On density of subgraphs of halved cubes

Victor Chepoi, Arnaud Labourel, Sébastien Ratel

On density of subgraphs of halved cubes

In memory of Michel Deza

Victor Chepoi, Arnaud Labourel, and Sébastien Ratel

Laboratoire d'Informatique Fondamentale, Aix-Marseille Université and CNRS,
Faculté des Sciences de Luminy, F-13288 Marseille Cedex 9, France
{victor.chepoi, arnaud.labourel, sebastien.ratel}@lif.univ-mrs.fr

Abstract. Let $\mathcal{S}$ be a family of subsets of a set X of cardinality m and VC-dim($\mathcal{S}$) be the Vapnik-Chervonenkis dimension of $\mathcal{S}$. Haussler, Littlestone, and Warmuth (Inf. Comput., 1994) proved that if $G_1(\mathcal{S}) = (V, E)$ is the subgraph of the hypercube Q_m induced by $\mathcal{S}$ (called the 1-inclusion graph of $\mathcal{S}$), then $\frac{|E|}{|V|} \leq$ VC-dim($\mathcal{S}$). Haussler (J. Combin. Th. A, 1995) presented an elegant proof of this inequality using the shifting operation.

In this note, we adapt the shifting technique to prove that if $\mathcal{S}$ is an arbitrary set family and $G_{1,2}(\mathcal{S}) = (V, E)$ is the 1,2-inclusion graph of $\mathcal{S}$ (i.e., the subgraph of the square Q_m^2 of the hypercube Q_m induced by $\mathcal{S}$), then $\frac{|E|}{|V|} \leq \binom{d}{2}$, where $d :=$ cVC-dim$^*(\mathcal{S})$ is the clique-VC-dimension of $\mathcal{S}$ (which we introduce in this paper). The 1,2-inclusion graphs are exactly the subgraphs of halved cubes and comprise subgraphs of Johnson graphs as a subclass.

1. Introduction

Let $\mathcal{S}$ be a family of subsets of a set X of cardinality m and VC-dim($\mathcal{S}$) be the Vapnik-Chervonenkis dimension of $\mathcal{S}$. Haussler, Littlestone, and Warmuth [19, Lemma 2.4] proved that if $G_1(\mathcal{S}) = (V, E)$ is the subgraph of the hypercube Q_m induced by $\mathcal{S}$ (called the *1-inclusion graph* of $\mathcal{S}$), then the following fundamental inequality holds: $\frac{|E|}{|V|} \leq$ VC-dim($\mathcal{S}$). They used this inequality to bound the worst-case expected risk of a prediction model of learning of concept classes $\mathcal{S}$ based on the bounded degeneracy of their 1-inclusion graphs. Haussler [18] presented an elegant proof of this inequality using the shifting (push-down) operation. 1-Inclusion graphs have many other applications in computational learning theory, for example, in sample compression schemes [21]. They are exactly the induced subgraphs of hypercubes and in graph theory they have been studied under the name of *cubical graphs* [14]. Finding a densest n-vertex subgraph of the hypercube Q_m (i.e., an n-vertex subgraph G of Q_m with the maximum number of edges) is equivalent to finding an n-vertex subgraph G of Q_m with the smallest edge-boundary (the number of edges of Q_m running between V and its complement in Q_m). This is the classical *edge-isoperimetric problem* for hypercubes [3, 17]. Harper [16] nicely characterized the solutions of this problem: for any n, this is the subgraph of the hypercube induced by the initial segment of length n of the *lexicographic numbering* of the vertices of the hypercube. One elegant way of proving this result is using compression [17].

Generalizing the density inequality $\frac{|E|}{|V|} \leq$ VC-dim($\mathcal{S}$) of [18, 19] to more general classes of graphs is an interesting and important problem. In the current paper, we present a density result for 1,2-inclusion graphs $G_{1,2}(\mathcal{S})$ of arbitrary set families $\mathcal{S}$. The 1,2-inclusion graphs are the subgraphs of the square Q_m^2 of the hypercube Q_m and they are exactly the subgraphs of the halved cube $\frac{1}{2}Q_{m+1}$ (Johnson graphs and their subgraphs constitute an important subclass). Since 1,2-inclusion graphs may contain arbitrary large cliques for constant VC-dimension, we have to adapt the definition of classical VC-dimension to capture this phenomenon. For this purpose, we introduce the notion of *clique-VC-dimension* cVC-dim$^*(\mathcal{S})$ of any set family $\mathcal{S}$. Here is the main result of the paper:

Theorem 1. *Let $\mathcal{S}$ be an arbitrary set family of 2^X with $|X| = m$, let $d = \text{cVC-dim}^*(\mathcal{S})$ be the clique-VC-dimension of $\mathcal{S}$ and $G_{1,2}(\mathcal{S}) = (V, E)$ be the 1,2-inclusion graph of $\mathcal{S}$. Then $\frac{|E|}{|V|} \leq \binom{d}{2}$.*

2. Related work

2.1. **Haussler's proof of the inequality $\frac{|E|}{|V|} \leq \text{VC-dim}(\mathcal{S})$.** We briefly review the notion of VC-dimension and the shifting method of [18] of proving the inequality $\frac{|E|}{|V|} \leq \text{VC-dim}(\mathcal{S})$ (the original proof of [19] was by induction on the number of sets). In the same vein, see Harper's proof [17, Chapter 3] of the isoperimetric inequality via compression. We will use the shifting method in the proof of Theorem 1.

Let $\mathcal{S}$ be a family of subsets of a set $X = \{e_1, \dots, e_m\}$; $\mathcal{S}$ can be viewed as a subset of vertices of the m-dimensional hypercube Q_m. Denote by $G_1(\mathcal{S})$ the subgraph of Q_m induced by the vertices of Q_m corresponding to the sets of $\mathcal{S}$; $G_1(\mathcal{S})$ is called the *1-inclusion graph* of $\mathcal{S}$ [18,19]. Vice-versa, for any subgraph G of Q_m there exists a family of subsets $\mathcal{S}$ of 2^X such that G is the 1-inclusion graph of $\mathcal{S}$. A subset Y of X is *shattered* by $\mathcal{S}$ if for all $Y' \subseteq Y$ there exists $S \in \mathcal{S}$ such that $S \cap Y = Y'$. The *Vapnik-Chervonenkis's dimension* [28] $\text{VC-dim}(\mathcal{S})$ of $\mathcal{S}$ is the cardinality of the largest subset of X shattered by $\mathcal{S}$.

Theorem 2 ([18,19])**.** *If $G := G_1(\mathcal{S}) = (V, E)$ is the 1-inclusion graph of a set family $\mathcal{S} \subseteq 2^X$ with VC-dimension $\text{VC-dim}(\mathcal{S}) = d$, then $\frac{|E|}{|V|} \leq d$.*

For a set family $\mathcal{S} \subseteq 2^X$, the *shifting (push down or stabilization) operation* φ_e with respect to an element $e \in X$ replaces every set S of $\mathcal{S}$ such that $S \setminus \{e\} \notin \mathcal{S}$ by the set $S \setminus \{e\}$. Denote by $\varphi_e(\mathcal{S})$ the resulting set family and by $G' = G_1(\varphi_e(\mathcal{S})) = (V', E')$ the 1-inclusion graph of $\varphi_e(\mathcal{S})$. Haussler [18] proved that the shifting map φ_e has the following properties:

(1) φ_e is bijective on the vertex-sets: $|V| = |V'|$,
(2) φ_e is increasing the number of edges: $|E| \leq |E'|$,
(3) φ_e is decreasing the VC-dimension: $\text{VC-dim}(\mathcal{S}) \geq \text{VC-dim}(\varphi_e(\mathcal{S}))$.

Harper [17, p.28] called *Steiner operations* the set-maps $\varphi : 2^X \to 2^X$ satisfying (1), (2), and the following condition:

(4) $S \subseteq T$ implies $\varphi(S) \subseteq \varphi(T)$.

He proved that the compression operation defined in [17, Subsection 3.3] is a Steiner operation. Note that φ_e satisfies (4) (but is defined only on $\mathcal{S}$).

After a finite sequence of shiftings, any set family $\mathcal{S}$ can be transformed into a set family $\mathcal{S}^*$, such that $\varphi_e(\mathcal{S}^*) = \mathcal{S}^*$ holds for any $e \in X$. The resulting set family $\mathcal{S}^*$, a *complete shifting* of $\mathcal{S}$, is *downward closed* (i.e., is a *simplicial complex*). Consequently, the 1-inclusion graph $G_1(\mathcal{S}^*)$ of $\mathcal{S}^*$ is a *bouquet of cubes*, i.e., a union of subcubes of Q_m with a common origin $\varnothing$. Let $G^* = G_1(\mathcal{S}^*) = (V^*, E^*)$ and $d^* = \text{VC-dim}(\mathcal{S}^*)$. Since all shiftings satisfy the conditions (1)-(3), we conclude that $|V^*| = |V|$, $|E^*| \geq |E|$, and $d^* \leq d$. Therefore, to prove the inequality $\frac{|E|}{|V|} \leq d$ it suffices to show that $\frac{|E^*|}{|V^*|} \leq d^*$. Haussler deduced it from Sauer's lemma [26], however it is easy to prove this inequality directly, by bounding the degeneracy of G^*. Indeed, let v_0 be the vertex of G^* corresponding to the origin $\varnothing$ and let v be a furthest from v_0 vertex of G^*. Then v_0 and v span a maximal cube of G^* (of dimension $\leq d^*$) and v belongs only to this maximal cube of G^*. Therefore, if we remove v from G^*, we will also remove at most d^* edges of G^* and the resulting graph will be again a bouquet of cubes $G^- = (V^-, E^-)$ with one less vertex and dimension $\leq d^*$. Therefore, we can apply the induction hypothesis to this bouquet G^- and deduce that $|E^-| \leq |V^-|d^*$. Consequently, $|E^*| \leq d^* + |E^-| \leq d^* + (|V^*| - 1)d^* = |V^*|d^*$.

To extend Haussler's proof to subgraphs of halved cubes (and, equivalently, to subgraphs of squares of cubes), we need to appropriately define the shifting operation and the notion of VC-dimension, that satisfy the conditions (1)-(3). Additionally, the degeneracy of the 1,2-inclusion graph of the final shifted family must be bounded by a function of the VC-dimension. We will use the shifting operation with respect to pairs of elements (and not to single elements) and the notion of clique-VC-dimension instead of VC-dimension.

2.2. **Other results.** The inequality of Haussler et al. [19] as well as the notion of VC-dimension and Sauer lemma have been subsequently extended to subgraphs of Hamming graphs (i.e., from binary alphabets to arbitrary alphabets); see [20, 23–25]. Cesa-Bianchi and Haussler [6] presented a graph-theoretical generalization of the Sauer lemma for the m-fold $F^m = F \times \cdots \times F$ Cartesian products of arbitrary undirected graphs F. In [9], we defined a notion of VC-dimension for subgraphs of Cartesian products of arbitrary connected graphs (hypercubes are Cartesian products of K_2) and we established a density result $\frac{|E|}{|V|} \leq \text{VC-dim}(G) \cdot \alpha(H)$ for subgraphs G of Cartesian products of graphs not containing a fixed graph H as a minor ($\alpha(H)$ is a constant such that any graph not containing H as a minor has density at most $\alpha(H)$; it is well known [12] that if $r := |V(H)|$, then $\alpha(H) \leq cr\sqrt{\log r}$ for a universal constant c).

For edge- and vertex-isoperimetric problems in Johnson graphs (which are still open problems), some authors [1,11] used a natural *pushing to the left* (or *switching*, or *shifting*) operation. Let $\mathcal{S}$ consists only of sets of size r. Given an arbitrary total order $e_1, \ldots, e_m$ of the elements of X and two elements $e_i < e_j$, in the pushing to the left of $\mathcal{S}$ with respect to the pair e_i, e_j each set S of $\mathcal{S}$ containing e_j and not containing e_i is replaced by the set $S \setminus \{e_j\} \cup \{e_i\}$ if $S \setminus \{e_j\} \cup \{e_i\} \notin \mathcal{S}$. This operation preserves the size of $\mathcal{S}$, the cardinality r of the sets and do not decrease the number of edges, but the degeneracy of the final graph is not easy to bound.

Bousquet and Thomassé [4] defined the notions of 2-shattering and 2VC-dimension and established the Erdös-Pósa property for the families of balls of fixed radius in graphs with bounded 2VC-dimension. These notions have some similarity with our concepts of c-shattering and clique-VC-dimension because they concern shattering not of all subsets but only of a certain pattern of subsets (of all pairs). Recall from [4] that a set family $\mathcal{S}$ *2-shatters* a set Y if for any 2-set $\{e_i, e_j\}$ of Y there exists $S \in \mathcal{S}$ such that $Y \cap S = \{e_i, e_j\}$; the *2VC-dimension* of $\mathcal{S}$ is the maximum size of a 2-shattered set.

Halved cubes and Johnson graphs host several important classes of graphs occurring from metric graph theory [2]: basis graphs of matroids are isometric subgraphs of Johnson graphs [22] and basis graphs of even Δ-matroids are isometric subgraphs of halved cubes [7]. More general classes are the graphs isometrically embeddable into halved cubes and Johnson graphs. Similarly to Djoković's characterization of isometric subgraphs of hypercubes [13], isometric subgraphs of Johnson graphs have been characterized in [8] (the problem of characterizing isometric subgraphs of halved cubes has been raised in [10] and is still open). Shpectorov [27] proved that the graphs admitting an isometric embedding into an ℓ_1-space are exactly the graphs which admit a scale embedding into a hypercube and he proved that such graphs are exactly the graphs which are isometric subgraphs of Cartesian products of octahedra and of isometric subgraphs of halved cubes. For a presentation of most of these results, see the book by Deza and Laurent [10].

3. Preliminaries

3.1. **Degeneracy.** All graphs $G = (V, E)$ occurring in this note are finite, undirected, and simple. The *degeneracy* of G is the minimal k such that there exists a total order $v_1, \ldots, v_n$ of vertices of G such that each vertex v_i has degree at most k in the subgraph of G induced by

$v_i, v_{i+1}, \ldots, v_n$. It is well known and it can be easily shown that the degeneracy of every graph $G = (V, E)$ upper bounds the ratio $\frac{|E|}{|V|}$.

3.2. **Squares of hypercubes, halved cubes, and Johnson graphs.** The *m-dimensional hypercube* Q_m is the graph having all 2^m subsets of a set $X = \{e_1, \ldots, e_m\}$ as the vertex-set and two sets A, B are adjacent in Q_m iff $|A \triangle B| = 1$. The *halved cube* $\frac{1}{2}Q_m$ [5, 10] has the subsets of X of even cardinality as vertices and two such vertices A, B are adjacent in $\frac{1}{2}Q_m$ iff $|A \triangle B| = 2$ (one can also define halved cubes for subsets of odd size). Equivalently, the halved cube $\frac{1}{2}Q_m$ is the *square* Q^2_{m-1} of the hypercube Q_{m-1}, i.e., the graph formed by connecting pairs of vertices of Q_{m-1} whose distance is at most two in Q_{m-1}. For an integer $r > 0$, the *Johnson graph* $J(r, m)$ [5, 10] has the subsets of X of size r as vertices and two such vertices A, B are adjacent in $J(r, m)$ iff $|A \triangle B| = 2$. All Johnson graphs $J(r, m)$ are (isometric) subgraphs of the corresponding halved cube $\frac{1}{2}Q_m$. Notice also that the halved cube $\frac{1}{2}Q_m$ and the Johnson graph $J(r, m)$ are scale 2 embedded in the hypercube Q_m.

Let $\mathcal{S}$ be a family of subsets of a set $X = \{e_1, \ldots, e_m\}$. The *1,2-inclusion graph* $G_{1,2}(\mathcal{S})$ of $\mathcal{S}$ is the graph having $\mathcal{S}$ as the vertex-set and in which two vertices A and B are adjacent iff $1 \leq |A \Delta B| \leq 2$, i.e., $G_{1,2}(\mathcal{S})$ is the subgraph of the square Q^2_m of Q_m induced by $\mathcal{S}$. The graph $G_{1,2}(\mathcal{S})$ comprises all edges of the 1-inclusion graph $G_1(\mathcal{S})$ of $\mathcal{S}$ and of the subgraphs of the halved cubes induced by even and odd sets of $\mathcal{S}$. The latter edges of $G_{1,2}(\mathcal{S})$ are of two types: *vertical edges* SS' arise from sets S, S' such that $|S| = |S'| + 2$ or $|S'| = |S| + 2$ and *horizontal edges* SS' arise from sets S, S' such that $|S| = |S'|$.

If all sets of $\mathcal{S}$ have even cardinality, then we will call $\mathcal{S}$ an *even set family*; in this case, the 1,2-inclusion graph $G_{1,2}(\mathcal{S})$ coincides with the subgraph of the halved cube $\frac{1}{2}Q_m$ induced by $\mathcal{S}$. Since Q^2_m is isomorphic to $\frac{1}{2}Q_{m+1}$, any 1,2-inclusion graph is an induced subgraph of a halved cube. More precisely, any set family $\mathcal{S}$ of X can be lifted to an even set family $\mathcal{S}^+$ of $X \cup \{e_{m+1}\}$ in such a way that the 1,2-inclusion graphs of $\mathcal{S}$ and $\mathcal{S}^+$ are isomorphic: $\mathcal{S}^+$ consists of all sets of even size of $\mathcal{S}$ and of all sets of odd size of $\mathcal{S}$ to which the element e_{m+1} was added. The proof of the following lemma is straightforward:

Lemma 1. *For any set family $\mathcal{S}$, the lifted family $\mathcal{S}^+$ is an even set family and the 1,2-inclusion graphs $G_{1,2}(\mathcal{S})$ and $G_{1,2}(\mathcal{S}^+)$ are isomorphic.*

3.3. **Pointed set families and pointed cliques.** We will call a set family $\mathcal{S}$ a *pointed set family* if $\varnothing \in \mathcal{S}$. Any set family $\mathcal{S}$ can be transformed into a pointed set family by the operation of twisting. For a set $A \in \mathcal{S}$, let $\mathcal{S} \triangle A := \{S \triangle A : S \in \mathcal{S}\}$ and say that $\mathcal{S} \triangle A$ is obtained from $\mathcal{S}$ by applying a *twisting* with respect to A. Note that a twisting is a bijection between $\mathcal{S}$ and $\mathcal{S} \triangle A$ mapping the set A to $\varnothing$ (and therefore $\mathcal{S} \triangle A$ is a pointed set family). Notice that any twisting of an even set family $\mathcal{S}$ is an even set family. As before, let $G_1(\mathcal{S})$ denote the 1-inclusion graph of $\mathcal{S}$. The following properties of twisting are well-known and easy to prove:

Lemma 2. *For any $\mathcal{S} \subseteq 2^X$ and any $A \subseteq X$, $G_1(\mathcal{S} \triangle A) \backsimeq G_1(\mathcal{S})$ and* VC-dim$(\mathcal{S} \triangle A) =$ VC-dim$(\mathcal{S})$.

Analogously to the proof of the first assertion of Lemma 2, one can easily show that:

Lemma 3. *For any set family $\mathcal{S} \subseteq 2^X$ and any $A \subseteq X$, $G_{1,2}(\mathcal{S} \triangle A) \backsimeq G_{1,2}(\mathcal{S})$.*

We will say that a clique $\mathcal{C}$ of $\frac{1}{2}Q_m$ is a *pointed clique* if $\mathcal{C}$ is a pointed set family.

Lemma 4. *By a twisting, any clique $\mathcal{C}$ of $\frac{1}{2}Q_m$ can be transformed into a pointed clique.*

Proof. Let $\mathcal{C}$ be a clique of $\frac{1}{2}Q_m$. Let A be a set of maximal size which is a vertex of $\mathcal{C}$. Then the twisting of $\mathcal{C}$ with respect to A maps $\mathcal{C}$ into a pointed clique $\mathcal{C} \triangle A$ of $\frac{1}{2}Q_m$: indeed, if C', C''

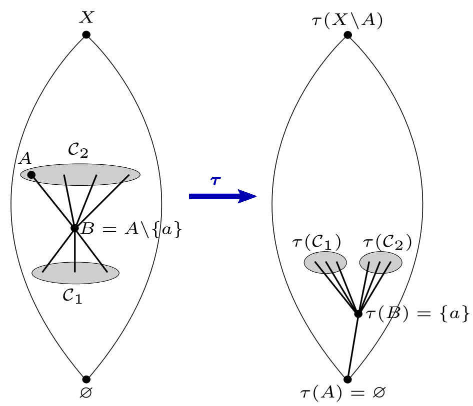

FIGURE 1. A twisting mapping $\tau : S \mapsto S \triangle A$ of a clique to a pointed clique.

are two vertices of $\mathcal{C}$, then $|(C' \triangle A) \triangle (C'' \triangle A)| = |C' \triangle C''| = 2$. Since $A \Delta A = \varnothing$, $\mathcal{C} \triangle A$ is a pointed clique (for an illustration, see Fig. 1). □

We describe now the structure of pointed cliques in halved cubes.

Lemma 5. *Any pointed maximal clique $\mathcal{C}$ of a halved cube $\frac{1}{2}Q_m$ is (a) a sporadic 4-clique of the form $\{\varnothing, \{e_i, e_j\}, \{e_i, e_k\}, \{e_j, e_k\}\}$ for arbitrary elements $e_i, e_j, e_k \in X$, or (b) a clique of size m of the form $\{\varnothing\} \cup \{\{e_i, e_j\} : e_j \in X \setminus \{e_i\}\}$ for an arbitrary but fixed element $e_i \in X$.*

Proof. Since $\mathcal{C}$ is a pointed clique, $\varnothing$ is a vertex of $\mathcal{C}$, denote it C_0. All other neighbors of C_0 in $\frac{1}{2}Q_m$ are sets of the form $\{e_i, e_j\}$ with $e_i, e_j \in X$, i.e., the neighborhood of C_0 in the halved cube $\frac{1}{2}Q_m$ is the line-graph of the complete graph K_m having X as the vertex-set. In particular, the clique $\mathcal{C}_0 := \mathcal{C} \setminus \{C_0\}$ corresponds to a set of pairwise incident edges of K_m. It can be easily seen that this set of edges defines either a triangle or a star of K_m. Indeed, pick an edge e_ie_j of K_m corresponding to a pair $\{e_i, e_j\} \in \mathcal{C}_0$. If the respective set of edges is not a star, then necessarily $\mathcal{C}_0$ contains two pairs of the form $\{e_i, e_k\}$ and $\{e_j, e_l\}$, both different from $\{e_i, e_j\}$. But then $k = l$, otherwise the edges e_ie_k and e_je_l would not be incident. Thus $\mathcal{C}_0$ contains the three pairs $\{e_i, e_j\}, \{e_i, e_k\}$, and $\{e_j, e_k\}$. If $\mathcal{C}_0$ contains yet another pair, then this pair will be necessarily disjoint from one of the three previous pairs, a contradiction. Thus in this case, $\mathcal{C} = \{\varnothing, \{e_i, e_j\}, \{e_i, e_k\}, \{e_j, e_k\}\}$. Otherwise, if the respective set of edges is a star with center e_i, then $\mathcal{C}_0$ is a clique of size $m - 1$ of the form $\{\{e_i, e_j\} : e_j \in X \setminus \{e_i\}\}$. □

4. The clique-VC-dimension

As we noticed above, the classical VC-dimension of set families cannot be used to bound the density of their 1,2-inclusion graphs. Indeed, the 1,2-inclusion graph of the set family $\mathcal{S}_0 := \{\{e_j\} : e_j \in X\}$ is a complete graph, while the VC-dimension of $\mathcal{S}_0$ is 1 (notice also that the 2VC-dimension of $\mathcal{S}_0$ is 0).

We will define a notion that is more appropriate for this purpose, which we will call clique-VC-dimension. The idea is to use the form of pointed cliques of $\frac{1}{2}Q_m$ established above and to shatter them. In view of Lemma 1, it suffices to define the clique-VC-dimension for even set families. First we present a generalized definition of classical shattering.

Let $X = \{e_1, \ldots, e_m\}$ and $\mathcal{S} \subseteq 2^X$. Let Y be a subset of X. Denote by $Q[Y]$ the subcube of Q_m consisting of all subsets of Y. Analogously, for two sets Y' and Y such that $Y' \subset Y$, denote by $Q[Y', Y]$ the smallest subcube of Q_m containing the sets Y' and Y: $Q[Y', Y] = \{Z \subset X :$

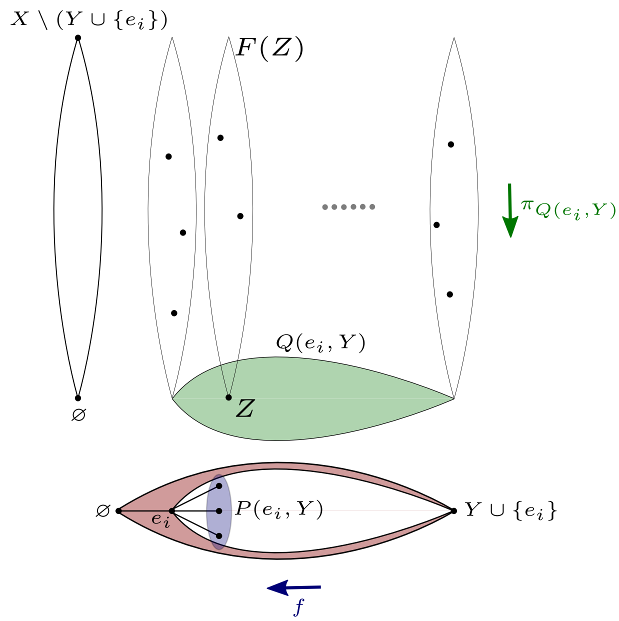

FIGURE 2. Example of a c-shattered pair (e_i, Y). $F(Z)$ is the fiber of Z in $Q(e_i, Y)$. The sets of $\mathcal{S}$ (black points) in the fibers of the sets of $Q(e_i, Y)$ are projected on $Q(e_i, Y)$ (in green). The vertices in $Q(e_i, Y)$ are then mapped to $P(e_i, Y)$ (in blue) by the c-shattering function f. The remaining vertices of $Q[\varnothing, Y \cup \{e_i\}]$ (in red) are "not used" for shattering.

$Y' \subseteq Z \subseteq Y\}$. In particular, $Q[Y] = Q[\varnothing, Y]$. For a vertex Z of $Q[Y', Y]$, call

$$F(Z) := \{Z \cup Z' : Z' \subseteq X \setminus Y\}$$

the *fiber* of Z with respect to the cube $Q[Y', Y]$. Let

$$\pi_{Q[Y',Y]}(\mathcal{S}) := \{Z \in Q[Y', Y] : F(Z) \cap \mathcal{S} \neq \varnothing\}$$

denote the *projection* of the set family $\mathcal{S}$ on $Q[Y', Y]$. Then the cube $Q[Y', Y]$ with $Y' \subseteq Y$ is shattered by $\mathcal{S}$ if $\pi_{Q[Y',Y]}(\mathcal{S}) = Q[Y', Y]$, i.e., for any $Y' \subseteq Z \subseteq Y$ the fiber $F(Z)$ contains a set of $\mathcal{S}$ (see Fig. 2). In particular, a subset Y is shattered by $\mathcal{S}$ iff $\pi_{Q[Y]}(\mathcal{S}) = Q[Y]$.

4.1. **The clique-VC-dimension of pointed even set families.** Let $\mathcal{S}$ be a pointed even set family of 2^X, i.e., a set family in which all sets have even size and $\varnothing \in \mathcal{S}$. Let Y be a subset of X and let e_i be an element of X not belonging to Y. Denote by $P(e_i, Y)$ the set of all 2-*sets*, i.e., pairs of the form $\{e_i, e_j\}$ with $e_j \in Y$. Then $Q[\{e_i\}, Y \cup \{e_i\}]$ is the smallest subcube of Q_m containing e_i and all the 2-sets of $P(e_i, Y)$. For simplicity, we will denote this cube by $Q(e_i, Y)$.

We will say that a pair (e_i, Y) with $Y \subset X$ and $e_i \notin Y$ is *c-shattered* by $\mathcal{S}$ if there exists a surjective function $f : \pi_{Q(e_i,Y)}(\mathcal{S}) \to P(e_i, Y)$ such that for any $S \in \pi_{Q(e_i,Y)}(\mathcal{S})$ the inclusion $f(S) \subseteq S$ holds. In other words, (e_i, Y) is c-shattered by $\mathcal{S}$ if each 2-set $\{e_i, e_j\} \in P(e_i, Y)$ admits an extension $S_j \in \pi_{Q(e_i,Y)}(\mathcal{S})$ such that $\{e_i, e_j\} \subseteq S_j$ and for any two 2-sets $\{e_i, e_j\}, \{e_i, e_{j'}\} \in P(e_i, Y)$ the sets S_j and $S_{j'}$ are distinct. Since $\varnothing \in \mathcal{S}$, the empty set $\varnothing$ is always shattered by $\mathcal{S}$.

For a pointed even set family $\mathcal{S}$, the *clique-VC-dimension* is

$$\text{cVC-dim}(\mathcal{S}) := \max\{|Y| + 1 : Y \subset X \text{ and } \exists e_i \in X \setminus Y \text{ such that } (e_i, Y) \text{ is c-shattered by } \mathcal{S}\}.$$

We continue with some simple examples of clique-VC-dimension:

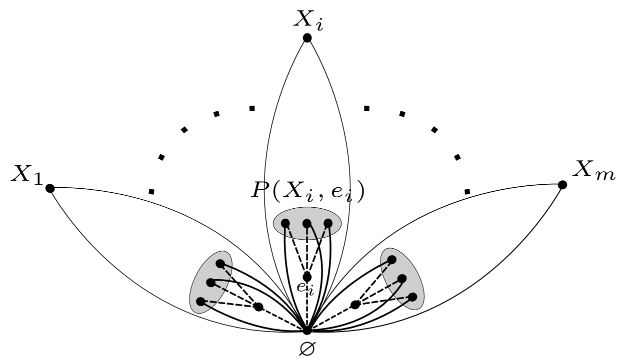

FIGURE 3. Illustration of Example 3

Example 1. For set family $\mathcal{S}_0 = \{\{e_j\} : e_j \in X\}$ introduced above, let $\mathcal{S}_0^+ = \{\{e_j, e_{m+1}\} : e_j \in X\}$ be the lifting of $\mathcal{S}_0$ to an even set family. For an arbitrary (but fixed) element e_i, let $\mathcal{S}_1 := \{\varnothing\} \cup \{\{e_i, e_j\} : e_j \neq e_i\}$. Then $\mathcal{S}_1$ coincides with $\mathcal{S}_0 \Delta \{e_i\}$ and with $\mathcal{S}_0^+ \Delta \{e_i, e_{m+1}\}$. $\mathcal{S}_1$ is an even set family, its 1,2-inclusion graph is a pointed clique, and cVC-dim$(\mathcal{S}_1) = |X| = m$.

Example 2. Let $\mathcal{S}_2 = \{\varnothing, \{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}\}$ be the sporadic 4-clique from Lemma 5. In this case, one can c-shatter any two of the pairs $\{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}$ but not all three. This shows that cVC-dim$(\mathcal{S}_2) = 2 + 1 = 3$.

Example 3. For arbitrary even integers m and k, Let X be a ground set of size $m+km$ which is the disjoint union of $m+1$ sets $X_0, X_1, \ldots, X_m$, where $X_0 = \{e_1, \ldots, e_m\}$ and $X_i = \{e_{i1}, \ldots, e_{ik}\}$ for each $i = 1, \ldots, m$. Let $\mathcal{S}_3$ be the pointed even set family consisting of the empty set $\varnothing$, the set X, and for each $i = 1, \ldots, m$ of all the 2-sets of $P(e_i, X_i) = \{\{e_i, e_{i1}\}, \ldots, \{e_i, e_{ik}\}\}$. Then $G_{1,2}(\mathcal{S}_3)$ consists of an isolated vertex X and m maximal cliques $\mathcal{C}_i := P(e_i, X_i) \cup \{\varnothing\}$ of size $k+1$ and these cliques pairwise intersect in a single vertex $\varnothing$. We assert that cVC-dim$(\mathcal{S}_3) = k + 2$. Indeed, let Y be the set consisting of X_i for a given $i \in \{1, \ldots, m\}$ plus the singleton $\{e_{(i+1)1}\}$. Then the pair (e_i, Y) is c-shattered by $\mathcal{S}_3$. The c-shattering map $f : \pi_{Q(e_i,Y)}(\mathcal{S}_3) \to P(e_i, Y)$ is defined as follows: every 2-set of $P(e_i, X_i) \subset Q(e_i, Y)$ is in $\mathcal{S}_3$ and is thus mapped to itself, $X \cap (Y \cup \{e_i\}) = Y \cup \{e_i\}$ is an extension of the remaining 2-set $\{e_i, e_{(i+1)1}\}$ in $Q(e_i, Y)$ and thus $f(Y \cup \{e_i\}) := \{e_i, e_{(i+1)1}\}$. Since $|Y| = k + 1$, we showed that cVC-dim$(\mathcal{S}_3) \geq k + 2$. On the other hand, cVC-dim$(\mathcal{S}_3) \leq k + 2$ because every element e from X is in at most $k + 1$ sets of $\mathcal{S}_3$. Therefore, cVC-dim$(\mathcal{S}_3) = k + 2$.

4.2. **The clique-VC-dimension of even and arbitrary set families.** The *clique-VC-dimension* cVC-dim$^*(\mathcal{S})$ of an even set family $\mathcal{S}$ is the minimum of the clique-VC-dimensions of the pointed even set families $\mathcal{S} \triangle A$ for $A \in \mathcal{S}$:

$$\text{cVC-dim}^*(\mathcal{S}) := \min\{\text{cVC-dim}(\mathcal{S} \triangle A) : A \in \mathcal{S}\}.$$

The *clique-VC-dimension* cVC-dim$^*(\mathcal{S})$ of an arbitrary set family $\mathcal{S}$ is the clique-VC-dimension of its lifting $\mathcal{S}^+$.

Remark 1. A simple analysis shows that for the even set families from Examples 1-3, we have cVC-dim$^*(\mathcal{S}_1) = m$, cVC-dim$^*(\mathcal{S}_2) = 3$, and cVC-dim$^*(\mathcal{S}_3) = k + 2$.

Remark 2. In fact, the set family $\mathcal{S}_3$ shows that the maximum degree of a 1,2-inclusion graph $G_{1,2}(\mathcal{S})$ of an even set family $\mathcal{S}$ can be arbitrarily larger than cVC-dim$^*(\mathcal{S})$. Indeed, $\varnothing$ is the vertex of maximum degree of $G_{1,2}(\mathcal{S}_3)$ and its degree is km.

The family $\mathcal{S}_3$ also explains why in the definition of the clique-VC-dimension of $\mathcal{S}$ we take the minimum over all $\mathcal{S}\triangle A, A \in \mathcal{S}$. Consider the twisting of $\mathcal{S}_3$ with respect to the set $X \in \mathcal{S}_3$. Then one can see that cVC-dim$(\mathcal{S}_3\triangle X) \geq (m-1)k+1$. Indeed, $\mathcal{S}_3\Delta X = \{\varnothing, X\} \cup (\bigcup_{(i,j)\in\{1,\dots,m\}\times\{1,\dots,k\}}\{X\backslash\{e_i,e_{ij}\}\})$. Let $Y := \{e_{ij} : i \in \{1,\dots,m-1\}$ and $j \in \{1,\dots,k\}\}$. We assert that (e_1, Y) is c-shattered by $\mathcal{S}_3\Delta X$. We set $\mathcal{S}_3' := \pi_{Q(e_1,Y)}(\mathcal{S}_3\Delta X)$, $S_{ij} := X \setminus \{e_i, e_{ij}\}$, and $S'_{ij} := \pi_{Q(e_1,Y)}(S_{ij})$. Let $f : \mathcal{S}_3' \to P(e_1, Y)$ be such that for all $i \in \{2,\dots,m\}$ and $j \in \{1,\dots,k\}$, we have $f(S'_{ij}) = \{e_1, e_{(i-1)j}\}$. Clearly, every $\{e_1, e_{(i-1)j}\}$ has an extension S'_{ij} with a non-empty fiber ($S_{ij} \in F(S'_{ij})$), and for all $S_{rl} \neq S_{ij}$, we have $S'_{rl} \neq S'_{ij}$, hence f is a surjection. Therefore, (e_1, Y) is c-shattered. Since $|Y| = (m-1)k$, whence cVC-dim$(\mathcal{S}_3\triangle X) \geq (m-1)k+1$.

5. Proof of Theorem 1

After the preparatory work done in previous three subsections, here we present the proof of our main result. We start the proof by defining the double shifting (d-shifting) as an adaptation of the shifting to pointed even families. We show that, similarly to classical shifting operation, d-shifting satisfies the conditions (1)-(3) and that the result of a complete sequence of d-shiftings is a bouquet of halved cubes (which is a particular pointed even set family). We show that the degeneracy of the 1,2-inclusion graph of such a bouquet $\mathcal{B}$ is bounded by $\binom{d}{2}$, where $d :=$ cVC-dim$(\mathcal{B})$. We conclude the proof of the theorem by considering arbitrary even set families $\mathcal{S}$ and applying the previous arguments to the pointed family $\mathcal{S}\Delta A$, where A is a set of $\mathcal{S}$ such that cVC-dim$(\mathcal{S}\Delta A) =$ cVC-dim$^*(\mathcal{S})$.

5.1. **Double shiftings of pointed even families.** For a pointed even set family $\mathcal{S} \subseteq 2^X$, the *double shifting* (*d-shifting* for short) with respect to a 2-set $\{e_i, e_j\} \subseteq X$ is a map $\varphi_{ij} : \mathcal{S} \to 2^X$ which replaces every set S of $\mathcal{S}$ such that $\{e_i, e_j\} \subseteq S$ and $S \setminus \{e_i, e_j\} \notin \mathcal{S}$ by the set $S \setminus \{e_i, e_j\}$:

$$\begin{aligned}\varphi_{ij} : \mathcal{S} &\to 2^X\\ S &\mapsto \begin{cases} S \setminus \{e_i, e_j\}, & \text{if } \{e_i, e_j\} \subseteq S \text{ and } S \setminus \{e_i, e_j\} \notin \mathcal{S}\\ S, & \text{otherwise.}\end{cases}\end{aligned}.$$

Proposition 1. *Let $\mathcal{S} \subseteq 2^X$ be a pointed even set family, let $\{e_i, e_j\} \subseteq X$ be a 2-set, and let $G_{1,2}(\mathcal{S}) = G = (V, E)$ and $G_{1,2}(\varphi_{ij}(\mathcal{S})) = G' = (V', E')$ be the subgraphs of the halved cube induced by $\mathcal{S}$ and $\varphi_{ij}(\mathcal{S})$, respectively. Then $|V| = |V'|$ and $|E| \leq |E'|$ hold.*

Proof. The fact that a d-shifting φ_{ij} preserves the number of vertices of an induced subgraph of halved cube immediately follows from the definition. Therefore we only need to show that φ_{ij} cannot decrease the number of edges, i.e., that there exists an injective map $\psi_{ij} : E \to E'$. We will call an edge SS' of G *stable* if $\varphi_{ij}(S) = S$ and $\varphi_{ij}(S') = S'$ hold and *shiftable* otherwise. For each stable edge SS' we will set $\psi_{ij}(SS') := SS'$.

Now, pick any shiftable edge SS' of E. Notice that in this case $\{e_i, e_j\} \subseteq S$ or $\{e_i, e_j\} \subseteq S'$. To define $\psi_{ij}(SS')$, we distinguish two cases depending on whether $\{e_i, e_j\}$ is a subset of only one of the sets S, S' or of both of them.

Case 1′. $\{e_i, e_j\} \subseteq S$ and $\{e_i, e_j\} \nsubseteq S'$ (the case $\{e_i, e_j\} \subseteq S'$ and $\{e_i, e_j\} \nsubseteq S$ is similar).

Since $\{e_i, e_j\} \nsubseteq S'$, necessarily $\varphi_{ij}(S') = S'$. Since SS' is shiftable, $\varphi_{ij}(S) \neq S$, i.e., $\varphi_{ij}(S) = S \setminus \{e_i, e_j\} =: Z$. We consider two cases depending on whether one of the elements e_i or e_j belongs to S' or not.

Subcase 1′.1. $e_i \in S'$ and $e_j \notin S'$ (the case $e_j \in S'$ and $e_i \notin S'$ is similar). In this case, there is an element $e_k \in X$ such that $S\Delta S' = \{e_j, e_k\}$. Observe that $S \nsubseteq S'$ since $e_j \notin S'$ and $e_j \in S$. Hence either $S' \subseteq S$ or there exists $A \subset X$ such that $S' = A \cup \{e_k\}$ and $S = A \cup \{e_j\}$. In the former case, we have $S = S' \cup \{e_j, e_k\}$, $Z = S' \cup \{e_k\} \setminus \{e_i\}$, and $Z\Delta S' = \{e_i, e_k\}$. In the later case, we have $Z = A \setminus \{e_i\}$ and $Z\Delta S' = \{e_i, e_k\}$. In both cases, $|Z\Delta S'| = 2$ and $ZS' \in E'$. We set $\psi_{ij}(SS') := ZS'$.

Subcase 1′.2. $e_i \notin S'$ and $e_j \notin S'$. Then $S\Delta S' = \{e_i, e_j\}$ and so $S \setminus \{e_i, e_j\} = Z = S'$. We obtain a contradiction that SS' is shiftable (i.e., $Z = \varphi_{ij}(S)$ cannot be in $\mathcal{S}$).

Case 2′. $\{e_i, e_j\} \subseteq S$ and $\{e_i, e_j\} \subseteq S'$.

Set $Z := S \setminus \{e_i, e_j\}$ and $Z' := S' \setminus \{e_i, e_j\}$. Then both sets Z, Z' belong to $\varphi_{ij}(\mathcal{S})$ and ZZ' defines an edge of G'. Since SS' is shiftable, at least one of the sets Z, Z' does not belong to $\mathcal{S}$.

Subcase 2′.1. $Z, Z' \notin \mathcal{S}$. Then $\varphi_{ij}(S) = Z$ and $\varphi_{ij}(S') = Z'$ and ZZ' is an edge of G'. In this case, we set $\psi_{ij}(SS') := ZZ'$.

Subcase 2′.2. $Z \in \mathcal{S}$ and $Z' \notin \mathcal{S}$ (the case $Z \notin \mathcal{S}$ and $Z' \in \mathcal{S}$ is similar). Then $\varphi_{ij}(S) = S, \varphi_{ij}(S') = Z'$, and ZZ' is an edge of G' but not of G. In this case, we set $\psi_{ij}(SS') := ZZ'$.

It remains to show that the map $\psi_{ij} : E \to E'$ is injective. Suppose by way of contradiction that G' contains an edge ZZ' for which there exist two distinct edges SS' and CC' of G such that $\psi_{ij}(SS') = \psi_{ij}(CC') = ZZ'$. Since at least one of the edges SS' and CC' is different from ZZ', from the definition of d-shifting we conclude that ZZ' is not an edge of G, say $Z' \notin \mathcal{S}$. This also implies that SS' and CC' are shiftable edges of G.

Case 1″. $Z \notin \mathcal{S}$.

From the definition of the map ψ_{ij} and since $Z, Z' \notin \mathcal{S}$, both edges SS' and CC' are in Subcase 2′.1. This shows that $Z = S \setminus \{e_i, e_j\}, Z' = S' \setminus \{e_i, e_j\}$, and $Z = C \setminus \{e_i, e_j\}, Z' = C' \setminus \{e_i, e_j\}$, yielding $S = C$ and $S' = C'$, a contradiction.

Case 2″. $Z \in \mathcal{S}$.

After an appropriate renaming of the sets S, S' and C, C', we can suppose that $\varphi_{ij}(S) = \varphi_{ij}(C) = Z$ and $\varphi_{ij}(S') = \varphi_{ij}(C') = Z'$. Since $Z' \notin \mathcal{S}$, from the definition of the map ψ_{ij}, we deduce that $S' = Z' \cup \{e_i, e_j\} = C'$. On the other hand, since $Z \in \mathcal{S}$, we have either $S = C = Z$ which contradicts the choice of $SS' \neq CC'$, or $S \setminus \{e_i, e_j\} = C = Z$ (or the symmetric possibility $C \setminus \{e_i, e_j\} = S = Z$) which contradicts the fact that SS' (or CC') is shiftable.

This shows that the map $\psi_{ij} : E \to E'$ is injective, thus $|E| \leq |E'|$. □

Lemma 6. *If φ_{ij} is a d-shifting of a pointed even family $\mathcal{S} \subset 2^X$, then* cVC-dim$(\varphi_{ij}(\mathcal{S})) \leq$ cVC-dim$(\mathcal{S})$.

Proof. Let (e, Y) be c-shattered by $\mathcal{S}_{ij} := \varphi_{ij}(\mathcal{S})$ (recall that $Y \subset X$ and $e \notin Y$). Let $\mathcal{S}' := \pi_{Q(e,Y)}(\mathcal{S})$ and $\mathcal{S}'_{ij} := \pi_{Q(e,Y)}(\varphi_{ij}(\mathcal{S}))$. By definition of c-shattering, there exists a surjective function f associating every element of $\mathcal{S}'_{ij}$ to a 2-set $\{e, e'\} \in P(e, Y)$. We will define a surjective function g from $\mathcal{S}'$ to $\mathcal{S}'_{ij}$, and derive from f a c-shattering function $f' := f \circ g$ from $\mathcal{S}'$ to $P(e, Y)$. Let $S_{e'} \in \mathcal{S}'_{ij}$ be a set such that $f(S_{e'}) = \{e, e'\}$. If $S_{e'} \in \mathcal{S}'$, then the 2-set $\{e, e'\}$ also has an extension in $\mathcal{S}'$ and we can set $g(S_{e'}) := S_{e'}$. If $S_{e'} \notin \mathcal{S}'$, it means that there exists a set $S \in \mathcal{S}$ such that $S \neq \varphi_{ij}(S)$ and $\varphi_{ij}(S)$ is in the fiber $F(S_{e'})$ of $S_{e'}$ with respect to $\pi_{Q(e,Y)}$ in $\mathcal{S}_{ij}$. The set S is in the fiber $F(S')$ of some set $S' \in \mathcal{S}'$ with respect to $\pi_{Q(e,Y)}$. Since $\varphi_{ij}(S) \subseteq S$, we have $S_{e'} \subseteq S'$ and $S' \in \mathcal{S}'$ is an extension of the 2-set $\{e, e'\}$. We set $g(S') := S_{e'}$. Moreover, for every set $S' \in \mathcal{S}' \setminus \mathcal{S}'_{ij}$, there is a set $S \in F(S')$ such that $\varphi_{ij}(S) \neq S$. In this case, there is a set $S_{e'} \in \mathcal{S}'_{ij}$ such that $\varphi_{ij}(S) \in F(S_{e'})$. We set $g(S') := S_{e'}$. We have $S_{e'} \subseteq S'$ since $\varphi_{ij}(S) \subset S$.

The function g is surjective by definition and maps every set of $\mathcal{S}'$ either on itself or on a subset of it. Since f is a c-shattering function, so is $f' := f \circ g$ and (e, Y) is c-shattered by $\mathcal{S}$.

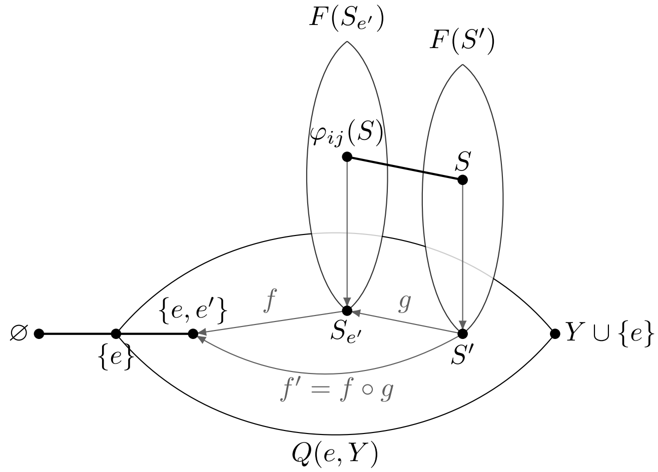

FIGURE 4. To the proof of Lemma 6.

Consequently, we have cVC-dim($\varphi_{ij}(\mathcal{S})$) $\leq$ cVC-dim($\mathcal{S}$) since every (e, Y) c-shattered by $\mathcal{S}_{ij}$ is also c-shattered by $\mathcal{S}$. $\square$

5.2. **Bouquets of halved cubes.** A *bouquet of cubes* (called usually a *downward closed family* or a *simplicial complex*) is a set family $\mathcal{B} \subseteq 2^X$ such that $S \in \mathcal{B}$ and $S' \subseteq S$ implies $S' \in \mathcal{B}$. Obviously $\mathcal{B}$ is a pointed family. Note that any bouquet of cubes $\mathcal{B}$ is the union of all cubes of the form $Q[\varnothing, S]$, where S is an inclusion-wise maximal subset of $\mathcal{B}$.

A *bouquet of halved cubes* is an even set family $\mathcal{B} \subseteq 2^X$ such that for any $S \in \mathcal{B}$, any subset S' of S of even size is included in $\mathcal{B}$. In other words, a bouquet of halved cubes $\mathcal{B}$ is the union of all halved cubes spanned by $\varnothing$ and inclusion-wise maximal subsets S of $\mathcal{B}$.

Lemma 7. *After a finite number of d-shiftings, any pointed even set family $\mathcal{S}$ of 2^X can be transformed into a bouquet of halved cubes.*

Proof. Let $\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, \ldots$ be a sequence of even set families such that $\mathcal{S}_0 = \mathcal{S}$ and, for any $i \geq 1$, $\mathcal{S}_i$ was obtained from $\mathcal{S}_{i-1}$ by a d-shifting and $\mathcal{S}_i \neq \mathcal{S}_{i-1}$. This sequence is necessarily finite because each d-shifting strictly decreases the sum of sizes of the sets in the family. Let $\mathcal{S}_r$ denote the last family in the sequence. This means that the d-shifting of $\mathcal{S}_r$ with respect to any pair of elements of X leads to the same set family $\mathcal{S}_r$. Therefore, for any set $S \in \mathcal{S}_r$ and for any pair $\{e_j, e_k\} \subseteq S$, the set $S \setminus \{e_j, e_k\}$ belongs to $\mathcal{S}_r$, i.e., $\mathcal{S}_r$ is a bouquet of halved cubes. $\square$

We continue with simple properties of bouquets of halved cubes.

Lemma 8. *Let $\mathcal{B} \subset 2^X$ be a bouquet of halved cubes of clique-VC-dimension $d :=$ cVC-dim($\mathcal{B}$). Then the following properties hold:*

(i) *for any element $e_i \in X$, $|\{\{e_i, e_j\} \in \mathcal{B} : e_j \in X \setminus \{e_i\}\}| \leq d - 1$;*
(ii) *if S is a set of $\mathcal{B}$, then $|S| \leq d$;*
(iii) *if S is a set of $\mathcal{B}$ maximal by inclusion, then $\mathcal{B} \setminus \{S\}$ is still a bouquet of halved cubes.*

Proof. The inequality $|\{\{e_i, e_j\} \in \mathcal{S} : e_j \in X \setminus \{e_i\}\}| \leq d-1$ directly follows from the definition of cVC-dim($\mathcal{B}$). The property (iii) immediately follows from the definition of a bouquet of halved cubes. To prove (ii), suppose by way of contradiction that $|S| > d$. Since $\mathcal{B}$ is a bouquet of halved cubes, every subset of S of even cardinality belongs to $\mathcal{B}$. Therefore, if we pick any $e \in S$ and if we set $Y := S \setminus \{e\}$, then all the 2-sets of the form $\{e, e'\}$ with $e' \in Y$ are subsets of S, and thus are sets of $\mathcal{B}$. Consequently (e, Y) is c-shattered by $\mathcal{B}$. Since $|Y| = |S| - 1 > d - 1$, this contradicts the assumption that $d = \text{cVC-dim}(\mathcal{B})$. $\square$

5.3. **Degeneracy of bouquets of halved cubes.** In this subsection we prove the following upper bound for degeneracy of 1,2-inclusion graphs of bouquets of halved cubes:

Proposition 2. *Let $\mathcal{B} \subset 2^X$ be a bouquet of halved cubes of clique-VC-dimension $d := \text{cVC-dim}(\mathcal{B})$, and let $G := G_{1,2}(\mathcal{B})$. Then the degeneracy of G is at most $\binom{d}{2}$.*

Proof. Let S be a set of maximal size of $\mathcal{B}$. By Lemma 8(iii), $\mathcal{B} \setminus \{S\}$ is a bouquet of halved cubes. Thus, it suffices to show that the degree of S in G is upper bounded by $\binom{d}{2}$. From Lemma 8(ii), we know that $|S| \leq d$. This implies that S is incident in G to at most $\binom{d}{2}$ vertical edges. Therefore, it remains to bound the number of horizontal edges sharing S. The following lemma will be useful for this purpose:

Lemma 9. *If $|S| = d - k \leq d$, then S is incident in G to at most $(d-k)k$ horizontal edges.*

Proof. Pick any $s \in S$ and set $Y := S \setminus \{s\}$. For an element $e \in X \setminus S$, let $S^e_s := Y \cup \{e\}$. Notice that such S^e_s are exactly the neighbors of S in $\frac{1}{2}Q_m$ connected by a horizontal edge. Let $X' = \{e \in X \setminus S : S^e_s \in \mathcal{B}\}$.

Pick any element $y \in Y$. Then $y \in S^e_s$ for any $e \in X'$. Since $\mathcal{B}$ is a bouquet of halved cubes, each of the $d-k-1$ pairs $\{y, e'\}$ with $e' \in S \setminus \{y\}$ belongs to $\mathcal{B}$ (yielding $P(y, S \setminus \{y\}) \subseteq \mathcal{B}$). To each set $S^e_s, e \in X'$, corresponds the unique pair $\{y, e\}$ and $\{y, e\} \in \mathcal{B}$ because $y, e \in S^e_s$. Therefore $P(y, X') \subset \mathcal{B}$. Since $|P(y, S \setminus \{y\})| + |P(y, X')| \leq d - 1$ and $|P(y, S \setminus \{y\})| = d - k - 1, |X'| = |P(y, X')|$, we conclude that $|X'| \leq k$. Therefore, for a fixed element $s \in S$, S has at most k neighbors of the form S^e_s with $e \in X'$. Since there are $|S| = d - k$ possible choices of the element s, S has at most $(d-k)k$ neighbors of cardinality $|S|$. $\square$

We now continue the proof of Proposition 2. Let $|S| = d - k \leq d$. Then S has $\binom{d-k}{2}$ neighbors of the form $S \setminus \{e, e'\}$ with $e \neq e' \in S$, i.e., S has $\binom{d-k}{2}$ incident vertical edges. It remains to bound the number of neighbors of S of the form $S \setminus \{e\} \cup \{e'\}$ with $e \in S$ and $e' \in X \setminus S$. By Lemma 9, S has at most $(d-k)k$ such neighbors. Summarizing, S possesses $(d-k)k + \binom{d-k}{2} = \frac{1}{2}(d^2 - d - k^2 + k)$ neighbors in G, and this number is maximal for $k = 0$ because

$$\frac{1}{2}(d^2 - d - k^2 + k) = \frac{1}{2}(d^2 - d) - \frac{1}{2}(k^2 - k) = \binom{d}{2} - \binom{k}{2} \leq \binom{d}{2}.$$

Hence, the degree of S in G is at most $\binom{d}{2}$, as asserted. $\square$

5.4. **Proof of Theorem 1.** First, let $\mathcal{S}$ be an even set family over X with $|X| = m$, $d = \text{cVC-dim}^*(\mathcal{S})$ be the clique-VC-dimension of $\mathcal{S}$, and $G_{1,2}(\mathcal{S}) = (V, E)$ be the 1,2-inclusion graph of $\mathcal{S}$. We have to prove that $\frac{|E|}{|V|} \leq \binom{d}{2} =: D$.

Let A be a set of $\mathcal{S}$ such that $\text{cVC-dim}(\mathcal{S}\Delta A) = \text{cVC-dim}^*(\mathcal{S}) = d$. By Lemma 3, $G_{1,2}(\mathcal{S} \triangle A) \backsimeq G_{1,2}(\mathcal{S})$. Thus it suffices to prove the inequality $\frac{|E(G_{1,2}(\mathcal{S}\triangle A))|}{|V(G_{1,2}(\mathcal{S}\triangle A))|} \leq D$. Consider a complete sequence of d-shiftings of $\mathcal{S} \triangle A$ and denote by $(\mathcal{S} \triangle A)^*$ the resulting set family. Since $\mathcal{S} \triangle A$ is a pointed even set family, applying Lemma 6 to each d-shifting, we deduce that $\text{cVC-dim}((\mathcal{S} \triangle A)^*) \leq \text{cVC-dim}(\mathcal{S} \triangle A) = d$. By Lemma 7, $(\mathcal{S} \triangle A)^*$ is a bouquet of halved cubes, thus, by Proposition 2, the degeneracy of its 1,2-inclusion graph $G^* = G_{1,2}((\mathcal{S} \triangle A)^*)$ is at most D. Therefore, if $G^* = (V^*, E^*)$, then $\frac{|E^*|}{|V^*|} \leq D$ (here we used the fact that the degeneracy of a graph $G = (V, E)$ is an upper bound for the ratio $\frac{|E|}{|V|}$). Applying Proposition 1 to each of the d-shiftings and taking into account that $G_{1,2}(\mathcal{S} \triangle A) \backsimeq G_{1,2}(\mathcal{S})$, we conclude that $\frac{|E|}{|V|} \leq \frac{|E^*|}{|V^*|}$, yielding the required density inequality $\frac{|E|}{|V|} \leq D$ and finishing the proof of Theorem 1 in case of even set families. If $\mathcal{S}$ is an arbitrary set family, then $\text{cVC-dim}^*(\mathcal{S}) = \text{cVC-dim}^*(\mathcal{S}^+)$, where $\mathcal{S}^+$ is the lifting of $\mathcal{S}$ to an even set family. Since by Lemma 1, $\mathcal{S}$ and $\mathcal{S}^+$ have isomorphic 1,2-inclusion graphs, the density result for $\mathcal{S}$ follows from the density result for $\mathcal{S}^+$. This concludes the proof of Theorem 1.

Example 4. As in the case of classical VC-dimension and Theorem 2, the inequality from Theorem 1 between the density of 1,2-inclusion graph $G_{1,2}(\mathcal{S})$ and the clique-VC-dimension of $\mathcal{S}$ is sharp in the following sense: there exist even set families $\mathcal{S}$ such that the degeneracy of $G_{1,2}(\mathcal{S})$ equals to $\binom{d}{2}$. For example, the sporadic clique $\mathcal{S}_2$ has clique VC-dimension 3 (see Examples 2 and remark 1), degeneracy 3, and density $\frac{3}{2}$). Notice that $G_{1,2}(\mathcal{S}_2)$ is the halved cube $\frac{1}{2}Q_3$. More generally, let $\mathcal{S}_4$ be the even set family consisting of all even subsets of an m-set X. Clearly $d := \text{cVC-dim}^*(\mathcal{S}_4) = |X| = m$ and $\mathcal{S}_4$ induces the halved cube $\frac{1}{2}Q_m$. We assert that $\frac{1}{2}Q_m$ has degeneracy $\binom{d}{2}$. Indeed, every $S \in \mathcal{S}_4$ is incident to $\binom{|X|-|S|}{2}$ supersets of cardinality $|S|+2$, to $\binom{|S|}{2}$ subsets of cardinality $|S|-2$, and to $|S|(|X|-|S|)$ sets of cardinality $|S|$. Setting $s := |S|$, we conclude that each set S has degree

$$\frac{(m-s)(m-s-1)+s(s-1)}{2} + s(m-s) = \frac{1}{2}(m^2-m) = \frac{1}{2}(d^2-d) = \binom{d}{2}.$$

Remark 3. In the following table, for pointed even set families $\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_4$ defined in Examples 1-3 and 4, we present their VC-dimension, the two clique VC-dimensions, the 2VC-dimension, the degeneracy, and the density.

$\mathcal{S}$	VC-dim	cVC-dim	cVC-dim*	degeneracy	density	2VC-dim
$\mathcal{S}_0$	1	–	m	$m-1$	$\frac{m-1}{2}$	0
$\mathcal{S}_1$	1	m	m	$m-1$	$\frac{m-1}{2}$	2
$\mathcal{S}_2$	2	3	3	3	$\frac{3}{2}$	3
$\mathcal{S}_3$	2	$k+2$	$k+2$	k	$\frac{k}{2}+o(1)$	2
$\mathcal{S}_3 \triangle X$	2	$(m-1)k+1$	$k+2$	k	$\frac{k}{2}+o(1)$	2
$\mathcal{S}_4$	$m-1$	m	m	$\binom{m}{2}$	$\frac{1}{2}\binom{m}{2}$	m

6. Final discussion

In this note, we adapted the shifting techniques to prove that if $\mathcal{S}$ is an arbitrary set family and $G_{1,2}(\mathcal{S}) = (V, E)$ is the 1,2-inclusion graph of $\mathcal{S}$, then $\frac{|E|}{|V|} \leq \binom{d}{2}$, where $d := \text{cVC-dim}^*(\mathcal{S})$ is the clique-VC-dimension of $\mathcal{S}$. The essential ingredients of our proof are Proposition 1 (showing that d-shiftings preserve the number of vertices and do not decrease the number of edges), Lemma 6 (showing that d-shiftings do not increase the clique-VC-dimension), and Proposition 2 (bounding the density of bouquets of halved cubes, resulting from complete d-shiftings), all established for even set families. While Propositions 1 and 2 are not very sensitive to the chosen definition of the clique-VC-dimension (but they require using the definition of 1,2-inclusion graphs as the subgraphs of the halved cube $\frac{1}{2}Q_m$), Lemma 6 strongly depends on how the clique-VC-dimension is defined. For example, this lemma does not hold for the notion of 2VC-dimension of [4] discussed in Section 2. Notice also that, differently from the classical VC-dimension and similarly to our notion of clique-VC-dimension, 2VC-dimension is not invariant under twistings.

In analogy to 2-shattering and 2VC-dimension, we can define the concepts of star-shattering and star-VC-dimension, which might be useful for finding sharper upper bounds (than those obtained in this paper) for density of 1,2-inclusion graphs. Let $Y \subset X$ and $e \notin Y$. We say that a set family $\mathcal{S}$ *star-shatters* (or *s-shatters*) the pair (e, Y) if for any $y \in Y$ there exists a set $S \in \mathcal{S}$ such that $S \cap (Y \cup \{e\}) = \{e, y\}$. The *star-VC-dimension* of a pointed set family $\mathcal{S}$ is

$$\text{sVC-dim}(\mathcal{S}) := \max\{|Y| + 1 : Y \subset X \text{ and } \exists e_i \in X \setminus Y \text{ such that } (e_i, Y) \text{ is s-shattered by } \mathcal{S}\}.$$

The difference with c-shattering is that, in the definition of s-shattering, a pair (e, Y) is s-shattered if all 2-sets of $P(e, Y)$ have non-empty fibers, i.e., if $P(e, Y) \subseteq \pi_{Q(e,Y)}(\mathcal{S})$. Consequently, any s-shattered pair (e, Y) is c-shattered, thus sVC-dim$(\mathcal{S}) \leq$ cVC-dim$(\mathcal{S})$. Since sVC-dim$(\mathcal{S}_3 \triangle X) = 3$ and $G_{1,2}(\mathcal{S}_3 \triangle X)$ contains a clique of size $k+1$, sVC-dim$(\mathcal{S})$ cannot be used directly to bound the density of 1,2-inclusion graphs. We can adapt this notion by taking the maximum over all twistings with respect to sets of $\mathcal{S}$: the *star-VC-dimension* sVC-dim$^*(\mathcal{S})$ of an arbitrary set family $\mathcal{S}$ is $\max\{\text{sVC-dim}(\mathcal{S}\Delta A) : A \in \mathcal{S}\}$[1]. Even if sVC-dim$(\mathcal{S}) \leq$ cVC-dim$(\mathcal{S})$ holds for pointed families, as the following examples show, there are no relationships between cVC-dim$^*(\mathcal{S})$ and sVC-dim$^*(\mathcal{S})$ for even families.

Example 5. Let $X = \{1, 2, \ldots, 2m-1, 2m\}$, where m is an arbitrary even integer, and let $\mathcal{S}_5 := \{\varnothing\} \cup \{\{1, 2, \ldots, 2i-1, 2i\} : i = 1, \ldots, m\}$. The nonempty sets of $\mathcal{S}_5$ can be viewed as intervals of even length of $\mathbb{N}$ with a common origin. The 1,2-inclusion graph of $\mathcal{S}_5$ is a path of length m. For any set $\{1, 2, \ldots, 2i\}$, the twisted family $\mathcal{S}_5^i := \mathcal{S}_5 \triangle \{1, 2, \ldots, 2i\}$ is the union of the set families $\mathcal{S}' := \{\varnothing, \{2i+1, 2i+2\}, \ldots, \{2i+1, 2i+2, \ldots, 2m\}\}$ and $\mathcal{S}'' := \{\{1, 2, \ldots, 2i-1, 2i\}, \ldots, \{2i-1, 2i\}\}$. We assert that for any $i = 1, \ldots, m$, we have sVC-dim$(\mathcal{S}_5^i) \leq 3$ and cVC-dim$(\mathcal{S}_5^i) = \max\{i, m-i\} + 1$. Indeed, for any element $j \in X$, $\mathcal{S}_5^i$ cannot simultaneously s-shatter two pairs $\{j, l_1\}$, $\{j, l_2\}$ with $j < l_1 < l_2$ because every set of $\mathcal{S}_5^i$ containing l_2 also contains l_1. Analogously, $\mathcal{S}_5^i$ cannot s-shatter two pairs $\{j, l_1\}$ and $\{j, l_2\}$ with $l_2 < l_1 < j$. Consequently, if the pair (j, Y) is s-shattered by $\mathcal{S}_5^i$, then $|Y| \leq 2$. This shows that sVC-dim$^*(\mathcal{S}_5) \leq 3$.

To see that cVC-dim$(\mathcal{S}_5^i) = \max\{i, m-i\}+1$, notice that $\mathcal{S}'$ c-shatters the pair $(2i+1, Y')$ with $Y' := \{2i+2, 2i+4, \ldots, 2m\}$ and $\mathcal{S}''$ c-shatters the pair $(2i, Y'')$ with $Y'' := \{1, 3, \ldots, 2i-1\}$. Since the minimum over all $i = 1, \ldots, m$ of $\max\{i, m-i\} + 1$ is attained for $i = \frac{m}{2}$, we conclude that cVC-dim$^*(\mathcal{S}_5) = \frac{m}{2} + 1$. Therefore sVC-dim$^*(\mathcal{S})$ can be arbitrarily smaller than cVC-dim$^*(\mathcal{S})$.

Example 6. Let $X = X_1 \dot\cup X_2$ with $X_1 = \{e_1, \ldots, e_m\}$ and $X_2 = \{x_1, \ldots, x_m\}$, and let $\mathcal{S}_6 := \{\varnothing, \{e_1, x_1\}\} \cup \{\{e_1, e_i, x_1, x_i\} : 2 \leq i \leq m\}$. The 1,2-inclusion graph of $\mathcal{S}_6$ is a star. One can easily see that sVC-dim$(\mathcal{S}_6) = m$. On the other hand, for the twisted family $\mathcal{S}_6' := \mathcal{S}_6 \triangle \{e_1, x_1\} = \{\varnothing\} \cup \{\{e_i, x_i\} : 1 \leq i \leq m\}$, one can check that cVC-dim$(\mathcal{S}_6') = 2$, showing that cVC-dim$^*(\mathcal{S}_6) = 2$ and sVC-dim$^*(\mathcal{S}_6) = m$. Therefore sVC-dim$^*(\mathcal{S})$ can be arbitrarily larger than cVC-dim$^*(\mathcal{S})$.

Therefore, it is natural to ask whether in Theorem 1 one can replace cVC-dim$^*(\mathcal{S})$ by sVC-dim$^*(\mathcal{S})$. However, we were not able to decide the status of the following question:

Question 1. Is it true that for any (even) set family $\mathcal{S}$ with the 1,2-inclusion graph $G_{1,2}(\mathcal{S}) = (V, E)$ and star-VC-dimension $d =$ sVC-dim$^*(\mathcal{S})$, we have $\frac{|E|}{|V|} = O(d^2)$?

The main difficulty here is that a d-shifting may increase the star-VC-dimension, i.e., Lemma 6 does no longer hold. The difference between the s-shattering and c-shattering is that a 2-set $\{e, y\}$ with $y \in Y$ can be s-shattered only by a set $S \in \mathcal{S}$ which belongs to the fiber $F(\{e, y\})$ (the requirement $Y \cap S = \{e, y\}$), while $\{e, y\}$ can be c-shattered by a set S if S just includes

[1] As noticed by one referee and O. Bousquet, in this form, the star-VC-dimension minus one coincides with the notion of *star number* that has been studied in the context of active learning [15, Definition 2].

this set (the requirement $\{e, y\} \subseteq S$). When performing a d-shifting φ_{ij} with respect to a pair $\{e_i, e_j\}$ such that $\{e_i, e_j\} \cap \{e, y\} = \varnothing$, a set $S \in \mathcal{S}$ can be mapped to a set $\varphi_{ij}(S)$ belonging to the fiber $F(\{e, y\})$. If $\varphi_{ij}(S)$ is used to c-shatter the 2-set $\{e, y\}$ by $\varphi_{ij}(\mathcal{S})$, then S can be used to shatter $\{e, y\}$ by $\mathcal{S}$ (the proof of Lemma 6). However, this is no longer true for s-shattering, because initially S may not necessarily belong to $F(\{e, y\})$.

Also we have not found a counterexample to the following question (where the square of the clique-VC-dimension or of the star-VC-dimension is replaced by the product of the classical VC-dimension of $\mathcal{S}$ and the clique number of $G_{1,2}(\mathcal{S})$):

Question 2. Is it true that for any set family $\mathcal{S}$ with 1,2-inclusion graph $G_{1,2}(\mathcal{S}) = (V, E)$, $d = \text{VC-dim}(\mathcal{S})$, and clique number $\omega = \omega(G_{1,2}(\mathcal{S}))$, we have $\frac{|E|}{|V|} = O(d \cdot \omega)$?

Hypercubes are subgraphs of Johnson graphs, therefore they are 1,2-inclusion graphs. This shows the necessity of both parameters (VC-dimension and clique number) in the formulation of Question 2. As above, the bottleneck in solving Question 2 via shifting is that this operation may increase the clique number of 1,2-inclusion graphs.

An alternative approach to Questions 1 and 2 is to adapt the original proof of Theorem 2 given in [19]. In brief, for a set family $\mathcal{S}$ of VC-dimension d and an element e, let $\mathcal{S}_e = \{S' \subseteq X \setminus \{e\} : S' = S \cap X \text{ for some } S \in \mathcal{S}\}$ and $\mathcal{S}^e = \{S' \subseteq X \setminus \{e\} : S' \text{ and } S' \cup \{e\} \text{ belong to } \mathcal{S}\}$. Then $|\mathcal{S}| = |\mathcal{S}_e| + |\mathcal{S}^e|$, $\text{VC-dim}(\mathcal{S}_e) \leq d$, and $\text{VC-dim}(\mathcal{S}^e) \leq d - 1$ hold. Denote by G_e and G^e the 1-inclusion graphs of $\mathcal{S}_e$ and $\mathcal{S}^e$. Then $|E(G_e)| \leq d|V(G_e)| = d|\mathcal{S}_e|$ and $E(G^e) \leq (d-1)|V(G^e)| = (d-1)|\mathcal{S}^e|$ by induction hypothesis. The proof of the required density inequality follows by induction from the equality $|V(G)| = |\mathcal{S}| = |\mathcal{S}_e| + |\mathcal{S}^e| = |V(G_e)| + |V(G^e)|$ and the inequality $|E(G)| \leq |E(G_e)| + |E(G^e)| + |V(G^e)|$. Unfortunately, as was the case for shiftings, the clique number of $G_{1,2}(\mathcal{S}_e)$ may be strictly larger than the clique number of $G_{1,2}(\mathcal{S})$. Also the inequality $|E(G)| \leq |E(G_e)| + |E(G^e)| + |V(G^e)|$ is no longer true in this form if instead of 1-inclusion graphs one consider 1,2-inclusion graphs.

Acknowledgements. We would like to acknowledge the anonymous referees for a careful reading of the previous version and many useful remarks. We are especially indebted to one of the referees who found a critical error in the previous proof of Proposition 1. We would like to acknowledge another referee and Olivier Bousquet for pointing to us the paper [15]. This work was supported in part by ANR project DISTANCIA (ANR-17-CE40-0015).

References

[1] R. Ahlswede and N. Cai, A counterexample to Kleitman's conjecture concerning an edge-isoperimetric problem, Combinatorics, Probability and Computing 8 (1999), 301–305.

[2] H.-J. Bandelt and V. Chepoi, Metric graph theory and geometry: a survey, in: J. E. Goodman, J. Pach, R. Pollack (Eds.), Surveys on Discrete and Computational Geometry. Twenty Years later, Contemp. Math., vol. 453, AMS, Providence, RI, 2008, pp. 49–86.

[3] S. L Bezrukov, Edge isoperimetric problems on graphs, Proc. Bolyai Math. Studies 449 (1998).

[4] N. Bousquet and S. Thomassé, VC-dimension and Erdös-Pósa property, Discr. Math. 338 (2015), 2302–2317.

[5] A.E. Brouwer, A.M. Cohen, and A. Neumaier, Distance Regular Graphs, Springer-Verlag Berlin, New York, 1989.

[6] N. Cesa-Bianchi and D. Haussler, A graph-theoretic generalization of Sauer-Shelah lemma, Discr. Appl. Math. 86 (1998), 27–35.

[7] V. Chepoi, Basis graphs of even Delta-matroids, J. Combin. Th. Ser B 97 (2007), 175–192.

[8] V. Chepoi, Distance-preserving subgraphs of Johnson graphs, Combinatorica (to appear).

[9] V. Chepoi, A. Labourel, and S. Ratel, On density of subgraphs of Cartesian products, (in preparation).

[10] M. Deza and M. Laurent, Geometry of Cuts and Metrics, Springer-Verlag, Berlin, 1997.

[11] V. Diego, O. Serra, and L. Vena, On a problem by Shapozenko on Johnson graphs, arXiv:1604.05084, 2016.

[12] R. Diestel, Graph Theory, Graduate texts in mathematics, Springer New York, Berlin, Paris, 1997.

[13] D.Ž. Djoković, Distance–preserving subgraphs of hypercubes, J. Combin. Th. Ser. B 14 (1973), 263–267.

[14] M.R. Garey and R.L. Graham, On cubical graphs, J. Combin. Th. B 18 (1975), 84–95.
[15] S. Hanneke and L. Yang, Minimax analysis of active learning, J. Mach. Learn. Res. 16 (2015), 3487–3602.
[16] L.H. Harper, Optimal assignments of numbers to vertices, SIAM J. Appl. Math., 12 (1964), 131–135.
[17] L.H. Harper, Global Methods for Combinatorial Isoperimetric Problems, Cambridge Studies in Advanced Mathematics (No. 90), Cambridge University Press 2004.
[18] D. Haussler, Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension, J. Comb. Th. Ser. A 69 (1995), 217–232.
[19] D. Haussler, N. Littlestone, and M. K. Warmuth, Predicting $\{0,1\}$-functions on randomly drawn points, Inf. Comput. 115 (1994), 248–292.
[20] D. Haussler and P.M. Long, A generalization of Sauer's lemma. J. Combin. Th. Ser. A 71 (1995), 219–240.
[21] D. Kuzmin and M.K. Warmuth, Unlabelled compression schemes for maximum classes, J. Mach. Learn. Res. 8 (2007), 2047–2081.
[22] S.B. Maurer, Matroid basis graphs I, J. Combin. Th. Ser. B 14 (1973), 216–240.
[23] B. K. Natarajan, On learning sets and functions, Machine Learning 4 (1989), 67–97.
[24] D. Pollard, Convergence of Stochastic Processes, Springer Science & Business Media, 2012.
[25] B.I. Rubinstein, P.L. Bartlett, and J.H. Rubinstein, Shifting: one-inclusion mistake bounds and sample compression, J. Comput. Syst. Sci. 75 (2009), 37–59.
[26] N. Sauer, On the density of families of sets, J. Combin. Th., Ser. A 13 (1972), 145–147.
[27] S.V. Shpectorov, On scale embeddings of graphs into hypercubes, Europ. J. Combin. 14 (1993), 117–130.
[28] V.N. Vapnik and A.Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Theory Probab. Appl. 16 (1971), 264–280.