mathrm{NCTD}(\mathcal{B}) \leq 5$ (its Theorem 13 in v2) and of its tightness question quoted in Section 6; it cites $\mathrm{VCD}(\mathcal{B}) \leq 4$ for planar $\mathcal{B}$ to [9].

[4] J. Liu and B. Li, *The no-clash teaching dimension is bounded by VC dimension*, arXiv:2603.23561 [cs.IT], **withdrawn**. Version history read from the arXiv abstract page on 3 September 2026: v1 24 March 2026, v2 26 March 2026 (withdrawn), v3 1 April 2026, v4 3 August 2026 (withdrawn); the comment field reads, in full, "The proof of Lemma 2 is wrong". The v3 e-print was read in full; all quotations in Section 5 are from it, and the statement numbers are its own. Cited here for the claim, the scheme and Lemma 2, and for nothing else.

[5] J. Chalopin, V. Chepoi, F. Mc Inerney and S. Ratel, *Non-clashing teaching maps for balls in graphs*, Proceedings of the Thirty Seventh Conference on Learning Theory (COLT 2024), Proceedings of Machine Learning Research **247** (2024), 840–875; arXiv:2309.02876. PDF read on 3 September 2026. Cited for the proposal of balls in planar graphs as a candidate class with NCTD > VCD, for its closing question about $\mathrm{NCTD}(\mathcal{B}^*(G))$ versus $\mathrm{VCD}(\mathcal{B}^*(G))$ for planar G, and and for its sentence "any cycle C_n with $n \geq 6$ admits an NCTM of size $2 < \mathrm{VCD}(\mathcal{B}^*(C_n)) = 3$", which the

solver used here reproduces (and it computes $\mathrm{NCTD}(\mathcal{B}^*(C_n)) = 2$ exactly, for $5 \le n \le 9$). Volume and page range from dblp.

[6] H. U. Simon, *Tournaments, Johnson graphs and NC-teaching*, Proceedings of the 34th International Conference on Algorithmic Learning Theory (ALT 2023), Proceedings of Machine Learning Research **201** (2023), 1411–1428; arXiv:2205.02792 [math.CO] (there titled *Tournaments, Johnson Graphs, and NC-Teaching*). The e-print was read on 3 September 2026 for the two statements cited: the characterisation of maximum classes of non-clashing teaching dimension 1 by tournaments, and the improvement of $|\mathcal{C}| \le 2^d\binom{n}{d}$ — which it attributes to [1] — to $(2\sqrt{2/(d+1)} - 2/(d+1)) \cdot 2^d\binom{n}{d}$, "a factor of order $\sqrt{d}$", by way of Johnson graphs. It is the only one of the eleven forward citations of [2] recorded by Semantic Scholar that could have contained a census, and it contains none. Volume and page range from dblp and the publisher's landing page.

[7] D. Kuzmin and M. K. Warmuth, *Unlabeled compression schemes for maximum classes*, Journal of Machine Learning Research **8** (2007), paper 70, 2047–2081. Cited only as the origin of the non-clashing condition, following the attribution made by [1] in the sentence introducing its Definition 7; not read here. Bibliographic data confirmed against the journal's landing page and against the bibliographies of [1, 2].

[8] M. Darnstädt, T. Kiss, H. U. Simon and S. Zilles, *Order compression schemes*, Theoretical Computer Science **620** (2016), 73–90, doi:10.1016/j.tcs.2015.10.038. Cited only through the attribution made by [1, 2], that Warmuth's class is the smallest concept class with PBTD > VCD; not read here. Bibliographic data taken from the bibliographies of [1, 2], which agree, and confirmed against Crossref.

[9] L. Duraj, F. Konieczny and K. Potępa, *Better diameter algorithms for bounded VC-dimension graphs and geometric intersection graphs*, Proceedings of the 32nd Annual European Symposium on Algorithms (ESA 2024), LIPIcs **308** (2024), 51:1–51:18, doi:10.4230/LIPIcs.ESA.2024.51. Cited only through [3], which attributes to it the bound $\mathrm{VCD}(\mathcal{B}(G)) \le 4$ for planar G; not read here. Bibliographic data taken from the bibliography of [3]'s e-print and confirmed against dblp.

[10] B. D. McKay and A. Piperno, *Practical graph isomorphism, II*, Journal of Symbolic Computation **60** (2014), 94–112. The `geng` and `planarg` programs of the `nauty` distribution, release 2.7r3, generated the graphs of Remark 6.2.

[11] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science, Springer, 2021, 625–635, doi:10.1007/978-3-030-79876-5_37.

[12] 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, 367–381, doi:10.1145/3372885.3373824.