This is a repository copy of *Vapnik–Chervonenkis dimension and density on Johnson and Hamming graphs*.

Version: Accepted Version

Article:

Vapnik-Chervonenkis Dimension and Density on Johnson and Hamming Graphs

Isolde Adler[b], Bjarki Geir Benediktsson[a,b], Dugald Macpherson[a]

[a]*School of Mathematics, University of Leeds*
[b]*School of Computing, University of Leeds*

Abstract

VC-dimension and VC-density are measures of combinatorial complexity of set systems. VC-dimension was first introduced in the context of statistical learning theory, and is tightly related to the sample complexity in PAC learning. VC-density is a refinement of VC-dimension. Both notions are also studied in model theory, in the context of *dependent* theories. A set system that is definable by a formula of first-order logic with parameters has finite VC-dimension if and only if the formula is a dependent formula.

In this paper we study the VC-dimension and the VC-density of the edge relation Exy on Johnson graphs and on Hamming graphs. On a graph G, the set system defined by the formula Exy is the vertex set of G along with the collection of all *open neighbourhoods* of G. We show that the edge relation has VC-dimension at most 4 on Johnson graphs and at most 3 on Hamming graphs and these bounds are optimal. We furthermore show that the VC-density of the edge relation on the class of all Johnson graphs is 2, and on the class of all Hamming graphs the VC-density is 2 as well. Moreover, we show that our bounds on the VC-dimension carry over to the class of all induced subgraphs of Johnson graphs, and to the class of all induced subgraphs of Hamming graphs, respectively. It also follows that the VC-dimension of the set systems of *closed neighbourhoods* in Johnson graphs and Hamming graphs is bounded.

Johnson graphs and Hamming graphs are well known examples of distance transitive graphs. Neither of these graph classes is nowhere dense nor is there a bound on their (local) clique-width. Our results contrast this by giving evidence of structural tameness of the graph classes.

Keywords: Johnson graphs, Hamming Graphs, VC-dimension, VC-density, graph theory

Email addresses: `I.M.Adler@leeds.ac.uk` (Isolde Adler),
`BjarkiGeirBenediktsson@gmail.com` (Bjarki Geir Benediktsson),
`H.D.MacPherson@leeds.ac.uk` (Dugald Macpherson)

1. Introduction

Vapnik-Chervonenkis dimension (VC-dimension) is a complexity measure of set systems. The related parameter VC-*density* provides a more refined picture of set systems that have bounded VC-dimension. First introduced in the context of statistical learning theory [25], VC-dimension also plays a key role in computational learning [24, 17, 14] as well as in model theory [4], and it has applications in numerous areas, including graph theory [6], computational geometry [8], database theory [22], and graph algorithms and complexity [7, 12]. For the definition of VC-dimension and VC-density see Section 2.

For fixed $k, m \in \mathbb{N}$ with $k \leq m$, the Johnson graph $J(m, k)$ has vertices that correspond to k-element subsets, of an underlying universe set of cardinality m, where two vertices are adjacent if their corresponding sets intersect in $k - 1$ elements. Figure 1 shows the Johnson graph $J(4, 2)$. We let $\mathcal{J} := \{J(m, k) : k, m \in \mathbb{N}, k \leq m\}$ denote the class of all Johnson graphs, and we let $\overline{\mathcal{J}}$ denote the closure of $\mathcal{J}$ under the induced subgraph relation. A first study of induced subgraphs of Johnson graphs has been done in [20].

Hamming graphs arise from Hamming schemes and they naturally model Hamming distance. For fixed $d, q \in \mathbb{N}$, let S be a set with $|S| = q$. The Hamming graph $H(d, q)$ has vertex set S^d, where two vertices are adjacent if they differ in precisely one coordinate. Figure 2 shows the Hamming graph $H(3, 2)$. We let $\mathcal{H} := \{H(d, q) : d, q \in \mathbb{N}\}$ denote the class of all Hamming graphs, and we let $\overline{\mathcal{H}}$ denote the closure of $\mathcal{H}$ under induced subgraphs. The class $\overline{\mathcal{H}}$ has been characterized in [19] via certain edge labellings. The classes $\mathcal{J}, \overline{\mathcal{J}}, \mathcal{H}$, and $\overline{\mathcal{H}}$ admit arbitrarily large cliques as subgraphs, but nevertheless come with a highly regular structure.

Johnson graphs and Hamming graphs are graphs of high regularity. They feature in different areas of computer science and mathematics, including coding theory, algebraic graph theory and model theory. Johnson graphs also appear in László Babai's algorithm for solving the graph isomorphism problem in quasipolynomial time [5], where they constitute the 'hard case'.

Our motivation for this work is multifaceted largely stemming from algorithmic graph theory, permutation group theory, and model theory as mentioned below. In algorithmic graph theory structural *tameness* is often linked to good algorithmic properties. Many problems on graphs, that are algorithmically hard (e.g. NP-hard) in general, can be solved efficiently on classes of graphs having a *tame* structure, such as graphs of bounded tree-width [10], planar graphs, graphs excluding a fixed minor, and nowhere dense classes of graphs [21]. Nowhere dense classes of graphs generalise the previously mentioned classes, and in [16] it was shown that on nowhere dense classes of graphs, every problem expressible in first-order logic is fixed-parameter tractable. All of these classes are sparse. In particular, they cannot contain arbitrarily large cliques. However, intuitively, cliques contain about as much information as independent sets. In [11], clique-width was introduced to address this (the class of all cliques has clique-width 2), and this was further generalised to graph classes of bounded local clique-width. That allowed fixed-parameter tractability for first-order logic [13]. Nowhere
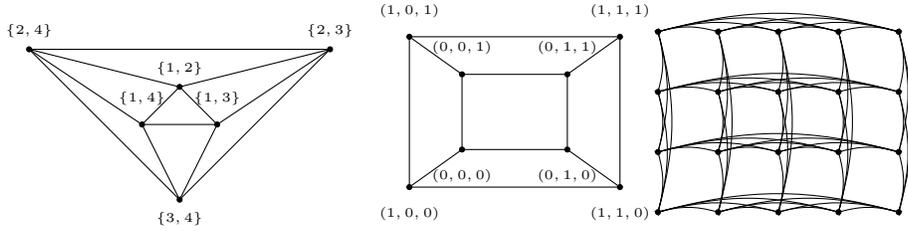
Figure 1: The Johnson graph $J(4,2)$.

Figure 2: The Hamming graph $H(3,2)$.

Figure 3: The rook's graph $R(5.4)$.

dense classes of graphs are closed under taking subgraphs, i.e. if $\mathcal{C}$ is a nowhere dense class of graphs, then the class obtained by closing $\mathcal{C}$ under subgraphs is also nowhere dense. Graph classes of bounded (local) clique-width are closed under taking *induced* subgraphs.

So-called *dependent* graph classes, i.e. graph classes where every first-order formula has bounded VC-dimension, are a common generalisation of both nowhere dense classes of graphs [1] and classes of bounded local clique-width [15]. We will discuss dependent classes below and we view dependence as an interesting notion of tameness. The classes $\mathcal{J}, \overline{\mathcal{J}}, \mathcal{H}$, and $\overline{\mathcal{H}}$ are somewhere dense, as arbitrarily large cliques occur as subgraphs, and they have unbounded local clique width. Indeed, the open neighbourhood of any vertex of $J(m,k)$ induces a *rook's graph $R(m-k,k)$*, cf. Figure 3, and the class of all rook's graphs has unbounded clique-width. Moreover, the open 2-neighbourhood in a Hamming graph $H(d,2)$ induces the 1-subdivision of the complete graph on d vertices, see Corollary 4.4, and it is known that the class of 1-subdivisions of complete graphs has unbounded clique-width (cf. e.g. [2]). While we do not give new algorithms in this paper, our results (see Theorem 1.1) provide evidence of structural tameness, despite unbounded local clique-width.

Hamming graphs and Johnson graphs are regular and have large vertex transitive automorphism groups making them of particular interest in permutation group theory. The symmetric group S_m is the full automorphism group of the Johnson graph $J(m,k)$ whenever $m \neq 2k$, and the wreath product $S_q \mathrm{wr} S_d$ is the full automorphism group of the Hamming graph $H(d,q)$. In both cases these groups act *distance-transitively*: if (u,v) and (u',v') are pairs of vertices with $d(u,v) = d(u',v')$ then there is an element g in the group with $g(u) = u'$ and $g(v) = v'$. This symmetry is exploited in some of our proofs to reduce the number of cases that need to be checked.

A major theme in recent model theory has been the study of structures which are dependent, that is, in which all formulas are dependent as described below. Suppose that M is a first-order structure over a language L, and $\phi(\bar{x}, \bar{y})$ is an L-formula with $\bar{x} = (x_i)_{i=1}^n$ and $\bar{y} = (y_i)_{i=1}^m$ (we write $|\bar{x}| = n$ and $|\bar{y}| = m$). For any $\bar{a} \in M^m$, put $\phi(M, \bar{a}) := \{\bar{x} \in M^n : M \models \phi(\bar{x}, \bar{a})\}$. Then $\{\phi(M, \bar{a}) : \bar{a} \in M^m\}$ is a set system in M^n. This set system has finite VC-dimension if and only if the formula $\phi(\bar{x}, \bar{y})$ is *dependent*, or *NIP* (does not

3

have the *independence property*). Dependent structures include structures with *stable* first-order theory, such as abelian groups, separably closed fields, and free groups, *o-minimal* structures (such as the real field, or even the real field equipped with the exponential function), and many Henselian valued fields such as $\mathbb{Q}_p$. From the viewpoint of model theory, VC-density seems to be both a more refined invariant than VC-dimension, and to be easier to compute. This is the viewpoint developed in the papers [4] and [3]. For background on dependent theories see [23].

If $\mathcal{C}$ is a class of structures in a fixed first-order language, then we say the formula $\phi(\bar{x}, \bar{y})$ is *dependent in* $\mathcal{C}$ if there is $d = d_\phi \in \mathbb{N}$ such that for every $M \in \mathcal{C}$, the set system $\{\phi(M, \bar{a}) : \bar{a} \in M^m\}$ has VC-dimension at most d, and the VC-*dimension of ϕ on the class* is the maximum VC-dimension, if it exists and ∞ otherwise, taken as M ranges through $\mathcal{C}$. The class $\mathcal{C}$ is *dependent* if all formulas are dependent in $\mathcal{C}$. It is known that for fixed integer k, the class $\{J(m, k) : m \in \mathbb{N}\}$ is dependent, because it is first-order definable in the class of all finite sets. Similarly it is also known that for a fixed integer d, the class $\{H(d, q) : q \in \mathbb{N}\}$ is dependent. The main results of this paper give tight bounds in the case that ϕ is the edge relation, i.e. $\phi(x, y) = Exy$, for the classes where *both* parameters vary.

Theorem 1.1. *The edge relation has:*

- VC-*dimension 4 on $\mathcal{J}$, the class of all Johnson graphs.*

- VC-*dimension 3 on $\mathcal{H}$, the class of all Hamming graphs.*

- VC-*density 2 on $\mathcal{J}$, the class of all Johnson graphs.*

- VC-*density 2 on $\mathcal{H}$, the class of all Hamming graphs.*

We show that the VC-dimension of the edge relation does not increase under vertex deletion, see Lemma 2.3 and hence it follows that the VC-dimension of the edge relation on $\overline{\mathcal{J}}$ is 4 and the VC-dimension of the edge relation on $\overline{\mathcal{H}}$ is 3.

It is known that boolean combinations of dependent formulas are dependent and since equality has VC-dimension at most 1 in any model it follows that any property expressible in the language of graphs without quantifiers is dependent in $\mathcal{J}$ and $\mathcal{H}$.

Using the well-known connection between VC-dimension and sample complexity in the probably approximately correct (PAC) model of computational learning theory, our results imply that if $\mathcal{C}$ is a subset of $\overline{\mathcal{J}}$ (or of $\overline{\mathcal{H}}$), then every concept class definable by a quantifier-free first-order formula on $\mathcal{C}$ is learnable with polynomial sample complexity in the PAC model, see e. g. [15, 18].

The techniques we use for the proofs include identifying structural graph properties and symmetries that allow breaking up the problem into a feasible number of cases.

In Section 2 we will cover the basic concepts and notations used throughout the paper. Section 3 contains the results related to Johnson graphs and Section 4 contains results on Hamming graphs.

4

2. Preliminaries

We let $\mathbb{N}$ denote the set of natural numbers including 0. For two sets X and Y we use $X \triangle Y$ to denote the symmetric difference of X and Y i.e. $X \triangle Y = (X \cup Y) \setminus (X \cap Y)$. We use $\mathcal{P}(X)$ to denote the power set of X. We call $|X|$ the *size* of X. For $k \in \mathbb{N}$ we let $\binom{X}{k}$ denote the set of all k-element subsets of X, i.e. $\binom{X}{k} = \{u \subseteq X \ : \ |u| = k\}$.

VC-dimension and VC-density.

Definition 2.1. *A* set system *is a pair* $(X, \mathcal{S})$ *consisting of a* universe set X *and a family* $\mathcal{S} \subseteq \mathcal{P}(X)$ *of subsets of* X.

Set systems are sometimes also referred to as *hypergraphs* or *range spaces*.

Definition 2.2. *Let* $(X, \mathcal{S})$ *be a set system and* $A \subseteq X$ *be a set. We say that* A *is* shattered *by* $\mathcal{S}$ *if the class of intersections of sets in* $\mathcal{S}$ *with* A *is the full powerset of* A, *i.e. if* $\{A \cap W \ : \ W \in \mathcal{S}\} = \mathcal{P}(A)$.

Definition 2.3. *We define the* shatter function $\pi_{\mathcal{S}} : \mathbb{N} \to \mathbb{N}$ *as*

$$\pi_{\mathcal{S}}(n) := \max \big\{ |\{S \cap A \ : \ S \in \mathcal{S}\}| : A \subseteq X, |A| = n \big\}.$$

We use a slight abuse of notation and say that a set A is *maximally shattered* for size n if $|A| = n$ and $\pi_{\mathcal{S}}(n) = |\{S \cap A \ : \ S \in \mathcal{S}\}|$.

Definition 2.4. *The* VC-dimension *of a set system* $(X, \mathcal{S})$ *is*

$$\mathrm{VC}((X, \mathcal{S})) = \begin{cases} \sup\{n \in \mathbb{N} \cup \{\infty\} : \begin{smallmatrix} X \ has \ a \ subset \ of \ size \\ n \ shattered \ by \ \mathcal{S} \end{smallmatrix}\} & \text{if } \mathcal{S} \neq \emptyset \\ -\infty & \text{if } \mathcal{S} = \emptyset. \end{cases}$$

In our work we expand the above concepts to apply to classes of finite graphs in the following way. For a class $\mathcal{C}$ of set systems the VC-dimension of the class is $\mathrm{VC}(\mathcal{C}) = \sup\{\mathrm{VC}(X, \mathcal{S}) : (X, \mathcal{S}) \in \mathcal{C}\}$ if it exists and ∞ otherwise, and the shatter function of $\mathcal{C}$ is $\pi_{\mathcal{C}}(n) = \max\{\pi_{\mathcal{S}}(n) \ : (X, \mathcal{S}) \in \mathcal{C}\}$.

We observe that the shatter function is 2^n for n smaller than the VC-dimension of the set system but for any n greater than the VC-dimension it is bounded above by a polynomial in n. This is due to the Sauer-Shelah Lemma.

Lemma 2.1 (Sauer-Shelah [4]). *If* $(X, \mathcal{S})$ *has finite* VC-*dimension* d *then* $\pi_{\mathcal{S}}(n) \leq \sum_{i=0}^{d} \binom{n}{i}$.

The bound on the degree of the polynomial derived from the VC-dimension need not be tight. Since the degree of the polynomial gives a more precise measure of the combinatorial complexity of a set system, this gives rise to the following definition, which here we only give for classes of finite set systems.

Definition 2.5. *For a class of $\mathcal{C}$ of set systems, the* VC-density *of $\mathcal{C}$ is*

$$\mathrm{vc}(\mathcal{C}) = \begin{cases} \inf\{r \in \mathbb{R}^+ : \pi_{\mathcal{C}}(n) \in \mathcal{O}(n^r)\}, & \text{if } \mathrm{VC}(\mathcal{C}) < \infty \\ \infty \text{ otherwise.} \end{cases}$$

Note that by the Sauer-Shelah Lemma $\mathrm{vc}(\mathcal{C}) \leq \mathrm{VC}(\mathcal{C})$.

Graphs. We consider simple, undirected graphs, i.e. graphs with no self-loops or parallel edges. A *graph G* is a pair $G = (V, E)$ where V is the set of *vertices* of G and $E \subseteq \binom{V}{2}$ is the set of *edges* of G. We also use $V(G)$ to denote the vertex set of G and $E(G)$ to denote the edge set of G. Two vertices u and v are *adjacent*, if $\{u, v\} \in E$. We denote by $N_G(v)$ the neighbourhood of v in G i.e. the set of vertices that are adjacent to v in G and when G is clear from the context we simply write $N(v)$. Note that $v \notin N(v)$. A graph $H = (V', E')$ is an *induced subgraph* of a graph $G(V, E)$, written $H = G[V']$ if $V' \subseteq V$, and $E' = E|_{V'}$, and we say that V' induces H as subgraph of G. A *complete* graph on n vertices, denoted K_n, is a graph (V, E) such that $|V| = n$ and $E = \binom{V}{2}$. For a graph G we say that a set $A \subseteq V(G)$ is a *clique* if it induces a complete graph. We say that A is a *maximal clique* if it is a clique and there is no vertex v such that $A \subseteq N(v)$. A *path* is a sequence $(v_i)_{i=0}^k$ of pairwise distinct vertices such that v_i is adjacent to v_{i+1}, and we say that k is the *length* of the path. The *distance* from vertex v to u, denoted $d(v, u)$, is the minimum length of a path from v to u. The 1-*subdivision* of a graph G is the graph obtained from G by replacing all edges of G by (pairwise internally disjoint) paths of length 2.

Definition 2.6. *For $m, n \in \mathbb{N}$, the* rook's graph *$R(m, n)$ is the graph whose vertex set is $R \times C$ where $|R| = m$ and $|C| = n$ and two distinct vertices $(i, j), (k, l)$ are adjacent if and only if $i = k$ or $j = l$. For a fixed i we call $\{(i, j) : j \in C\}$ the i-th row and $\{(j, i) : j \in R\}$ the i-th column of $R(m, n)$.*

Definition 2.7 (Johnson graphs). *Let $m, k \in \mathbb{N}$ with $m \geq k$ and X be a set with $|X| = m$. The Johnson graph $J(m, k)$ is the graph whose vertex set is $\binom{X}{k}$, where two vertices are adjacent if and only if their intersection has size $k - 1$ i.e. if their symmetric difference has size 2. We call X the underlying set of $J(m, k)$.*

We let $\mathcal{J} := \{J(m, k) : k, m \in \mathbb{N}, k \leq m\}$ denote the class of all Johnson graphs and $\overline{\mathcal{J}}$ its closure under taking induced subgraphs..

Examples of Johnson graphs include the octahedral graph $J(4, 3)$ and the complete graph $K_n = J(n, 1)$. The following lemma is easy to verify.

Lemma 2.2 ([9]). *Let u and v be vertices in a Johnson graph. Then $d(u, v) = |u \triangle v|/2$.*

Definition 2.8 (Hamming graph). *Let $d, q \in \mathbb{N}$ and let S a set with $|S| = q$. The Hamming graph $H(d, q)$ is the graph whose vertices correspond to elements of S^d, where two vertices are adjacent if they agree in all but one coordinate.*

We let $\mathcal{H} := \{H(d,q) \ : \ d, q \in \mathbb{N}\}$ denote the class of all Hamming graphs and $\overline{\mathcal{H}}$ its closure under taking induced subgraphs. Note that $H(2,n) = R(n,n)$.

First-order logic of graphs. The set of all formulas of first-order logic of graphs is defined recursively from the *atomic* formulas 'Exy' and '$x = y$', where x and y are variables and 'Exy' expresses that x and y are joined by an edge, and it is closed under Boolean connectives $\neg, \wedge$ and $\vee$ and existential quantification ($\exists$) and universal quantification ($\forall$) over vertices of the graph. A formula is *quantifier free*, if it does not contain a quantifier. Since we study undirected graphs, for us E is a binary relation that is symmetric and irreflexive. We write $G \models \phi$ to say that the graph G satisfies formula ϕ. In this paper we will focus on the atomic formula Exy, more precisely we are looking at the set systems obtained by it. The set system for Exy in a graph G is $(V(G), \mathcal{S}_E)$ where

$$\mathcal{S}_E := \big\{\{x \ : \ G \models Exy\} \ : \ y \in V(G)\big\} = \{N(v) \ : \ v \in V(G)\}.$$

We say a set A is *shattered by the edge relation* in a graph G if A is a shattered in $(V(G), \mathcal{S}_E)$. Moreover we will say the edge relation has any characteristic (VC-dimension, shatter function, and VC-density) on a graph G that the set system for the edge relation on G has. We write $\mathrm{VC}_E(G)$ for the VC-dimension of the edge relation on a graph G.

Lemma 2.3. *Let G be a graph and $G' := G[V(G) \setminus \{u\}]$ be a graph obtained from G by deleting a single vertex u. Then $\mathrm{VC}_E(G') \leq \mathrm{VC}_E(G)$.*

Proof. For the edge relation we have $\mathcal{S} = \{N(v) \ : \ v \in V(G)\}$. If we delete a vertex u the edge relation on the resulting subgraph G' will give us the class $\mathcal{S}' = \{N(v) \setminus \{u\} \ : \ v \in V(G) \setminus \{u\}\}$. Now assume that $\mathrm{VC}_E(G) < \mathrm{VC}_E(G')$. Then there exists a set $A \subseteq V(G) \setminus \{u\}$ such that $|A| > \mathrm{VC}_E(G)$ and A is shattered by $\mathcal{S}'$. Since $u \notin A$ we have that for all $S \subseteq V(G)$ we get $A \cap S = A \cap (S \setminus \{u\})$, so $\mathcal{P}(A) = \{A \cap S \ : \ S \in \mathcal{S}'\} \subseteq \{A \cap S : S \in \mathcal{S}\}$. That means that $\mathcal{S}$ shatters A, in contradiction with $|A| > \mathrm{VC}_E(G)$. $\square$

3. Johnson Graphs

In this section we will present our results on the VC-dimension and VC-density of the edge relation in Johnson graphs.

Lemma 3.1. *Let v be a vertex in the Johnson graph $J(m,k)$. Then $N(v)$ induces the rook's graph $R(k, m-k)$ as a subgraph of $J(m,k)$.*

Proof. Let v be a vertex in the Johnson graph $J(m,k)$ and without loss of generality assume $v = [1,k] \cap \mathbb{N}$. Every vertex in $N(v)$ has the form $(v \setminus \{a\}) \cup \{x\}$ where $a \in v$ and $x \in [k+1, m] \cap \mathbb{N}$. The mapping $(v \setminus \{a\}) \cup \{x\} \mapsto (a, x-k)$ is a graph isomorphism $J(m,k)[N(v)] \to R(k, m-k)$. $\square$

Lemma 3.2. *Let v and w be vertices in a Johnson graph with $d(v, w) = 1$. Write $w = (v \setminus \{a\}) \cup \{x\}$. Then we have $u \in N(v) \cap N(w)$ if and only if $u = (v \setminus \{c\}) \cup \{z\}$ with exactly one of $c = a$ or $z = x$.*

Proof. Assume $u \in N(v) \cap N(w)$. Then since $d(v, u) = 1$ we must have $u = (v \setminus \{c\}) \cup \{z\}$ for some c and z. Now assume $c \neq a$ and $z \neq x$. Then we have $u \triangle w = \{a, c, x, z\}$ so $|u \triangle v| = 4$ contradicting that $d(u, w) = 1$. So we must have either $c = a$ or $z = x$.

Conversely assume $u = (v \setminus \{a\}) \cup \{z\}$ with $x \neq z$. Then $u \cap v = v \setminus \{a\}$ which has size $k - 1$ so $u \in N(v)$. Also $u \cap w = v \setminus \{a\}$ which has size $k - 1$ so $u \in N(w)$. Thus we have $u \in N(v) \cap N(w)$.

Assume $u = (v \setminus \{c\}) \cup \{x\}$. Then $u \cap v = v \setminus \{c\}$ which has size $k - 1$ so $u \in N(v)$. Also $u \cap w = (v \setminus \{a, c\}) \cup \{x\}$ which has size $k - 1$ so $u \in N(w)$. Thus we have $u \in N(v) \cap N(w)$.

Note that if we have both $c = a$ and $z = x$ then $u = w$ in contradiction with Euw. $\square$

Lemma 3.3. *Let v and w be vertices in a Johnson graph with $d(v, w) = 2$. We can write $w = (v \setminus \{a, b\}) \cup \{x, y\}$. Then we have $u \in N(v) \cap N(w)$ if and only if $u = (v \setminus \{c\}) \cup \{z\}$ with $c \in \{a, b\}$ and $z \in \{x, y\}$.*

Proof. Assume $u \in N(v) \cap N(w)$. Then since $d(v, u) = 1$ we must have $u = (v \setminus \{c\}) \cup \{z\}$ for some $c \in v$ and $z \notin v$.

Now assume $c \notin \{a, b\}$. Then we have $u \triangle w \supseteq \{a, b, c\}$ contradicting that $|u \triangle w| = 2$.

Similarly $z \in \{x, y\}$ as otherwise we have $u \triangle w \supseteq \{x, y, z\}$ in contradiction with $|u \triangle w| = 2$. So we must have $c \in \{a, b\}$ and $z \in \{x, y\}$.

Conversely assume $u = (v \setminus \{c\}) \cup \{z\}$ with $c \in \{a, b\}$ and $z \in \{x, y\}$. Assume without loss of generality $u = (v \setminus \{a\}) \cup \{x\}$. Then $u \cap v = v \setminus \{a\}$ which has size $k - 1$ so $u \in N(v)$. Also $u \cap w = (v \setminus \{a, b\}) \cup \{x\}$ which has size $k - 1$ so $u \in N(w)$. Thus we have $u \in N(v) \cap N(w)$. $\square$

Lemma 3.4. *Let u and v be vertices in the Johnson graph $J(m, k)$ then*

$$
|N(u) \cap N(v)| = \begin{cases} k(m - k) & \text{if } d(u, v) = 0 \\ m - 1 & \text{if } d(u, v) = 1 \\ 4 & \text{if } d(u, v) = 2 \\ 0 & \text{if } d(u, v) \geq 3 \end{cases}
$$

Proof. This follows immediately from Lemmas 3.1, 3.2, and 3.3 $\square$

Lemma 3.5. *Let A be a set of vertices in a Johnson graph shattered by the edge relation and assume $|A| \geq 4$. Then there do not exist three vertices in A pairwise at distance 2 from each other.*

8

Proof. Let v be a vertex such that $A \subseteq N(v)$ and A contains three vertices that are pairwise of distance 2 from each other. That is to say we have $(v \setminus \{a\}) \cup \{x\} \in A$, $(v \setminus \{b\}) \cup \{y\} \in A$, $(v \setminus \{c\}) \cup \{z\} \in A$ where a, b, c, x, y, z are all distinct.

Let w be a vertex such that $N(w) \cap A = \{(v \setminus \{a\}) \cup \{x\}, (v \setminus \{b\}) \cup \{y\}, (v \setminus \{c\}) \cup \{z\}\}$.

If $d(v, w) = 1$ we can write $w = (v \setminus \{a_1\}) \cup \{x_1\}$ by Lemma 3.2. We know that since $(v \setminus \{a\}) \cup \{x\} \in N(w)$ we have $a_1 = a$ or $x_1 = x$.

Assume $a_1 = a$. Then since $(v \setminus \{b\}) \cup \{y\} \in N(w)$ we must have $x_1 = y$, so we have $w = (v \setminus \{a\}) \cup \{y\}$. However $(v \setminus \{c\}) \cup \{z\} \notin N((v \setminus \{a\}) \cup \{y\})$ in contradiction to $N(w) \cap A = \{(v \setminus \{a\}) \cup \{x\}, (v \setminus \{b\}) \cup \{y\}, (v \setminus \{c\}) \cup \{z\}\}$.

Alternatively assume $x_1 = x$. Then since $(v \setminus \{b\}) \cup \{y\} \in N(w)$ we have $a_1 = b$ so we have $w = (v \setminus \{b\}) \cup \{x\}$. However $(v \setminus \{c\}) \cup \{z\} \notin N((v \setminus \{b\}) \cup \{x\})$ in contradiction to $N(w) \cap A = \{(v \setminus \{a\}) \cup \{x\}, (v \setminus \{b\}) \cup \{y\}, (v \setminus \{c\}) \cup \{z\}\}$.

So we must have $d(v, w) = 2$ and write $w = (v \setminus \{a_1, a_2\}) \cup \{x_1, x_2\}$. By Lemma 3.3 we know that since $(v \setminus \{a\}) \cup \{x\} \in N(w)$ we have $a \in \{a_1, a_2\}$ and $x \in \{x_1, x_2\}$. Without loss of generality we assume $a_1 = a$ and $x_1 = x$.

Similarly, since $(v \setminus \{b\}) \cup \{y\} \in N(w)$, we have $b \in \{a, a_2\}$ and $y \in \{x, x_2\}$ so we have $w = (v \setminus \{a, b\}) \cup \{x, y\}$. But then $(v \setminus \{c\}) \cup \{z\} \notin N(w)$, a contradiction.

$\square$

Theorem 3.6. *The* VC-*dimension of the edge relation in a Johnson graph is at most* 4.

Proof. The proof goes through a series of cases demonstrating that no vertex set of size 5 in a Johnson graph can be shattered. We rely on the fact that every set A shattered by the edge relation must have $A \subseteq N(v)$ for some vertex v and that every subset of a shattered set is also shattered which allows us to drastically reduce the number of cases we need to check.

Observe that in $J(m, k)$ we can pick an element of the underlying set and the set of all vertices not containing that element induces $J(m - 1, k)$ as a subgraph of $J(m, k)$ and the set of all vertices containing that element induces $J(m - 1, k - 1)$. Thus we can assume m and k to be arbitrarily large and since by Lemma 2.3 taking induced subgraphs can only decrease the VC-dimension, our argument then holds for all m and k.

We will start by computing the number of configurations that can be obtained by picking 4 vertices out of $N(v)$. Formally the configurations, which we label Case I - Case XVI, are the orbits of the group of automorphisms fixing v in its action on 4 element subsets of $N(v)$. There are 16 and out of those 8 are shattered by the edge relation and 8 are not. We will then go through them one by one. For those cases that are not shattered by the edge relation we will give a proof of why they are not shattered, and in the shattered cases, we will demonstrate that whichever way we choose a fifth vertex to add to those collections we will always end up with a set that is not shattered by the edge relation.

Let A be a set of vertices in a Johnson graph with $|A| = 4$, and v be a vertex such that $A \subseteq N(v)$.

Let $v_i = (v \setminus \{a_i\}) \cup \{x_i\}$ for $i \in \{1, 2, 3, 4\}$ be the four vertices of A. Let $\sim_x$ be the equivalence relation $v_i \sim_x v_j$ if and only if $x_i = x_j$ and $\sim_a$ be the equivalence relation $v_i \sim_a v_j$ if and only if $a_i = a_j$. Note that if we have $v_i \sim_x v_j$ and $v_i \sim_a v_j$ then $v_i = v_j$ and by our assumption that the four vertices are distinct we have $i = j$.

There are 5 ways, up to permutation, to split a set of size 4 into equivalence classes. These correspond to the ways of summing up to 4. Not every combination of equivalence classes for $\sim_a$ and $\sim_x$ is possible. We will now look at each of the ways $\sim_x$ can split A and give the available ways for $\sim_a$ to split A. Note that the equivalence classes of $\sim_a$ and $\sim_x$ correspond to the columns and rows of the rook's graph induced by $N(v)$. We now look at each of the different ways of summing up to 4.

4 In this case we have $x_1 = x_2 = x_3 = x_4$ and we therefore must have $a_i \neq a_j$ whenever $i \neq j$. This means $\sim_a$ has 4 equivalence classes of size 1. This gives us Case IX.

$3 + 1$ Without loss of generality we assume $x_1 = x_2 = x_3 \neq x_4$. Then there are two ways for $\sim_a$ to split A into equivalence classes. It can either have $2 + 1 + 1$ or $1 + 1 + 1 + 1$ as the partition. In the former case we can assume without loss of generality that $a_1 = a_4$ and this yields Case X. In the latter we have $a_i \neq a_j$ whenever $i \neq j$ and this gives us Case I.

$2 + 2$ Without loss of generality we assume $x_1 = x_2 \neq x_3 = x_4$. Note that this implies $a_1 \neq a_2$ and $a_3 \neq a_4$. We now have three ways that $\sim_a$ can split A into equivalence classes.

$2 + 2$ We assume without loss of generality $a_1 = a_3$ and $a_2 = a_4$, giving us Case II.

$2 + 1 + 1$ We assume without loss of generality $a_1 = a_3 \neq a_2$, $a_1 \neq a_4$ and $a_2 \neq a_4$. This gives us Case XI.

$1 + 1 + 1 + 1$ We have $a_i \neq a_j$ whenever $i \neq j$, yielding Case XII.

$2 + 1 + 1$ Without loss of generality we assume $x_1 = x_2 \neq x_3 \neq x_4$ and additionally assume $x_4 \neq x_1$. We can have four ways for $\sim_a$ to split A into equivalence classes.

$3 + 1$ Without loss of generality we can assume $a_1 = a_3 = a_4 \neq a_2$. This is Case $XIII$.

$2 + 2$ Without loss of generality we can assume $a_1 = a_3$ and $a_2 = a_4$. This is Case XIV.

$2 + 1 + 1$ In this instance we have two ways of grouping the vertices with $\sim_a$ that are not equivalent with relabeling.

By making $a_1 = a_3$ we get Case III.

By making $a_3 = a_4$ we get Case IV.

1 + 1 + 1 + 1 We have $a_i \neq a_j$ whenever $i \neq j$, giving us Case V.

$1 + 1 + 1 + 1$ Here we have x_1, x_2, x_3, x_4 all distinct. We can have four ways for $\sim_a$ to split A into equivalence classes.

4 Here we have $a_1 = a_2 = a_3 = a_4$ which is Case XV.

3 + 1 Without loss of generality we may assume $a_1 = a_2 = a_3 \neq a_4$, giving us Case VI.

2 + 2 Without loss of generality we can assume $a_1 = a_2 \neq a_3 = a_4$ which yields Case XVI.

2 + 1 + 1 Without loss of generality we can assume $a_1 = a_2 \neq a_3 \neq a_4$ and $a_1 \neq a_4$ which gives us Case VII.

1 + 1 + 1 + 1 We have $a_i \neq a_j$ whenever $i \neq j$. This gives us Case $VIII$.

We now have 16 cases and will go through them one by one demonstrating that in each case either A can not be shattered or that adding a fifth vertex to A will always result in a set that cannot be shattered. This is sufficient to ensure that no set of size 5 can be shattered. When proving a configuration does not shatter we have to prove that there exists a subset $B \subseteq A$ such that there exists no w for which $N(w) \cap A = B$. In all cases we will have $B \neq A$ so we have to check the cases $d(v, w) = 1$ and $d(v, w) = 2$.

When we have to add a fifth vertex we will have to check every possible combination of $\sim_a$ and $\sim_x$ between the fifth vertex and the previous four vertices, up to a relabeling of the x_i and a_i. In these subcases we will often simply observe that the fifth vertex along with 3 of the original 4 vertices is identical to a case which is separately proved not to shatter.

We will for each case give a diagram showing those vertices of $N(v)$ we are taking to be in A arranged in rows and columns as they would be in the rook graph induced by $N(v)$. In those cases where we do not give a proof that the four vertices selected cannot form a shattered set we will have a choice of how to pick our fifth vertex to add to A. The fifth vertex we will label with the associated subcase rather than v_5 to avoid confusion and save space on the diagrams. Note that row permutations just correspond to relabeling of the equivalence classes of $\sim_x$ and column permutations correspond to relabeling of equivalence classes of $\sim_a$.

There is one more symmetry that we take advantage of which is the isomorphism between $J(m, k)$ and $J(m, m - k)$ which takes a vertex v and maps it to $X \setminus v$. In our diagrams this corresponds to swapping the rows and columns, i.e. transposing the diagram. In those cases where we have already addressed the transpose we will simply refer to the previous case, unless the case is easily resolved.

Case I

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_1\}$$
$$v_4 = (v \setminus \{a_4\}) \cup \{x_2\}$$

Let w be such that $N(w) \cap A = \{v_2, v_3, v_4\}$. We have 2 cases.

(a) $w = (v \setminus \{a\}) \cup \{x\}$. Since we have to exclude v_1 from $N(w)$ we must by Lemma 3.2 have that $a \neq a_1$ and $x \neq x_1$. So in order to have $v_2 \in N(w)$ we must have $a = a_2$ and in order to have $v_3 \in N(w)$ we must have $a = a_3$. But then $a_2 = a_3$ in contradiction with $v_1 \neq v_2$.

(b) $w = (v \setminus \{a, b\}) \cup \{x, y\}$. From Lemma 3.3 we get that $v_2 \in N(w)$ yields $a_2 \in \{a, b\}$ and $x_1 \in \{x, y\}$; $v_3 \in N(w)$ yields $a_3 \in \{a, b\}$ and $x_1 \in \{x, y\}$; $v_4 \in N(w)$ yields $a_4 \in \{a, b\}$ and $x_2 \in \{x, y\}$. Thus $\{a_2, a_3, a_4\} \subseteq \{a, b\}$, contradicting that a_2, a_3, a_4 are all distinct.

Case II

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_2\}) \cup \{x_2\}$$

Let w be such that $N(w) \cap A = \{v_2, v_3, v_4\}$. We have 2 cases.

(a) $w = (v \setminus \{a\}) \cup \{x\}$. Since $v_4 \in N(w)$ we have $w \neq v_1$. Since we have to exclude v_1 from $N(w)$, by Lemma 3.2 we must have that $a \neq a_1$ and $x \neq x_1$. So in order to have $v_2 \in N(w)$ we must have $a = a_2$ and in order to have $v_3 \in N(w)$ we must have $x = x_2$. But then $w = v_4$ in contradiction with $v_4 \in N(w)$.

(b) $w = (v \setminus \{a, b\}) \cup \{x, y\}$. From Lemma 3.3 we get that $v_3 \in N(w)$ yields $a_1 \in \{a, b\}$ and $x_2 \in \{x, y\}$; $v_2 \in N(w)$ yields $a_2 \in \{a, b\}$ and $x_1 \in \{x, y\}$; hence $w = (v \setminus \{a_1, a_2\}) \cup \{x_1, x_2\}$, contradicting that $v_1 \notin N(w)$.

Case III

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_3\}) \cup \{x_3\}$$

The vertices v_2, v_3, v_4 are pairwise at distance 2 from each other so by Lemma 3.5 A is not shattered.

Case IV
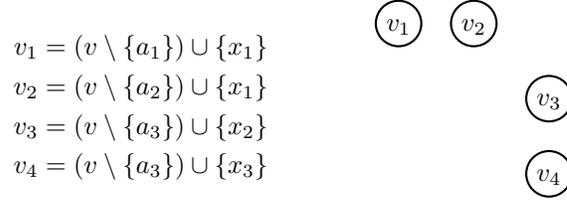
$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_3\}) \cup \{x_3\}$$

(v_1) (v_2)

(v_3)

(v_4)

Let w be such that $N(w) \cap A = \{v_2, v_3, v_4\}$. We have 2 cases.

(a) $w = (v \setminus \{a\}) \cup \{x\}$. Since we have to exclude v_1 from $N(w)$ by Lemma 3.2 we must have that $a \neq a_1$ and $x \neq x_1$. So in order to have $v_2 \in N(w)$ we must have $a = a_2$ and in order to have $v_3 \in N(w)$ we must have $x = x_2$. But then $w = (v \setminus \{a_2\}) \cup \{x_2\}$ in contradiction with $v_4 \in N(w)$.

(b) $w = (v \setminus \{a, b\}) \cup \{x, y\}$. From Lemma 3.3 we get that $v_2 \in N(w)$ yields $a_2 \in \{a, b\}$ and $x_1 \in \{x, y\}$; $v_3 \in N(w)$ yields $a_3 \in \{a, b\}$ and $x_2 \in \{x, y\}$. Thus $w = (v \setminus \{a_2, a_3\}) \cup \{x_1, x_2\}$ in contradiction with $v_4 \in N(w)$.

Case V
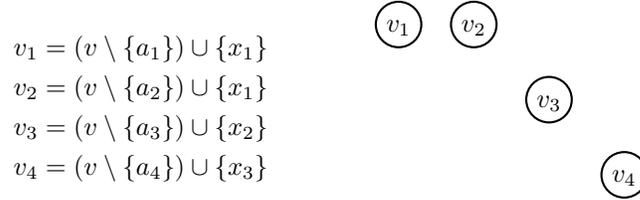
$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_4\}) \cup \{x_3\}$$

(v_1) (v_2)

(v_3)

(v_4)

The vertices v_2, v_3, v_4 are at distance 2 from each other so by Lemma 3.5 A is not shattered.

Case VI

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_3\}$$
$$v_4 = (v \setminus \{a_2\}) \cup \{x_4\}$$

(v_1)

(v_2)

(v_3)

(v_4)
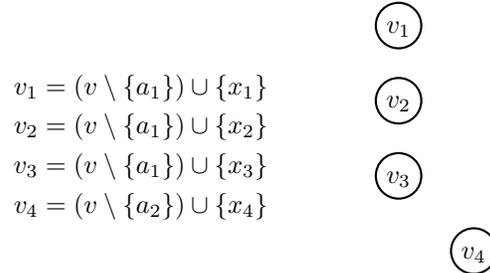
This case is the transpose of Case I which we have already shown cannot shatter so this case cannot shatter either.

Case VII

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_3 = (v \setminus \{a_2\}) \cup \{x_3\}$$
$$v_4 = (v \setminus \{a_3\}) \cup \{x_4\}$$
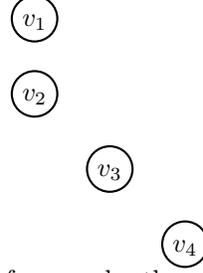
The vertices v_2, v_3, v_4 are at distance 2 from each other so by Lemma 3.5 A is not shattered.

Case VIII

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_2\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_3\}$$
$$v_4 = (v \setminus \{a_4\}) \cup \{x_4\}$$
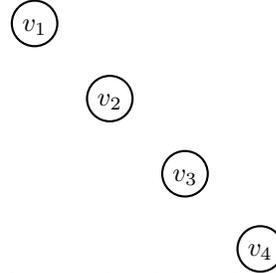
The vertices v_2, v_3, v_4 are at distance 2 from each other so by Lemma 3.5 A is not shattered.

The remaining cases shatter, so we look at the different ways a fifth vertex can be added to the collection and demonstrate that the result cannot be a shattered set.

Case IX

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_1\}$$
$$v_4 = (v \setminus \{a_4\}) \cup \{x_1\}$$
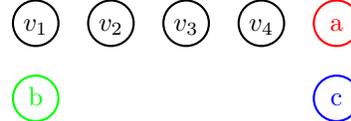
This and each of the remaining cases, are shattered sets so we take a closer look at what configurations are obtainable by adding a fifth vertex to each of them.

a $v_5 = (v \setminus \{a_5\}) \cup \{x_1\}$. Let w be such that $N(w) \cap A = \{v_1, v_2, v_3\}$. Observe that $w \neq v_4$ since $v_5 \in N(v_4)$ so we will need an alternative w. We have 2 cases: either $d(v, w) = 1$ or $d(v, w) = 2$.

Let $w = (v \setminus \{a\}) \cup \{x\}$. Since we have to exclude v_5 from $N(w)$ then by Lemma 3.2 we cannot have $x = x_1$. So in order to have $v_1 \in N(w)$ we must have $a = a_1$ but then in order to have $v_2 \in N(w)$ we must have $x = x_1$, a contradiction.

Let $w = (v \setminus \{a, b\}) \cup \{x, y\}$. In order to have $v_1 \in N(w), v_2 \in N(w)$ and $v_3 \in N(w)$, Lemma 3.3 gives us $\{a_1, a_2, a_3\} \subseteq \{a, b\}$, a contradiction.

14

b $v_5 = (v \setminus \{a_1\}) \cup \{x_2\}$. Here v_2, v_3, v_4, v_5 form case I.

c $v_5 = (v \setminus \{a_5\}) \cup \{x_2\}$. Here v_1, v_2, v_3, v_5 form case I.

Case X

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_1\}$$
$$v_4 = (v \setminus \{a_1\}) \cup \{x_2\}$$
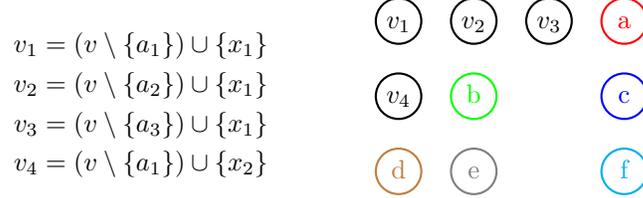
a $v_5 = (v \setminus \{a_4\}) \cup \{x_1\}$. Then v_2, v_3, v_4, v_5 forms case I.

b $v_5 = (v \setminus \{a_2\}) \cup \{x_2\}$. Then v_1, v_2, v_4, v_5 forms case II.

c $v_5 = (v \setminus \{a_4\}) \cup \{x_2\}$. Then v_1, v_2, v_3, v_5 forms case I.

d $v_5 = (v \setminus \{a_1\}) \cup \{x_3\}$. Then v_2, v_3, v_4, v_5 forms case IV.

e $v_5 = (v \setminus \{a_2\}) \cup \{x_3\}$. Then v_1, v_3, v_4, v_5 forms case III.

f $v_5 = (v \setminus \{a_4\}) \cup \{x_3\}$. Then v_1, v_2, v_4, v_5 forms case III.

Case XI

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_3\}) \cup \{x_2\}$$
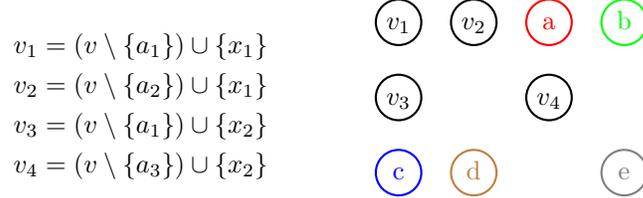
a $v_5 = (v \setminus \{a_3\}) \cup \{x_1\}$. Here v_1, v_3, v_4, v_5 form case II.

b $v_5 = (v \setminus \{a_4\}) \cup \{x_1\}$. Here v_1, v_2, v_4, v_5 form case I.

c $v_5 = (v \setminus \{a_1\}) \cup \{x_3\}$. Here v_2, v_4, v_5 all have distance 2 from each other and thus by Lemma 3.5 A is not shattered.

d $v_5 = (v \setminus \{a_2\}) \cup \{x_3\}$. Here v_1, v_4, v_5 all have distance 2 from each other and thus by Lemma 3.5 A is not shattered.

e $v_5 = (v \setminus \{a_4\}) \cup \{x_3\}$ In this case v_1, v_4, v_5 all have distance 2 from each other and thus by Lemma 3.5 A is not shattered.

Case XII

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_3\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_4\}) \cup \{x_2\}$$
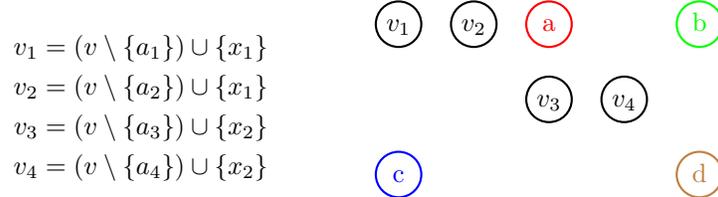
a $v_5 = (v \setminus \{a_3\}) \cup \{x_1\}$. Then v_1, v_2, v_4, v_5 form case I.

b $v_5 = (v \setminus \{a_5\}) \cup \{x_1\}$. Then v_1, v_2, v_3, v_5 form case I.

c $v_5 = (v \setminus \{a_1\}) \cup \{x_3\}$. In this case v_1, v_3, v_4, v_5 form case IV.

d $v_5 = (v \setminus \{a_5\}) \cup \{x_3\}$. In this case v_1, v_3, v_5 all have distance 2 from each other and thus by Lemma 3.5 A is not shattered.

Case XIII

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_1\}) \cup \{x_3\}$$

This is the transpose of **Case X** which we have already shown cannot be contained in a 5 vertex shattered set.

Case XIV

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_2\}) \cup \{x_1\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_4 = (v \setminus \{a_2\}) \cup \{x_3\}$$

This case is the transpose of **Case XI** which we have already shown cannot be a part of a shattered set of size 5.

Case XV

$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_3 = (v \setminus \{a_1\}) \cup \{x_3\}$$
$$v_4 = (v \setminus \{a_1\}) \cup \{x_4\}$$

We note that this is the transpose of **Case IX** which we have already shown is not a part of a 5 vertex shattered set and thus neither can this configuration.

Case XVI
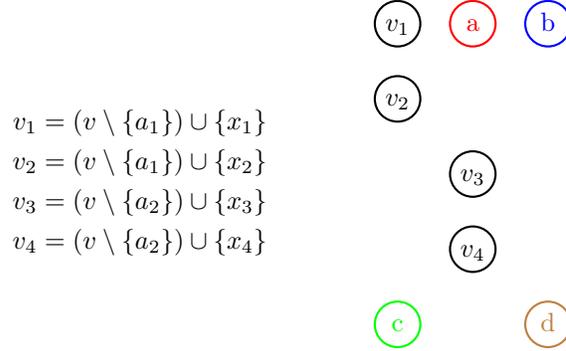
$$v_1 = (v \setminus \{a_1\}) \cup \{x_1\}$$
$$v_2 = (v \setminus \{a_1\}) \cup \{x_2\}$$
$$v_3 = (v \setminus \{a_2\}) \cup \{x_3\}$$
$$v_4 = (v \setminus \{a_2\}) \cup \{x_4\}$$

This case is the transpose of **Case XII** which we have already shown is not a part of any 5 vertex shattered set.

$\square$

Theorem 3.7. *The* VC-*dimension of the edge relation in the Johnson graph* $J(m,k)$ *is* 4 *if and only if* $1 < k < m - 1$ *and* $|V(J(m,k))| = \binom{m}{k} \geq 16$.

Proof. If $|V(J(m,k))| < 16 = 2^4$ then the set system induced by the edge relation has fewer than 16 sets. Thus by the pigeonhole principle the VC-dimension of the edge relation is less than 4.

Assume $\binom{m}{k} \geq 16$ and $1 < k < m - 1$. Here we again rely on $J(m-1, k-1)$ and $J(m-1, k)$ being induced subgraphs of $J(m, k)$. We also observe that $J(m, k)$ is isomorphic to $J(m, m-k)$. So since $\binom{m}{k} \geq 16$ then either $J(7, 2)$ or $J(6, 3)$ are induced subgraphs of $J(m, k)$.

Since removing vertices from a graph can only decrease VC-dimension it now suffices to show that the edge relation has VC-dimension 4 in $J(7, 2)$ and $J(6, 3)$. In Figure 4 we show choices for vertices v_1, v_2, v_3, v_4 such that $A = \{v_1, v_2, v_3, v_4\}$ is shattered by the edge relation, along with how each subset of A can be obtained.

So the VC-dimension of the edge relation is at least 4 in both $J(6, 3)$ and $J(7, 2)$. This shows that the VC-dimension of the edge relation is at least 4 in all Johnson graphs $J(m, k)$ where $\binom{m}{k} \geq 16$ and $1 < k < m - 1$. Theorem 3.6 shows us that the edge relation has VC-dimension at most 4 in all Johnson graphs so this bound is tight whenever $\binom{m}{k} \geq 16$ and $1 < k < m - 1$. $\square$

It is known that boolean combinations of formulas with bounded VC-dimension also have finite VC-dimension (See Lemma 2.9 in [23]). Since the only relations in the language of graphs are the edge relation and equality, and equality always has a VC-dimension at most 1 we get.

Corollary 3.8. *Every quantifier free formula in the language of graphs has finite* VC-*dimension on* $\mathcal{J}$.

Theorem 3.9. *The* VC-*density of the edge relation on* $\mathcal{J}$ *is* 2.

$J(7,2).$

$$v_1 = \{1,3\}$$
$$v_2 = \{1,4\}$$
$$v_3 = \{1,5\}$$
$$v_4 = \{1,6\}$$

$$A \cap N(\{2,7\}) = \emptyset$$
$$A \cap N(\{3,7\}) = \{v_1\}$$
$$A \cap N(\{4,7\}) = \{v_2\}$$
$$A \cap N(\{5,7\}) = \{v_3\}$$
$$A \cap N(\{6,7\}) = \{v_4\}$$
$$A \cap N(\{3,4\}) = \{v_1,v_2\}$$
$$A \cap N(\{3,5\}) = \{v_1,v_3\}$$
$$A \cap N(\{3,6\}) = \{v_1,v_4\}$$
$$A \cap N(\{4,5\}) = \{v_2,v_3\}$$
$$A \cap N(\{4,6\}) = \{v_2,v_4\}$$
$$A \cap N(\{5,6\}) = \{v_3,v_4\}$$
$$A \cap N(v_4) = \{v_1,v_2,v_3\}$$
$$A \cap N(v_3) = \{v_1,v_2,v_4\}$$
$$A \cap N(v_2) = \{v_1,v_3,v_4\}$$
$$A \cap N(v_1) = \{v_2,v_3,v_4\}$$
$$A \cap N(\{1,2\}) = A$$

$J(6,3).$

$$v_1 = \{2,3,4\}$$
$$v_2 = \{1,3,4\}$$
$$v_3 = \{1,3,5\}$$
$$v_4 = \{1,2,5\}$$

$$A \cap N(\{4,5,6\}) = \emptyset$$
$$A \cap N(\{2,3,6\}) = \{v_1\}$$
$$A \cap N(\{v_1\}) = \{v_2\}$$
$$A \cap N(\{v_4\}) = \{v_3\}$$
$$A \cap N(\{1,2,6\}) = \{v_4\}$$
$$A \cap N(\{3,4,6\}) = \{v_1,v_2\}$$
$$A \cap N(v_2) = \{v_1,v_3\}$$
$$A \cap N(\{2,4,5\}) = \{v_1,v_4\}$$
$$A \cap N(\{1,3,6\}) = \{v_2,v_3\}$$
$$A \cap N(v_3) = \{v_2,v_4\}$$
$$A \cap N(\{1,5,6\}) = \{v_3,v_4\}$$
$$A \cap N(\{3,4,5\}) = \{v_1,v_2,v_3\}$$
$$A \cap N(\{1,2,4\}) = \{v_1,v_2,v_4\}$$
$$A \cap N(\{2,3,5\}) = \{v_1,v_3,v_4\}$$
$$A \cap N(\{1,4,5\}) = \{v_2,v_3,v_4\}$$
$$A \cap N(\{1,2,3\}) = A$$

Figure 4: Examples of shattered sets of size 4 in $J(7,2)$ and $J(6,3)$

Proof. First we show that the VC-density is at least 2. Assume without loss of generality that $m > 2k$ and let X be the underlying set of $J(m, k)$. Fix a vertex $v = \{a_i : 1 \leq i \leq k\}$ in $J(m, k)$, and let $(x_i)_{i=1}^k$ be distinct elements of X such that $x_i \notin v$ for all i. Define $A := \{(v \setminus \{a_i\}) \cup \{x_i\} : 1 \leq i \leq k\}$. Then for any pair of vertices $v_i := (v \setminus \{a_i\}) \cup \{x_i\}$ and $v_j = (v \setminus \{a_j\}) \cup \{x_j\}$ we have that $N((v \setminus \{a_i\}) \cup \{x_j\}) \cap A = \{v_i, v_j\}$. There are $\frac{|A|^2 - |A|}{2}$ such pairs so the VC-density of the edge relation on $\mathcal{J}$ is at least 2.

Now we show that the VC-density of the edge relation on $\mathcal{J}$ is at most 2. Let A be a set of vertices in $J(m, k)$, and $\pi(n)$ be the shatter function for the edge relation on $J(m, k)$. Let $|A| = n$ and A be maximally shattered by the edge relation for sets of size n. Let

$$
\begin{aligned}
S(A) &= \{N(u) \cap A : u \in V(G)\}, \\
C_1(A) &= \{N \in S(A) : N \text{ is a clique}\}, \text{ and} \\
C_2(A) &= \{N \in S(A) : N \text{ is not a clique}\}.
\end{aligned}
$$

By our assumption that A is maximally shattered we have $|S(A)| = \pi(n)$. Note also that $S(A) = C_1(A) \cup C_2(A)$ so we deal with those two cases separately.

$|C_1(A)| \leq \frac{5|A|^2 + 3|A|}{2}$: There are at most $\frac{|A|^2 + |A|}{2}$ cliques of size 2 or less in $S(A)$. There are at most $|A|$ cliques C in $S(A)$ such that $C = A \cap N(v)$ for some $v \in A$.

Now assume we have $C = A \cap N(v)$ for some $v \notin A$ and further assume that $|C| \geq 3$. We want to show that then the clique C is of the form $A \cap Q$ for some maximal clique Q of $J(m, k)$. We then argue that there can be at most $2|A|^2$ maximal cliques of $J(m, k)$ that intersect A in more than one vertex.

Note that in any graph G a maximal clique Q of G is contained in $N(u) \cup \{u\}$ for all $u \in Q$ so $Q \setminus \{u\}$ is a maximal clique in $G[N(u)]$. It is easy to see that the maximal cliques of the rook's graph $R(m, k)$ are the rows and columns. So by Lemma 3.1 we find that for every vertex u in $J(m, k)$ the maximal cliques of $J(m, k)$ that u belongs to are of the form $Z \cup \{u\}$ where Z is a row or a column of the rook's graph $J(m, k)[N(u)]$.

Since $|C| \geq 3$ we know by Lemma 3.2 the only vertices connected to all vertices in C are v and those vertices that share that row or column with all of C, in the rook's graph induced by $N(v)$, and therefore lie in $N(v)$. It follows that $C = A \cap Q$ for some maximal clique Q of $J(m, k)$.

For every vertex $u \in A$ we have that A intersects at most $|A|$ rows and at most $|A|$ columns of the rook's graph induced by $N(u)$. So u can be a member of at most $2|A|$ maximal cliques of $J(m, k)$ that intersect A in more than two vertices. So the number of maximal cliques of $J(m, k)$ that intersect A in more than two vertices is at most $2|A|^2$.

$|C_2(A)| \leq 4|A|^2$: This holds since every pair of vertices at distance 2 from each other can by Lemma 3.4 be contained in the neighbourhood of at most 4 vertices and there are at most $|A|^2$ such pairs.

So we get that $|S(A)| \leq |C_1(A)| + |C_2(A)| \leq \frac{5|A|^2 + 3|A|}{2} + 4|A|^2 = \frac{13|A|^2 + 3|A|}{2} \in \mathcal{O}(|A|^2)$. $\square$

4. Hamming Graphs

In this section we will give technical lemmas for dealing with Hamming graphs and prove our main results on the VC-dimension and VC-density of the edge relation in such graphs.

Lemma 4.1. *Let v be a vertex in the Hamming graph $H(d,q)$. Then $N(v)$ induces a disjoint union of d copies of K_{q-1}.*

Proof. We observe that for each coordinate j the set of neighbours of v that disagree with v in the j-th coordinate has size $q-1$ and since those vertices all agree in all but the j-th coordinate they form a clique. If two vertices $u, w \in N(v)$ disagree with v in different coordinates, say i and j respectively, then u and w disagree with each other in the i-th and the j-th coordinate and thus they are non-adjacent. $\square$

Lemma 4.2. *Let u and v be vertices in the Hamming Graph $H(d,q)$ with $d(u,v) = 1$. Let $1 \leq i \leq d$ be such that u and v agree on all but the i-th coordinate. Then $N(u) \cap N(v)$ is a clique of size $q-2$ whose members are all vertices w that agree with u and v in all but the i-th coordinate.*

Proof. Since u and v are neighbours we know that they agree in all but one coordinate namely the i-th. All vertices that agree with u and v on all coordinates except the i-th form a clique. Since each coordinate can have q different values there are $q-2$ such vertices that are neither u nor v.

Take any vertex w neighboring u that agrees with u on all coordinates except the j-th for a $j \neq i$. Now we know that w and u agree on the i-th coordinate but since u and v disagree on this coordinate we get that w and v disagree on the i-th and the j-th coordinate and therefore do not have an edge between them. $\square$

Lemma 4.3. *Let $u = (u_k)_{k=1}^d$ and $v = (v_k)_{k=1}^d$ be vertices in the Hamming Graph $H(d,q)$ with $d(u,v) = 2$. Let $1 \leq i < j \leq d$ be such that $u_i \neq v_i$, $u_j \neq v_j$ and $u_k = v_k$ for every $k \notin \{i,j\}$. Then $N(u) \cap N(v)$ has exactly two vertices and they are not connected, namely $x = (x_k)_{k=1}^d$ and $y = (y_k)_{k=1}^d$ where $x_i = u_i$ and $x_k = v_k$ for all $k \neq i$, and $y_i = v_i$ and $y_k = u_k$ for all $k \neq i$.*

Proof. Since u and v disagree on both the j-th and the i-th coordinates any vertex $w \in N(u) \cap N(v)$ will have to agree with u on either the i-th or the j-th coordinate and with v on the other one of those. $\square$

Lemma 4.3 implies the following.

Corollary 4.4. *The open 2-neighborhood of a vertex in $H(d,2)$ contains the 1-subdivision of K_{d-1} as an induced subgraph.*

Since $H(d,2)$ is an induced subgraph of $H(d,q)$ for $q \geq 2$ it follows that $\mathcal{H}$ has unbounded local clique-width as mentioned in the introduction.

Lemma 4.5. *Let u and v be vertices in the Hamming graph $H(d, q)$ then*

$$|N(v) \cap N(w)| = \begin{cases} d(q-1) & \text{if } d(u,v) = 0 \\ q-2 & \text{if } d(u,v) = 1 \\ 2 & \text{if } d(u,v) = 2 \\ 0 & \text{if } d(u,v) \geq 3 \end{cases}$$

Proof. This follows immediately from Lemmas 4.1, 4.2, and 4.3. □

Theorem 4.6. *The VC-dimension of the edge relation in a Hamming graph is at most 3.*

Proof. Assume there is a set A' with $|A'| > 3$ which is shattered by the edge relation. Then there is a set $A \subseteq A'$ with $|A| = 4$ which is shattered by the edge relation. Let $A = \{v_1, v_2, v_3, v_4\}$, let v be such that $N(v) \cap A = A$ and w be a vertex such that $N(w) \cap A = \{v_1, v_2, v_3\}$. Since $v \neq w$ and $|N(v) \cap N(w)| > 2$ we have that $d(v, w) = 1$, so the intersection of $N(v)$ and $N(w)$ is a clique. Now we have two cases: either $v_4 = w$ or $v_4 \neq w$.

Assume $v_4 = w$ so A induces a clique. Let u be such that $N(u) \cap A = \{v_1, v_2\}$. Since A is a clique we know that $u \notin A$. More importantly u cannot belong to the copy of K_{q-1} in $N(v)$ that contains A so by Lemma 4.1 $d(u, v) = 2$. But then $N(v) \cap N(u)$ by Lemma 4.3 has two vertices that are not adjacent, in contradiction with A being a clique.

Assume $v_4 \neq w$. Then we know that $v_4 \notin N(v) \cap N(v_1)$ since otherwise it would be in $N(w)$ in contradiction with $N(w) \cap A = \{v_1, v_2, v_3\}$. Then $d(v_4, v_1) = 2$ and similarly $d(v_4, v_2) = 2$. Let u be a vertex such that $N(u) \cap A = \{v_1, v_2, v_4\}$. Since $u \neq v$ and $|N(u) \cap N(v)| > 2$ we have by Lemma 4.2 that $N(u) \cap A \subseteq N(u) \cap N(v)$ is a clique, in contradiction with $d(v_4, v_1) = 2$. □

Theorem 4.7. *The VC-dimension of the edge relation on the Hamming graph $H(d, q)$ is 3 if and only if at least one of the following holds.*

1. *$d \geq 3$ and $q \geq 3$.*

2. *$d \geq 2$ and $q \geq 4$.*

3. *$d \geq 4$ and $q \geq 2$.*

Proof. It is clear that the VC-dimension is less than 3 whenever $q^d < 8$. Since $H(1, q) = K_q$, it follows that the VC-dimension is 1 when $d < 2$.

Note that if $d \leq d'$ and $q \leq q'$ then $H(d, q)$ is an induced subgraph of $H(d', q')$. Since removing vertices from a graph can only decrease VC-dimension it now suffices to show that $H(2, 3)$ and $H(3, 2)$ have no shattered set of size 3, and that the edge relation has VC-dimension 3 in $H(3, 3)$, $H(2, 4)$, and $H(4, 2)$.

Note that a set A shattered by the edge relation must have $A \subseteq N(v)$ for some vertex v. Moreover since Hamming graphs are vertex transitive we have that for any vertex u there is an automorphism f such that $u = f(v)$ and

A	$H(2,4)$	$H(3,3)$	$H(4,2)$
v_1	$(0,1)$	$(0,0,1)$	$(0,0,0,1)$
v_2	$(0,2)$	$(0,1,0)$	$(0,0,1,0)$
v_3	$(0,3)$	$(1,0,0)$	$(0,1,0,0)$
$A \cap N(x)$	$H(2,4)$	$H(3,3)$	$H(4,2)$
$\emptyset$	$(1,0)$	$(1,1,1)$	$(1,1,1,1)$
$\{v_1\}$	$(1,1)$	$(0,0,2)$	$(1,0,0,1)$
$\{v_2\}$	$(2,2)$	$(0,2,0)$	$(1,0,1,0)$
$\{v_3\}$	$(3,3)$	$(2,0,0)$	$(1,1,0,0)$
$\{v_1,v_2\}$	$(0,3)$	$(0,1,1)$	$(0,0,1,1)$
$\{v_1,v_2\}$	$(0,2)$	$(0,0,1)$	$(0,1,0,1)$
$\{v_1,v_2\}$	$(0,1)$	$(1,1,0)$	$(0,1,1,0)$
$\{v_1,v_2,v_3\}$	$(0,0)$	$(0,0,0)$	$(0,0,0,0)$

Table 1: Examples of shattered sets in $H(2,4)$, $H(3,3)$ and $H(4,2)$

$f[A] \subseteq N(u)$. In other words, if there is a shattered set A of size 3 then for any vertex u there is some shattered set A' of size 3 such that $A' \subseteq N(u)$. To show that $H(2,3)$ and $H(3,2)$ have VC-dimension less than 3, it is therefore sufficient to show that for a given vertex v, no subset of $N(v)$ of size 3 is shattered.

$H(2,3)$. The neighbourhood of a vertex v in $H(2,3)$ has 4 vertices each adjacent to exactly 1 other vertex in $N(v)$. Thus a 3 element subset A of $N(v)$ contains two adjacent vertices u,w and one vertex connected to neither of them. By Lemma 4.3 we get that $|N(u) \cap N(w)| = 1$ which gives us $N(u) \cap N(w) = \{v\}$. Since $N(v) \cap A = A$ there is no vertex v' such that $N(v') = \{u,w\}$.

$H(3,2)$. Since for any vertex v in $H(3,2)$ we have $|N(v)| = 3$ then a set A of size 3 shattered by the edge relation must have the form $A = N(v)$ for some v. It is therefore sufficient to show that $N(v)$ is not shattered for some v. Without loss of generality assume $v = (a,a,a)$. Observe that $N((a,a,a)) = \{(a,a,b),(a,b,a),(b,a,a)\}$ and any neighbour of (a,a,b) is also a neighbour of (a,b,a) or (b,a,a). Thus there does not exist a vertex u such that $N(u) \cap N((a,a,a)) = \{(a,a,b)\}$ so $N((a,a,a))$ is not a shattered set, and thus $H(3,2)$ has no shattered set of size 3.

In Table 1 we give examples of shattered sets $A = \{v_1, v_2, v_3\}$ in $H(2,4)$, $H(3,3)$, and $H(4,2)$. The last 3 columns in the second table show choices of x, in the different graphs, such that $A \cap N(x)$ is the subset shown in the first column.

$\square$

Theorem 4.8. *The* VC-*density of the edge relation on $\mathcal{H}$ is 2.*

Proof. First we observe that for $d > 1$ a set such that any two vertices agree on all but the first two coordinates has the property that $\forall u, v \in A \exists w (A \cap N(w) = \{u,v\})$ so the VC-density is at least 2.

We now need to show that $\pi_{\mathcal{H}}(n) \in \mathcal{O}(n^2)$ where $\pi_{\mathcal{H}}$ is the shatter function for the edge relation on $\mathcal{H}$. We do this by giving a bound on a recursive formula for $\pi(n)$ and showing that it has a $\mathcal{O}(n^2)$ closed form.

Let A be a maximally shattered set of size n in the Hamming graph $H(d, q)$. Let $v \in A$ and $\mathcal{S}$ be the class of all neigbourhoods in $H(d, q)$. Let $S_1 = \{A \cap S : S \in \mathcal{S} \wedge v \in S\}$. Let $S_2 = \{A \cap S : S \in \mathcal{S} \wedge v \notin S\}$. Note that $|S_1 \cup S_2| = \pi(n)$ and $|S_2| \leq \pi(n-1)$.

Every member of S_1 is an intersection of A with a neighborhood of neighbour of v. Also, $N(v)$ induces a disjoint union of d copies of K_{q-1} with no edges between the cliques. Let

$$D_0 = \{v\}, \qquad\qquad D_1 = A \cap N(v),$$
$$D_2 = \{u \in A : d(u, v) = 2\}, \text{ and} \qquad D_3 = \{u \in A : d(u, v) > 2\}$$

Then D_3 intersects no member of S_1 by definition of D_3. By Lemma 4.3 every element of D_2 can be a member of at most 2 sets of S_1 thus the total number of distinct sets containing v and intersecting D_2 is at most $2|D_2| < 2n$.

Since we have counted all members of S_1 that intersect D_2, and no members of S_1 intersect D_3 we only have left to count those members of S_1 that are subsets of $D_0 \cup D_1$. By Lemma 4.1, $N(v)$ induces a disjoint union of d copies of K_{q-1}. Let $(Q_i)_{i=1}^d$ be sets such that for each i, Q_i is the set of all vertices $u \in D_1$ that disagree with v in the i-th coordinate. Note that $D_1 = \bigcup_{i=1}^d Q_i$ and any element of S_1 which is a subset of $D_0 \cup D_1$ is a subset of $D_0 \cup Q_i$ for some i.

Moreover every subset of $D_0 \cup Q_i$ that is an element of S_1 is either: $(D_0 \cup Q_i) \setminus \{u\}$ for some $u \in Q_i$, or $D_0 \cup Q_i$, or $\{v\}$, thus the number of distinct elements of S_1 contained in $D_0 \cup D_1$ is at most

$$\sum_{i=1}^d |Q_i| + \min(d, n) + 1 = |D_1| + \min(d, n) + 1 \leq n + \min(d, n) + 1.$$

So we have

$$\pi(n) = |S_1| + |S_2| \leq |S_1| + \pi(n-1) \leq 2|D_2| + n + \min(d, n) + 1 + \pi(n-1)$$
$$\leq 2n + n + n + 1 + \pi(n-1) \leq 4n + 1 + \pi(n-1).$$

By induction we get that $\pi(n) \leq 4n^2 + n$ for all n and thus $\pi(n) \in \mathcal{O}(n^2)$. This tells us that the VC-density is at most 2. We have thus demonstrated that the VC-density of the edge relation on $\mathcal{H}$ is at least 2 and at most 2 and conclude that it must be 2. $\qquad\square$

References

[1] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *Eur. J. Comb.*, 36:322–330, 2014. doi: 10.1016/j.ejc.2013.06.048. URL https://doi.org/10.1016/j.ejc.2013.06.048.

[2] Isolde Adler, Binh-Minh Bui-Xuan, Yuri Rabinovich, Gabriel Renault, Jan Arne Telle, and Martin Vatshelle. On the boolean-width of a graph: Structure and applications. In Dimitrios M. Thilikos, editor, *Graph Theoretic Concepts in Computer Science - 36th International Workshop, WG 2010, Zarós, Crete, Greece, June 28-30, 2010 Revised Papers*, volume 6410 of *Lecture Notes in Computer Science*, pages 159–170, 2010. doi: 10.1007/978-3-642-16926-7_16. URL `https://doi.org/10.1007/978-3-642-16926-7_16`.

[3] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, Sergei Starchenko, et al. Vapnik–Chervonenkis density in some theories without the independence property, II. *Notre Dame Journal of Formal Logic*, 54(3-4):311–363, 2013.

[4] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-Chervonenkis density in some theories without the independence property, I. *Transactions of the American Mathematical Society*, 368(8):5889–5949, 2016.

[5] László Babai. Graph isomorphism in quasipolynomial time. *CoRR*, abs/1512.03547, 2015. URL `http://arxiv.org/abs/1512.03547`.

[6] Nicolas Bousquet and Stéphan Thomassé. VC-dimension and Erdős–Pósa property. *Discrete Mathematics*, 338(12):2302 – 2317, 2015. ISSN 0012-365X. doi: https://doi.org/10.1016/j.disc.2015.05.026. URL `http://www.sciencedirect.com/science/article/pii/S0012365X15002174`.

[7] Karl Bringmann, László Kozma, Shay Moran, and N. S. Narayanaswamy. Hitting set for hypergraphs of low VC-dimension. In Piotr Sankowski and Christos D. Zaroliagis, editors, *24th Annual European Symposium on Algorithms, ESA 2016, August 22-24, 2016, Aarhus, Denmark*, volume 57 of *LIPIcs*, pages 23:1–23:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi: 10.4230/LIPIcs.ESA.2016.23. URL `https://doi.org/10.4230/LIPIcs.ESA.2016.23`.

[8] Bernard Chazelle and Emo Welzl. Quasi-optimal range searching in space of finite VC-dimension. *Discret. Comput. Geom.*, 4:467–489, 1989. doi: 10.1007/BF02187743. URL `https://doi.org/10.1007/BF02187743`.

[9] Sebastian M Cioabă, Brandon D Gilbert, Jack H Koolen, and Brendan D McKay. Addressing Johnson graphs, complete multipartite graphs, odd cycles and other graphs. *arXiv preprint arXiv:1808.04757*, 2018.

[10] Bruno Courcelle. Graph rewriting: An algebraic and logic approach. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 193–242. MIT Press, 1990. ISBN 0-444-88074-7.

[11] Bruno Courcelle and Stephan Olariu. Upper bounds to the clique width of graphs. *Discrete Applied Mathematics*, 101(1-3):77–114, 2000.

[12] Kord Eickmeyer, Archontia C. Giannopoulou, Stephan Kreutzer, O-joung Kwon, Michal Pilipczuk, Roman Rabinovich, and Sebastian Siebertz. Neighborhood complexity and kernelization for nowhere dense classes of graphs. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, volume 80 of *LIPIcs*, pages 63:1–63:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. ISBN 978-3-95977-041-5. doi: 10.4230/LIPIcs.ICALP.2017.63. URL `https://doi.org/10.4230/LIPIcs.ICALP.2017.63`.

[13] Martin Grohe. Logic, graphs, and algorithms. In *Logic and Automata – History and Perspectives*, pages 357–422. 2007.

[14] Martin Grohe and Martin Ritzert. Learning first-order definable concepts over structures of small degree. In *32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, Reykjavik, Iceland, June 20-23, 2017*, pages 1–12. IEEE Computer Society, 2017. doi: 10.1109/LICS.2017.8005080. URL `https://doi.org/10.1109/LICS.2017.8005080`.

[15] Martin Grohe and Gyorgy Turán. Learnability and definability in trees and similar structures. *Theory Comput. Syst.*, 37(1):193–220, 2004.

[16] Martin Grohe, Stephan Kreutzer, and Sebastian Siebertz. Deciding first-order properties of nowhere dense graphs. In D. Shmoys, editor, *STOC*, pages 89–98. ACM, 2014. doi: 10.1145/2591796.2591851. URL `http://doi.acm.org/10.1145/2591796.2591851`.

[17] Lisa Hellerstein, Krishnan Pillaipakkamnatt, Vijay Raghavan, and Dawn Wilkins. How many queries are needed to learn? *J. ACM*, 43(5):840–862, 1996. doi: 10.1145/234752.234755. URL `https://doi.org/10.1145/234752.234755`.

[18] Michael J. Kearns and Umesh V. Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, 1994. ISBN 978-0-262-11193-5. URL `https://mitpress.mit.edu/books/introduction-computational-learning-theory`.

[19] Sandi Klavzar and Iztok Peterin. Characterizing subgraphs of Hamming graphs. *Journal of Graph Theory*, 49(4):302–312, 2005. doi: 10.1002/jgt.20084. URL `https://doi.org/10.1002/jgt.20084`.

[20] Ramin Naimi and Jeffrey Shaw. Induced subgraphs of Johnson graphs. *Involve*, 5, 08 2010. doi: 10.2140/involve.2012.5.25.

[21] Jaroslav Nešetřil and Patrice Ossona de Mendez. On nowhere dense graphs. *Eur. J. Comb.*, 32(4):600–617, 2011. doi: 10.1016/j.ejc.2011.01.006. URL `http://dx.doi.org/10.1016/j.ejc.2011.01.006`.

[22] Matteo Riondato, Mert Akdere, Uğur Çetintemel, Stanley B. Zdonik, and Eli Upfal. The VC-dimension of SQL queries and selectivity estimation through sampling. In Dimitrios Gunopulos, Thomas Hofmann, Donato Malerba, and Michalis Vazirgiannis, editors, *Machine Learning and Knowledge Discovery in Databases*, pages 661–676, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[23] Pierre Simon. *A guide to NIP theories*. Lecture Notes in Logic. Cambridge University Press, 2015. doi: 10.1017/CBO9781107415133.002.

[24] Leslie G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, 1984. doi: 10.1145/1968.1972. URL https://doi.org/10.1145/1968.1972.

[25] Vladimir Vapnik and Alexey Chervonenkis. *On the Uniform Convergence of the Frequencies of Occurrence of Events to Their Probabilities*, page 16:264–280. 1971.