Karlsruher Institut für Technologie

Bachelor Thesis

Induced Turán problems

Jakob Zimmermann

May 20, 2024

Reviewers: Dr. Torsten Ueckerdt

T.T.-Prof. Dr. Thomas Bläsius

Prof. Dr. Maria Axenovich

Advisor: Prof. Dr. Maria Axenovich

Department of Computer Science
Department of Mathematics

Karlsruhe Institute of Technology

Statement of Authorship

I hereby declare that this document has been composed by myself and describes my own work, unless otherwise acknowledged in the text. I declare that I have observed the *Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis*, as amended.
Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde, sowie die *Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis* in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe Germany, May 20, 2024.

Abstract

Extremal Graph Theory is a branch of combinatorics that studies the phenomena that global constraints on graphs, as edge density, force the existence of local substructure. Given two graphs G and H the classical extremal function $\mathrm{ex}(G, H)$ is defined as the maximal number of edges in a subgraph of G that does not contain any copies of H, where a copy of H in G is defined as a subgraph of G that is isomorphic to H.
Recently the interest arose in studying subgraphs of some host graph G that do not contain induced copies of H or biinduced copies of $\tilde{H}$, where $\tilde{H}$ is a bipartite graph. We denote the corresponding extremal functions by $\mathrm{ex}(G, H\text{-ind})$ and $\mathrm{ex}\left(G, \tilde{H}\text{-biind}\right)$ respectively. Since these Definitions are trivial in case that H is no induced copy of G or $\tilde{H}$ is no biinduced copy of G respectively, we are also interested in various cases of mixed restrictions.
For a third graph F the extremal function $\mathrm{ex}(G, \{F, H\text{-ind}\})$ is defined to be the maximal number of edges in a subgraph of G, that neither contains any copy of F nor any induced copy of H. Determining the latter extremal function reduces to finding either $\mathrm{ex}(G, F)$ or $\mathrm{ex}(G, H)$ unless H is a biclique or both F and H are bipartite.
By strengthening a result of Sudakov and Tomon we show that for any $d, t \in \mathbb{N}$ with $t \geq d \geq 2$ and any $K_{d,d}$-free bipartite graph H where each vertex in one of its partite sets is either complete or has degree at most d, one has $\mathrm{ex}(K_n, \{K_{t,t}, H\text{-biind}\}) = o\left(n^{2-\frac{1}{d}}\right)$. This provides an upper bound on the biinduced extremal function for a wide class of bipartite graphs and implies in particular an extremal result for bipartite graphs of bounded Vapnik Chervonenkis dimension by Janzer and Pohoata. This result is also part of the paper [4] by Maria Axenovich and the author, that arose during the creation of this thesis.
Furthermore, for graphs G, H and F, where G does not contain any copy of F, we are interested in counting induced copies of H inside G. The lower bounds, we obtain in case that H is a bipartite graph fulfilling certain degree conditions and F is a biclique of specific size, asymptotically imply a result on the extremal function $\mathrm{ex}(K_n, \{F, H\text{-ind}\})$ by Hunter, Milojevic, Sudakov and Tomon. Furthermore, under even stricter degree conditions on H, it matches the lower bound for the number of graph homomorphisms from H to G given by Sidorenkos Conjecture up to constants.
Apart from these results we provide a comprehensive introduction into Extremal Graph Theory and important connections to various notions of Vapnik Chervonenkis dimension. Hereby we develop some examples where it is possible to determine the extremal function exactly. Moreover, we present common Reduction lemmas with reworked constants and exponents for more convenient application. We study the Vapnik Chervonenkis dimension of hereditary graph properties and geometrically motivated set systems, especially the k-fold union of halfspaces. The introduction to the Vapnik Chervonenkis dimension leads to the presentation of a powerful Packing lemma for hypergraphs by Fox, Pach, Sheffer, Suk and Zahl.
At last, we give a full and simplified proof of the Erdős-Hajnal conjecture for graphs with bounded Vapnik Chervonenkis dimension, a major and very recent breakthrough by Nguyen, Scott and Seymour. Interestingly here the distinction between the restriction of forbidding induced and biinduced copies of some bipartite graph plays a crucial role. Our presentation includes the generalization of the Ultra Strong Regularity lemma for graphs with bounded Vapnik Chervonenkis dimension to uniform hypergraphs by Fox, Pach and Suk.

Abstrakt

Extremale Graphentheorie untersucht das Phänomen, dass globale Eigenschaften von Graphen, wie die Kantendichte, lokale Substrukturen erzwingen. Gegeben zwei Graphen G und H definiert sich die klassische extremale Funktion $\mathrm{ex}\,(G, H)$ als die maximale Kantenanzahl eines Subgraphen von G, der keine Kopie von H enthält. Unter einer Kopie von H in G verstehen wir hier einen zu H isomorphen Subgraphen von G.
Vor kurzem entstand Interesse an der Untersuchung der Menge an Subgraphen von G, die keine induzierten Kopien von H oder biinduzierten Kopien von $\tilde{H}$ enthalten, wobei $\tilde{H}$ ein bipartiter Graph ist. Wir bezeichnen die entsprechenden extremalen Funktionen mit $\mathrm{ex}\,(G, H\text{-ind})$ bzw. $\mathrm{ex}\left(G, \tilde{H}\text{-biind}\right)$. Da sich diese Definitionen jedoch als trivial erweisen, wenn G keine induzierte Kopie von H bzw. keine biinduzierte Kopie von $\tilde{H}$ enthält, interessieren wir uns besonders für die Fälle gemischter Restriktionen.
Gegeben einen dritten Graphen F definieren wir die extremale Funktion $\mathrm{ex}\,(G, \{F, H\text{-ind}\})$ als die maximale Anzahl an Kanten in einem Subgraphen von G, der weder eine Kopie von F noch eine induzierte Kopie von H enthält. Gegeben den Fall, dass H keine Biklique und einer der beiden Graphen F und H nicht bipartit ist, reduziert sich diese extremale Funktion entweder auf $\mathrm{ex}\,(G, F)$ oder auf $\mathrm{ex}\,(G, H)$.
Durch die Verallgemeinerung eines Ergebnisses von Sudakov und Tomon zeigen wir $\mathrm{ex}\,(K_n, \{K_{t,t}, H\text{-biind}\}) = o\left(n^{2-\frac{1}{d}}\right)$, wobei $d, t \in \mathbb{N}$ mit $t \geq d \geq 2$ zwei beliebige natürliche Zahlen sind und H ein $K_{d,d}$-freier bipartiter Graph ist, bei dem jeder Knoten in einer seiner Partitionsmengen entweder vollständig ist oder höchstens Grad d hat. Dies liefert eine obere Schranke für die biinduzierte extremale Funktion für eine weite Klasse an bipartiten Graphen und impliziert insbesondere ein extremales Resultat von Janzer und Pohoata zu bipartiten Graphen mit beschränkter Vapnik-Chervonenkis Dimension. Das Resultat ist auch Teil der im Rahmen der Bachelor Arbeit entstandenen Veröffentlichung [4].
Unser zweites Resultat handelt von Graphen G, H und F, bei denen G keine Kopie von F enthält, und gibt eine untere Schranke für die Anzahl induzierter Kopien von H in G. Im Falle, dass H ein bipartiter Graph ist, der bestimmte Bedingungen an seine Grade erfüllt, und F eine Biklique einer bestimmten Größe ist, impliziert unser Theorem das Resultat über $\mathrm{ex}\,(K_n, \{F, H\text{-ind}\})$ von Hunter, Milojevic, Sudakov und Tomon.
Neben diesen neuen Ergebnissen präsentieren wir eine umfassende Einführung in die Extremale Graphentheorie und bauen die Verbindung zu der so genannten Vapnik-Chervonenkis Dimension. Hierbei entwickeln wir einige anschauliche Beispiele und präsentieren nützliche Reduktionslemmatas mit überarbeiteten Konstanten und Exponenten für eine bequemere Anwendung. Wir untersuchen die Vapnik-Chervonenkis Dimension von hereditären Graphenfamilien und geometrisch motivierten Mengensystemen, insbesondere der k-fachen Vereinigung von Halbräumen. Die Einführung in die Vapnik-Chervonenkis Dimension mündet in der Präsentation des mächtigen Theorems über Hypergraphen-Packungen von Fox, Pach, Sheffer, Suk und Zahl.
Zuletzt geben wir einen vollständigen und vereinfachten Beweis der Erdős-Hajnal-Vermutung für Graphen mit beschränkter Vapnik-Chervonenkis Dimension, ein bedeutender und sehr aktuellen Durchbruch von Nguyen, Scott und Seymour. Unsere Darstellung umfasst die Verallgemeinerung des ultra-starken Regularitätslemmas für Graphen mit beschränkter Vapnik-Chervonenkis Dimension auf uniforme Hypergraphen von Fox, Pach und Suk.

Contents

1 Introduction

The questions considered by Extremal Graph Theory are among the most natural ones in mathematics but lead to a very broad and rich theory that has many connections to functional analysis, number theory and geometry. Extremal Graph Theory finds various application in Computational Geometry. Historically, the roots of the field lie in Mantels theorem, the characterization of edge maximal graphs that do not contain triangles, which was discovered in 1907. In 1938 Paul Erdős studied edge maximal bipartite graphs that do not contain a four-cycle to tackle the multiplicative Sidon problem from number theory. The generalization of Mantels theorem, a characterization of edge maximal graphs that do not contain cliques of fixed size, was found by the Hungarian mathematician Pál Turán in 1941. Until today, we refer to problems related to the forbidden subgraph problem by *Turán type problems* [24].
Many breakthroughs in Computational Geometry originated from results in extremal combinatorics. Interestingly many natural occurring set systems such as intersection hypergraphs show unusually strong Ramsey-type properties, meaning they contain very large cliques or independent sets. Furthermore, they often do not allow for large δ-packings, meaning that high edge density implies similarity between the edges. One explanation for their simple structure could be that many geometrically defined set systems have a bounded Vapnik Chervonenkis dimension [20].
In this thesis we consider the classical forbidden subgraph problem as well as the problem of isomorphism counting in an induced setting. We compare our results to the classical, non-induced case. Furthermore, we draw the connection to the concept of Vapnik Chervonenkis dimension, a complexity measure for hypergraphs. Another angle on the graph property $\text{Free}\,(H\text{-ind})$ of graphs G that do not contain another graph H as an induced subgraph is provided by the longstanding Erdős-Hajnal conjecture. Here the problem is to find graphs in $\text{Free}\,(H\text{-ind})$ that do not contain large homogeneous sets, where a homogeneous set is defined to be either a clique or an independent set. The Conjecture states that graphs in a proper hereditary graph property have polynomially large homogeneous sets. During the work on this thesis, building on work of Fox, Pach and Suk in [20], Nguyen, Scott and Seymour proved the Conjecture for graphs of bounded VC dimension in [40]. However, the general Conjecture is still open.

We have structured the thesis as follows. In section 2 we define the forbidden subgraph problem in its most general form, give examples, and present cornerstone Theorems of Extremal Graph Theory and their analogue in the induced setting. Furthermore, we make preparations for our main results, including a counting Lemma for independent sets in graphs that do not contain a copy of $K_{s,s}$ as well as a powerful Reduction lemma, originating from work of Tao Jiang and Robert Seiver in [34].
In section 3 we give an extensive introduction to various notions of Vapnik Chervonenkis dimension. Aside some important standard results we provide a full proof of the celebrated Packing lemma, Theorem 10, where we rely on the work of Fox, Pach, Sheffer, Suk and Zahl in [18]. Furthermore, as a case study from Computational Geometry, we give a refined proof of the asymptotics for the VC dimension of so-called k-fold unions of halfspaces with respect to their dimension, based on the results of Csikós, Mustafa and Kupavskii in [14] as well as Kupavskii, Nabil, Pach in [35]. Bridging to the induced forbidden subgraph problem we study the Vapnik Chervonenkis dimension of hereditary graph properties and present a short proof for a result of Bousquet, Lagoutte, Li, Parreau and Thomassé in [9].
Our main results can be found in section 4. Building on work of Sudakov and Tomon in [44] as well as Janzer and Pohoata [32] in Theorem 16 for $d \in \mathbb{N}$ we give an upper bound on $\text{ex}\,(n, \{K_{s,s}, H\text{-biind}\})$ for $K_{d,d}$-free bipartite graphs H with one partite set in which every vertex has either a full degree or degree at most d. Furthermore, in Theorem 19 we present a counting framework for the number of induced isomorphisms from a bipartite graph H to some host graph G, in case that H fulfills some degree condition with parameter $d \in \mathbb{N}$ and G is a dense $K_{d+1,d+1}$-free graph. Those bounds imply state-of-the-art bounds for the extremal function $\text{ex}\,(n, \{K_{d+1,d+1}, H\text{-ind}\})$.
In section 5 we introduce the Erdős-Hajnal conjecture, collate related results and draw the connection to the polynomial Rödl property. We present a reworked and self-sustained proof of the Erdős-Hajnal conjecture

for graph properties of bounded Vapnik Chervonenkis dimension, Theorem 25, where we rely on the work of Nguyen, Scott and Seymour in [40] as well as Fox, Pach and Suk in [20].

1.1 Preliminaries

In this section we introduce notation for graphs and hypergraphs as well as simple inequalities and underlying Theorems we are going to use throughout the thesis. For most of the standard notion of graphs, as completeness, independence number, regularity etc. we refer the reader to the introduction section in Diestel [15].

Let $n \in \mathbb{Z}$ and X be an arbitrary set. We denote $[n] \coloneqq \{1, \ldots, n\}$. Notice that for $n \leq 0 : \ [n] = \emptyset$. Furthermore, let us define $\binom{X}{n} \coloneqq \{ \ A \subseteq X \mid |A| = n \ \}$. We remark that in case $n > |X|$ or $n < 0$ by Definition $\binom{X}{n} = \emptyset$. For $k \in \mathbb{N}$ we denote a sequence of k elements in X by $(x_j)_{j\in[k]}$. For convenience, we often write $(x_j)_{j\in[k]} \subseteq X$ instead of $(x_j)_{j\in[k]} \in X^k$. We write 2^X for the power set of X. Furthemore, for an other set Y we define all mappings from Y to X by X^Y. For $x \in \mathbb{R}$ and $z \in \mathbb{N}_0$ we define

$$\binom{x}{z} \coloneqq \mathbb{1}\{x \geq z\} \prod_{0 \leq j < z} \frac{x-j}{z-j}.$$

Let $n \in \mathbb{N}$ and let $(X_j)_{j\in[n]}$ be a sequence of pairwise disjoint sets. In this case we denote the disjoint union of all the sets by $\sum_{j\in[n]} X_j$ and the disjoint union of the two sets X_1 and X_2 by $X_1 \cup\!\!\!\cdot\, X_2$. Furthermore, we denote the symmetrical difference of two sets A, B by $A\Delta B \coloneqq (A \cup B) \setminus (A \cap B)$.

Let $G = (V,\ E)$ be a graph and $v \in V,\ A, B \subseteq V$. We remark that in this thesis all considered graphs are finite and simple. We denote $V(G) \coloneqq V$ and $E(G) \coloneqq E$.
We define neighborhood by $N_A(v) \coloneqq \{ \ w \in A \mid \{v, w\} \in E \ \}$ and $N_A(B) \coloneqq \bigcap_{b\in B} N_A(b)$ respectively. We remark that in this thesis we interpret the empty intersection as the whole set, meaning that $N_A(\emptyset) = A$. Notice further that A and B do not necessarily have to be disjoint. Our notion of degree follows this convention: $\deg_A(v) \coloneqq |N_A(v)|$ and $\deg_A(B) \coloneqq |N_A(B)|$. We denote the minimal degree of G by $\delta(G)$ and the maximal degree of G by $\Delta(G)$. Furthermore, we denote the average degree of G by $\text{avdeg}(G)$.
We denote the vertex count of G by $|G|$ and the edge count by $\|G\|$. Furthermore, we introduce the notation $E(A, B) \coloneqq \{ \ \{a, b\} \in E(G) \mid a \in A,\ b \in B \ \}$ as well as $\|A, B\| \coloneqq |E(A, B)|$. We say A *sends an edge towards* B if $E(A, B) \neq \emptyset$. In this case and if $A = \{ \ a \ \}$ we also say that a *sends an edge towards* B. We call a vertex $v \in V(G)$ *complete* if it is adjacent to all vertices in $V(G) \setminus \{v\}$.

We call a graph $G = (V,\ E)$ *bipartite* in case that there is a partition of V into two independent sets A, B. Of course, the partition A, B is not necessarily unique. We call A, B *partite sets*. Throughout the thesis we think of this bipartition implicitly fixed to a bipartite graph. By denoting $G = (A \cup\!\!\!\cdot\, B,\ E)$ we implicitly fix the partite sets of the bipartite graph as the tuple (A, B). In a bipartite graph we call a vertex *complete* if it is adjacent to all vertices in the partite set, that it does not belong to.

Furthermore, we introduce notation for induced and biinduced subgraphs, where we assume A and B to be disjoint.

Definition 1 (Induced subgraph). $G[A] \coloneqq \left(A,\ \left(E(G) \cap \binom{A}{2}\right)\right)$.

Definition 2 (Biinduced subgraph). $G[A, B] \coloneqq (A \cup\!\!\!\cdot\, B,\ (E(G) \cap \{ \ \{a, b\} \mid a \in A \text{ and } b \in B \ \}))$.

We use the standard notation for edge and vertex deletion. Let $E' \subseteq E(G)$ and $e \in E(G)$.

Definition 3 (Vertex deletion). $G - A \coloneqq G[V(G) \setminus A]$ and $G - v \coloneqq G - \{v\}$.

Definition 4 (Edge deletion). $G - E' \coloneqq (V(G),\ E(G) \setminus E')$ and $G - e \coloneqq G - \{e\}$.

There is a variety of different kinds of graph homomorphisms. We settle for the following three. Let H,G be graphs and $\tilde{H} := (A \cup B,\ E)$ be a bipartite graph.

Definition 5 (Graph homomorphism)**.**

$$\mathrm{Hom}(H,\ G) := \left\{ \Phi :\ V(H) \longrightarrow V(G) \ \middle|\ \forall \{u,v\} \in \binom{V(H)}{2} :\ \{u,v\} \in E(H) \Longrightarrow \{\Phi(u), \Phi(v)\} \in E(G) \right\}.$$

$$\mathrm{Hom}_{\mathrm{ind}}(H,\ G) := \left\{ \Phi :\ V(H) \longrightarrow V(G) \ \middle|\ \forall \{u,v\} \in \binom{V(H)}{2} :\ \{u,v\} \in E(H) \Longleftrightarrow \{\Phi(u), \Phi(v)\} \in E(G) \right\}.$$

$$\mathrm{Hom}_{\mathrm{biind}}\left(\tilde{H},\ G\right) := \left\{ \Phi :\ V\left(\tilde{H}\right) \longrightarrow V(G) \ \middle|\ \forall u \in A,\ v \in B :\ \{u,v\} \in E(H) \Longleftrightarrow \{\Phi(u), \Phi(v)\} \in E(G) \right\}.$$

Definition 6 (Graph isomorphism)**.**

$$\mathrm{Isom}(H,\ G) := \{\ \Phi \in \mathrm{Hom}(H,\ G) \mid \Phi \text{ is injective} \}.$$

$$\mathrm{Isom}_{\mathrm{ind}}(H,\ G) := \{\ \Phi \in \mathrm{Hom}_{\mathrm{ind}}(H,\ G) \mid \Phi \text{ is injective} \}.$$

$$\mathrm{Isom}_{\mathrm{biind}}\left(\tilde{H},\ G\right) := \left\{\ \Phi \in \mathrm{Hom}_{\mathrm{ind}}\left(\tilde{H},\ G\right) \ \middle|\ \Phi \text{ is injective} \right\}.$$

In this thesis we use the equal sign to express that two graphs are *isomorphic*, meaning that there are (surjective) graph isomorphisms from the one graph to the other. For sake of simplicity we sometimes even identify bipartite graphs if they are isomorphic, but we have fixed different partite sets for them. This ambiguity is common in the literature and never leads to heavyweight confusion. Furthermore, we refer to graphs that are isomorphic to H as *copies* of H.

Let us introduce some special subgraph notation. For this purpose let $H, \tilde{H}, G, \tilde{G}$ be graphs where $\tilde{H} = (A \cup B,\ F)$ and $\tilde{G} = (X \cup Y,\ E)$ are bipartite graphs. We call H a subgraph of G in case there is an isomorphism from H to G. In this case we also write $H \subseteq G$. We call H a *proper* subgraph of G in case that $H \subseteq G$ but $H \neq G$.

Definition 7 (Asymmetric subgraph)**.** Write $\tilde{H} \subseteq^* \tilde{G}$ if there is a copy of $\tilde{H}$ in $\tilde{G}$ where the vertices corresponding to A lie in X and the vertices corresponding to B lie in $\tilde{Y}$.

Definition 8 (Induced subgraph)**.** Write $H \underset{ind}{\subseteq} G$ in case that there is $V' \subseteq V(G)$ such that $H = G[V]$.

Definition 9 (Biinduced subgraph)**.** Write $H \underset{biind}{\subseteq} G$ if there are disjoint subsets $X, Y \subseteq V(G)$ such that $H = G[X, Y]$.

Definition 10 (Hereditary graph property)**.** We call a (possibly infinite) set of graphs $\mathscr{C}$ a *graph property.* We call it *proper* if it is not empty and does not contain all graphs. Furthermore, we call a graph property $\mathscr{C}$ *hereditary* if $\forall G \in \mathscr{C},\ H \underset{ind}{\subseteq} G :\ H \in \mathscr{C}$.

We want to introduce notation for graph complements. Let $G = (V,\ E)$ be a graph and $H = (A \cup B,\ F)$ be a bipartite graph.

Definition 11 (Graph complement)**.** $\overline{G} := \left(V,\ \binom{V}{2} \setminus E\right)$.

Definition 12 (Bipartite graph complement)**.** $\overline{\overline{H}} := (A \cup B,\ \{\ \{a,b\} \mid a \in A,\ b \in B\ \} \setminus F)$.

Furthermore, we introduce four binary graph operations. Let G, H be graphs.

Definition 13 (Disjoint sum)**.** Let us take a copy $\tilde{H}$ of H such that $V\left(\tilde{H}\right) \cap V(G) = \emptyset$. Define the *disjoint sum* of H and G by

$$G + H := G \cup H := \left(V(G) \cup V\left(\tilde{H}\right),\ E(G) \cup E\left(\tilde{H}\right)\right).$$

Furthermore, we inductively define

$$1H := H,$$
$$nH := H + (n-1)H,\ n \in \mathbb{N}_{\geq 2}.$$

Definition 14 (Union)**.** $G \cup H := (V(G) \cup V(H),\ E(G) \cup E(H))$.

Definition 15 (Disjoint product)**.** Let us take a copy $\tilde{H}$ of H such that $V\left(\tilde{H}\right) \cap V(G) = \emptyset$. Define the *disjoint product* of H and G by

$$G \times H := \overline{\left(\overline{G} + \overline{H}\right)} = \left(V(G) \cup V\left(\tilde{H}\right),\ E(G) \cup E\left(\tilde{H}\right) \cup \left\{ \{v, w\} \,\middle|\, v \in V(G), w \in V\left(\tilde{H}\right) \right\}\right).$$

Definition 16 (Tensor product)**.** Define the *tensor product* of H and G by

$$G \otimes H = (V(G) \times V(H),\ \{ \{(a_G, a_H), (b_G, b_H)\} \mid \{a_G, b_G\} \in E(G) \text{ and } \{a_H, b_H\} \in E(H) \}).$$

In case that $H = (\{v\},\ \emptyset)$ we also use the abbreviations $G + v := G + H$ and $G \times v = G \times H$.

Let $p \in (0,1)$, $n \in \mathbb{N}$ and let S be a set. We denote the binomial distribution by $\mathrm{Bin}(n,p)$. In case that $n = 1$ we call it *Bernoulli*-distribution and denote it by $\mathrm{Be}(p) := \mathrm{Bin}(1,p)$. Furthermore, we denote the uniform distribution over the elements of S by $\mathcal{U}(S)$.

For $n \in \mathbb{N}$ we denote the permutations of $[n]$ by $\mathscr{S}_n$. Furthermore we denote a permutation $\sigma \in \mathscr{S}_n$ by a vector $(\sigma(j))_{j \in [n]}$.

The naturals $\mathbb{N}$ do not include zero. We denote non-negative integers by $\mathbb{N}_0$. We use similar notation for the real numbers. $\mathbb{R}_0 := \{ x \in \mathbb{R} \mid x \geq 0 \}$. $\mathbb{R}_+ := \{ x \in \mathbb{R} \mid x > 0 \}$. Furthermore, we interpret $\min\{\emptyset\} := -\infty$ and $\max\{\emptyset\} := \infty$.
For $r > 0$ we denote the logarithm with basis r as $\log_r(\bullet)$. However, we denote the natural logarithm with $\ln(\bullet)$.

Let us introduce notation for sequences.

Definition 17 (Restricted sequence)**.** For some index sets $\mathscr{I}$, $\mathscr{J}$ with $\mathscr{J} \subseteq \mathscr{I}$ and some sequence $v = (v_i)_{i \in \mathscr{I}}$ let us introduce the notation

$$v\big|_{\mathscr{J}} := (v_j)_{j \in \mathscr{J}}.$$

Furthermore, for some set V of sequences with index set $\mathscr{I}$ we introduce the notation

$$V\big|_{\mathscr{J}} := \left\{ v\big|_{\mathscr{J}} \,\middle|\, v \in V \right\}.$$

Let $\mathscr{F} = (V,\ \mathscr{E})$ be a hypergraph, this means $\mathscr{E} \subseteq 2^V$. In some cases we allow V to be infinite and in some cases we allow $\mathscr{E}$ to be a multiset. However, we are always going to mark those cases. In case that there is $k \in \mathbb{N}$ such that $\mathscr{E} \subseteq \binom{V}{k}$ we call $\mathscr{F}$ *k-uniform.* Most of the notation for graphs can be directly generalized to hypergraphs. In most cases we use the notation without adjusting the definition to hypergraphs since the generalization is obvious. In some rare cases we identify $\mathscr{F}$ with its edges for notational convenience.

We define the incidence graph of $\mathscr{F}$ simply to be the bipartite graph where the vertices and edges of $\mathscr{F}$ represent the two partition classes and an edge is adjacent to a vertex if it contains it.

Definition 18 (Incidence graph)**.** $\mathrm{Incidence}(\mathscr{F}) := (V(\mathscr{F}) \cup E(\mathscr{F}),\ \{ \{a, A\} \mid a \in A \in E(\mathscr{F}) \})$.

For $j \in \mathbb{N}$ we define a one-sided j-blowup of the incidence graph.

Definition 19 (Blown up incidence graph)**.**

$$\text{Incidence}_j(\mathscr{F}) := (V(\mathscr{F}) \cup ([j] \times E(\mathscr{F})), \ \{ \{a, (i, A)\} \mid a \in A \in E(\mathscr{F}), i \in [j] \}).$$

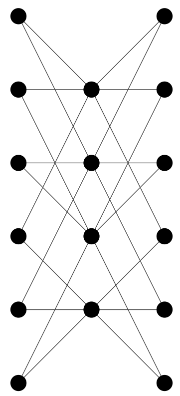

Figure 1: Rendering of a $(4,2,2)-$hedgehog.

Definition 20 ((k,d,j)-hedgehog)**.** For non-negative integers k, d, j where $k \geq d$ we define a (k, d, j)-*hedgehog* as

$$H(k, d, j) := \text{Incidence}_j\left(\binom{[k]}{d}\right).$$

We call the partite set $[k]$ the *body* of $H(k, d, r)$.

Definition 21 (Path)**.** For $l \in \mathbb{N}$ we define a *path of length* l by $P_l := ([l], \ \{ \{j, j+1\} \mid j \in [l-1] \})$. Given a graph P isomorphic to some path we introduce $\text{length}(P) := |P|$. Furthermore, we denote P by $(x_1, \ldots, x_l)$ where $l := \text{length}(P)$ and $P = (\{ x_j \mid j \in [l] \}, \ \{ \{x_j, x_{j+1}\} \mid j \in [l-1] \})$.

Definition 22 (Boolean Hypercube)**.** For $d \in \mathbb{N}$ let us introduce the *d-dimensional Boolean Hypercube*

$$Q_d := \left(2^{[d]}, \ \left\{ \{A, B\} \subseteq 2^{[d]} \;\middle|\; |A \Delta B| = 1 \right\}\right).$$

Finally, let us introduce notation for complete multipartite (hyper-)graphs.

Definition 23 (Complete multipartite (hyper-)graphs)**.** We say that a d-uniform hypergraph is *complete multipartite* if one can partition its vertex set such that any d-set of vertices is an edge if and only if it does not contain two vertices of the same partition class.
For $k, d \in \mathbb{N}$ and $(s_j)_{j \in [k]} \subseteq \mathbb{N}$ we denote the generic complete multipartite graph with partition classes of sizes $(s_j)_{j \in [k]}$ by

$$K^{(d)}_{(s_j)_{j \in [k]}} := \left(\{ (i, \ j) \mid i \in [k], \ j \in [s_i] \}, \ \bigcup_{I \in \binom{[k]}{d}} \left\{ \{ (i, \ \alpha_i) \mid i \in I \} \;\middle|\; \alpha \in \bigtimes_{i \in I} [s_i] \right\} \right).$$

In case $d = 2$ we simply write $K_{(s_j)_{j \in [k]}} := K^{(2)}_{(s_j)_{j \in [k]}}$. The case $d = 1$ is trivial since the edges of the hypergraph are isomorphic to its vertices. Furthermore, in case that $\exists s \in \mathbb{N} \ \forall j \in [k] : \ s_k = s$ we want to introduce the notation

$$K^{(d)}_k(s) := K^{(d)}_{(s_k)_{j \in [k]}}.$$

Definition 24 (Equitable partition)**.** Let $k \in \mathbb{N}, \ X$ be a finite set and let $(U_j)_{j \in [k]}$ be a partition of X. We call $(U_j)_{j \in [k]}$ *equitable* if $\forall i, j \in [k] : \ ||U_j| - |U_i|| \leq 1$.

If the partition of a complete multipartite graph is equitable we call it *Turán graph*.

Definition 25 (Turán graph)**.** Let $n, k \in \mathbb{N}$. Define $r := (n \bmod k)$ and $s := \lfloor \frac{n}{k} \rfloor$ as well as $s_j := s + \mathbb{1}\{j \in [r]\}$ for $j \in [k]$. Then we set

$$T(n,k) := K_{(s_j)_{j\in[k]}}.$$

Observe that for $n \leq k$ the corresponding Turán graph is the complete graph on n vertices. For $n = k+1$ the Turán graph is a complete graph on n vertices with one missing edge. For $k = 1$ the Turán graph is the empty graph on n vertices.

Observation 1. Let $n, k \in \mathbb{N}$. Then $\left(1 - \frac{1}{k}\right)\binom{n}{2} \leq \|T(n,k)\| \leq \left(1 - \frac{1}{k}\right)\frac{n^2}{2}$.

Proof of Observation 1. For the lower bound observe that $\delta(T(n,k)) = (n-1) - \left(\lceil \frac{n}{k} \rceil - 1\right) \geq \left(1 - \frac{1}{k}\right)(n-1)$, where in the inequality we used $k\lceil \frac{n}{k} \rceil \leq n + k - 1$. Hence, by the Handshake Lemma

$$\|T(n,k)\| \geq \frac{n}{2}\delta(T(n,k)) \geq \left(1 - \frac{1}{k}\right)\binom{n}{2}.$$

For the upper bound let set $r := (n \bmod k)$ and $s := \lfloor \frac{n}{k} \rfloor$ and calculate

$$\|T(n,k)\| = \binom{n}{2} - r\binom{s+1}{2} - (k-r)\binom{s}{2},$$

where in case $n < k$ we used that $\binom{0}{2} = \binom{1}{2} = 0$.
Observe that $z \mapsto \binom{z}{2} = \frac{z(z-1)}{2}$ is a convex function. Thus, by the Definition of convexity for a real valued function

$$\frac{r}{k}\binom{s+1}{2} + \frac{k-r}{k}\binom{s}{2} \geq \binom{\frac{r(s+1)+(k-r)s}{k}}{2} = \binom{\frac{n}{k}}{2} = \frac{\frac{n}{k}\left(\frac{n}{k} - 1\right)}{2} = \frac{1}{k^2}\binom{n}{2} - \frac{n}{2k}\left(1 - \frac{1}{k}\right).$$

Hence, we obtain $\|T(n,k)\| = \binom{n}{2} - r\binom{s+1}{2} - (k-r)\binom{s}{2} \leq \left(1 - \frac{1}{k}\right)\left(\binom{n}{2} + \frac{n}{2}\right) = \left(1 - \frac{1}{k}\right)\frac{n^2}{2}.$ □

The following simple bounds on the binomial coefficient are going to be used frequently.

Observation 2 (Bounds on the binomial coefficient)**.** Let $n, k \in \mathbb{N}$ with $n \geq k$. Then $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$.

Proof of Observation 2. Observe that $\binom{n}{k} = \prod_{0 \leq j < k} \frac{n-j}{k-j}$. The lower bound now simply follows by the fact that for $j \in [k] : \frac{n-j}{k-j} \geq \frac{n}{k}$. Regarding the upper bound, using a standard series representation of the Eulerian constant, we observe

$$\binom{n}{k} = \prod_{0 \leq j < k} \frac{n-j}{k-j} \leq \frac{n^k}{k!} = \left(\frac{n}{k}\right)^k \frac{k^k}{k!} \leq \left(\frac{n}{k}\right)^k \sum_{j \in \mathbb{N}_0} \frac{k^j}{j!} = \left(\frac{n}{k}\right)^k e^k.$$ □

Observation 3 (Bernoullis inequality)**.** $\forall x \in \mathbb{R},\ x \geq -1,\ n \in \mathbb{N} :\ (1+x)^n \geq 1 + nx.$

We consider Observation 3 to be common mathematical knowledge and omit a proof. For the proof of following Theorem we again refer the reader to a standard presentation in Diestel [15].

Theorem 1 (Ramseys Theorem for uniform hypergraphs)**.** $\forall d, c \in \mathbb{N},\ (q_j)_{j\in[c]} \subseteq \mathbb{N}\ \exists R := \mathrm{R}^{(d)}\left((q_j)_{j\in[c]}\right) \in \mathbb{N}$ such that for any edge coloring using colors $[c]$ of the complete d-uniform hypergraph on at least R vertices there is some color $j \in [c]$ such that there is a monochromatic q_j-clique in color j. Formally

$$\forall n \geq R\ \forall \phi : \binom{n}{d} \longrightarrow [c]\ \ \exists j \in [c]\ \ \exists X \in \binom{[n]}{q_j}\ \forall S \in \binom{X}{d} :\ \phi(S) = j.$$

In case that $d = 2$ we call R the $(q_j)_{j\in[c]}$-*Ramsey number.* We call it *off-diagonal* in case that there are $i, j \in [c]$ such that $q_i \neq q_j$.

The proof of the following Lemma can be found in any undergraduate text book on stochastics, e.g. in [33]. We remark that in this thesis we only consider discrete random variables. We consider the Definition of a convex function as common mathematical knowledge.

Lemma 1 (Jensens inequality)**.** Let $A \subseteq \mathbb{R}$ be an interval and X be a random variable with values in A and finite first momentum as well as $f : A \to \mathbb{R}$ be a convex function. Then

$$\mathbb{E}\left[f(X)\right] \geq f(\mathbb{E}\left[X\right]).$$

2 Introduction to (induced) Turán problems

In this section we give an introduction to induced Turán problems and make preparations for our main Theorems in section 4. At first, in section 2.1 we formally introduce the forbidden subgraph problem in its most general form, giving an illustrative example along the way. In section 2.2 we present the cornerstone Theorems of Extremal Graph Theory and consider the question of how the choice of the host graph changes the forbidden subgraph problem.
In section 2.3 we present some standard techniques used in the Theorems of this thesis. Here we present the technique of Dependent Random Choice as well as the Hypergraph Removal lemma. The latter is going to be one of the main tools in the proof of our main result, Theorem 16. Afterwards, in section 2.3.3 we are going to provide a useful formalization of the probabilistic method for the forbidden subgraph problem. In section 2.3.4 we develop some counting tool for independent sets it $K_{s,s}$-free graphs, that will be crucial in the counting result Theorem 19.
In section 2.4 we are going to reduce the problem of determining $\mathrm{ex}\left(G, \{F, H\text{-ind}\}\right)$ for graphs G, F and H in case that H is no biclique and H or F are not bipartite. Furthermore, we present a result on $\mathrm{ex}\left(G, \{F, H\text{-ind}\}\right)$ in case that H is a biclique by Loh, Tait, Timmons, Zhou in [37].
In section 2.5 we present a Reduction lemma used by proofs for upper bounds of extremal functions, where we put in some effort to simplify the constants and exponents. The Reduction lemma originates from Jiang and Seiver in [34].

2.1 Definition of the forbidden subgraph problem

Let H, F and G be graphs. The basic problem of Extremal Graph Theory is to explore the subgraphs $G' \subseteq G$ such that $F \nsubseteq G'$.

Definition 26 (Free graphs)**.** $\mathrm{Free}\left(G, F\right) := \{\, G' \subseteq G \mid F \nsubseteq G' \text{ and } |G'| = |G| \,\}$. For $G \in \mathrm{Free}\left(G, F\right)$ we say that G is *F-free*. Furthermore, for $n \in \mathbb{N}$ we introduce the abbreviated notation $\mathrm{Free}\left(n, F\right) := \mathrm{Free}\left(K_n, F\right)$. In this setting we often call G the *host graph*.

The most natural and first studied problem in the area of Extremal Graph Theory is finding edgemaximal graphs given some forbidden subgraph restriction.

Definition 27 (Extremal functions)**.** $\mathrm{ex}\left(G, H\right) := \max_{G' \in \mathrm{Free}(G,H)} \|G'\|$.

We call a subgraph G' of some host graph G *extremal* if it is edge-maximal with respect of some forbidden subgraph restriction.

Definition 28 (Extremal graph)**.** $\mathrm{Ex}\left(G, H\right) := \{\, G' \in \mathrm{Free}\left(G, H\right) \mid \|G'\| = \mathrm{ex}\left(G, H\right) \,\}$.

We can generalize the forbidden subgraph restriction to an induced and biinduced version. For this purpose let $\tilde{H}$ be a bipartite graph.

Definition 29 (Induced free graphs)**.** $\mathrm{Free}\left(G, H\text{-ind}\right) := \left\{ G' \subseteq G \;\middle|\; H \underset{ind}{\nsubseteq} G' \text{ and } |G'| = |G| \right\}.$

Definition 30 (Biinduced free graphs)**.** $\mathrm{Free}\left(G, \tilde{H}\text{-biind}\right) := \left\{ G' \subseteq G \;\middle|\; \tilde{H} \underset{biind}{\nsubseteq} G' \text{ and } |G'| = |G| \right\}.$

Furthermore, we want to define the graph property of all H-free graphs

Definition 31. $\mathrm{Free}(H) := \bigcup_{n\in\mathbb{N}} \mathrm{Free}(n, H)$.

Lastly, we want to generalize the forbidden subgraph restriction to sets of graphs. Let G be a graph and $\mathscr{G}$ be a graph property.

Definition 32 (Notation for graph properties)**.** $\mathrm{Free}(G, \mathscr{G}) := \bigcap_{F\in\mathscr{G}} \mathrm{Free}(G, F)$.

We remark that we are going to use the notation freely, e.g. for two graphs F_1, F_2 and any graph G we say that a graph G' is $\left\{F_1, F_2, H\text{-ind}, \tilde{H}\text{-biind}\right\}$*-free* if it lies in $\mathrm{Free}\left(n, \left\{F_1, F_2, H\text{-ind}, \tilde{H}\text{-biind}\right\}\right)$. Furthermore, for $n \in \mathbb{N}$ we denote

$$\mathrm{ex}\left(n, \left\{F_1, F_2, H\text{-ind}, \tilde{H}\text{-biind}\right\}\right)$$

for the maximal number of edges of a graph on n vertices that neither contains F_1 or F_2 as a subgraph nor H as an induced subgraph nor $\tilde{H}$ as a biinduced subgraph. We consider the set expression simply as notation and do not care about a rigid Definition of the mathematical object $\left\{F_1, \tilde{H}\text{-biind}\right\}$. However, we are going to use the union operator to combine subgraph restrictions.

We remark that in case $H \underset{ind}{\nsubseteq} G$ the problem of determining $\mathrm{ex}(G, H\text{-ind})$ is trivial.

For a better understanding of the notation we give a simple but illustrating example, where it is possible to determine all the extremal graphs.

Example 1. Let $k, r \in \mathbb{N}_{\geq 2}$ and $G := K_k(r)$ be the complete multipartite graph on k partition classes of size r each. Then

$$\mathrm{ex}(G, \{K_{r,r}\text{-ind}\}) = \|G\| - \binom{k}{2}.$$

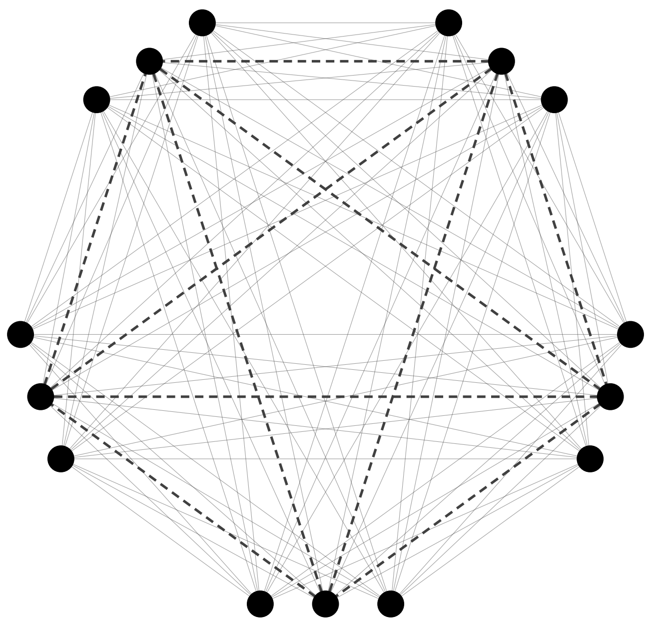

Figure 2: A graph in $\mathrm{Ex}(G, \{K_{r,r}\text{-ind}\})$ with $r = 3,\ k = 5$.

Proof of Example 1. First we may check the lower bound. For this we construct one extremal graph, see Figure 2. Let X be a set of vertices, one from each partition class of G. Let us define $G' := G - \binom{X}{2}$. Obviously $\|G'\| = \|G\| - \binom{k}{2}$. We show $K_{r,r} \underset{ind}{\nsubseteq} G'$. Assume otherwise, meaning there are disjoint independent vertex sets A, B in G of size r each such that $G[A, B]$ is complete bipartite.

case There are two vertices $a_1, a_2 \in A$ that lie in different partition classes of G. In this case we know that $a_1, a_2 \in X$. Notice that any other vertex a_3 in A has to lie outside of the partition classes of a_1 and a_2 since otherwise $\{a_1, a_3\} \in E(G')$ or $\{a_2, a_3\} \in E(G')$. By the same argument we know that all vertices of A lie in pairwise different partition classes of G. It is easy to see that $A \subseteq X$. Thus, we know that $B \subseteq V(G) \setminus X$. However, since any partition class contains one vertex of X we know that B lies in at least two partition classes, a contradiction to its independence.
case A represents a partition class of G. Since A intersects X we know that $B \subseteq V(G) \setminus X$. However, this leads to the same contradiction as in the previous case.
We remark that G is not a unique extremal graph. Indeed, it is easy to check that

$$\operatorname{Ex}(G, \{K_{r,r}\text{-ind}\}) = \{\, G - E(Q) \mid Q \in \mathscr{Q} \,\},$$

where $\mathscr{Q}$ is the set of subgraphs $Q \subseteq G$ such that for any pair of partition classes X, Y of G on has $|E(Q) \cap E(X,Y)| = 1$ and any clique in Q of order r is connected to any other clique of order r and sends one edge towards any partition class of G.
Regarding the upper bound assume for a contradiction that there would be a graph $G'' \in \operatorname{Free}(G, K_{r,r}\text{-ind})$ with $\|G''\| > \|G\| - \binom{k}{2}$. Then there would be two distinct partition classes A, B in G such that $G'[A,B]$ is complete. A contradiction. ☐

Observation 4. In case H is an empty graph, by Ramseys Theorem 1, graphs that are large enough either contain H as an induced subgraph or a clique of size $|F|$. Hence, it could happen that $\operatorname{Free}(G, \{F, H\text{-ind}\}) = \emptyset$. Regarding this case we remind the reader that we interpret the maximum of the empty set as infinity.

In case of a complete bipartite host we want to furthermore introduce an asymmetric problem. For this purpose let $\tilde{G}, \tilde{F}$ be bipartite graphs.

Definition 33 (Asymmetric free graphs)**.** $\operatorname{Free}^*\left(\tilde{G}, \tilde{F}\right) := \left\{ \tilde{G}' \subseteq^* \tilde{G} \,\middle|\, \tilde{F} \not\subseteq^* \tilde{G}' \text{ and } |\tilde{G}'| = |\tilde{G}| \right\}$.

Again we generalize the notation to induced and biinduced restrictions as well as forbidden graph properties. Furthermore, we mark the notation of the extremal function and graphs with a star to indicate that we are in the asymmetric setting. The Definitions made in this section allow us to compactly state the following simple Observations.

Observation 5. Let H, F, G be graphs with $H \subseteq F \subseteq G$ and $\tilde{H}, \tilde{F}, \tilde{G}$ be bipartite graphs with $\tilde{H} \subseteq^* \tilde{F} \subseteq^* \tilde{G}$. Then

$$\begin{aligned} \operatorname{ex}(G, H) \le \operatorname{ex}(G, \{F, H\text{-ind}\}) \quad &\le \operatorname{ex}(G, F). \\ \operatorname{ex}^*\left(\tilde{G}, \tilde{H}\right) \le \operatorname{ex}^*\left(\tilde{G}, \left\{\tilde{F}, \tilde{H}\text{-ind}\right\}\right) &\le \operatorname{ex}^*\left(\tilde{G}, \tilde{F}\right). \end{aligned}$$

Proof of Observation 5. Notice that the first line simply follows by the inclusions

$$\operatorname{Free}(G, H) \subseteq \operatorname{Free}(G, \{F, H\text{-ind}\}) \subseteq \operatorname{Free}(G, F).$$

Concerning the first inclusion consider $G' \in \operatorname{Free}(G, H)$ and assume that $G' \notin \operatorname{Free}(G, \{F, H\text{-ind}\})$. Then either $F \subseteq G'$ or $H \underset{ind}{\subseteq} G'$ both implying $H \subseteq G'$, a contradiction. Concerning the second inclusion consider $G' \in \operatorname{Free}(G, \{F, H\text{-ind}\})$. The Definition states that $F \not\subseteq G'$ which already shows $G' \in \operatorname{Free}(G, F)$. The second line of inequalities follows analogously. ☐

Observation 6 (Symmetric versus asymmetric)**.** Let $\tilde{H}, \tilde{F}$ and $\tilde{G}$ be bipartite graphs. Then

$$\begin{aligned} \operatorname{ex}\left(\tilde{G}, \tilde{H}\right) &\le \operatorname{ex}^*\left(\tilde{G}, \tilde{H}\right). \\ \operatorname{ex}\left(\tilde{G}, \left\{\tilde{F}, \tilde{H}\text{-ind}\right\}\right) &\le \operatorname{ex}^*\left(\tilde{G}, \left\{\tilde{F}, \tilde{H}\text{-ind}\right\}\right). \end{aligned}$$

Proof of Observation 6. The statement is an immediate consequence of the inclusions $\text{Free}\left(\tilde{G}, \tilde{H}\right) \subseteq \text{Free}^*\left(\tilde{G}, \tilde{H}\right)$ and $\text{Free}\left(\tilde{G}, \left\{\tilde{F}, \tilde{H}\text{-ind}\right\}\right) \subseteq \text{Free}^*\left(\tilde{G}, \left\{\tilde{F}, \tilde{H}\text{-ind}\right\}\right)$. □

2.2 Standard results about the Turán problem

2.2.1 Non-degenerate case

The general extremal function $\text{ex}(n, H)$ for non-bipartite graphs H is well-studied and understood. In the literature it is referred to as the *non-degenerate* case. When $H = K_k$ for some integer $k \geq 2$ by Turáns Theorem the single extremal graph for H in K_n is the Turán graph, see Definition 25. Two of its standard proofs can be found in Diestel [15].

Theorem 2 (Turán theorem)**.** $\forall k, n \in \mathbb{N},\ k \geq 2 :\ \text{Ex}(K_n, K_k) = \{T(n, k-1)\}$.

Corollary 1. Let $\epsilon \in (0,\ 1)$ and $n \in \mathbb{N}$. Then for any graph G on n vertices

$$\|G\| \geq (1-\epsilon)\frac{n^2}{2} \implies \omega(G) \geq \frac{1}{\epsilon}.$$

Note that ϵ has the implicit lower bound $\frac{1}{n}$ since otherwise the condition $\|G\| \geq (1-\epsilon)\frac{n^2}{2}$ could not hold.

Proof of Corollary 1. Let $k := \left\lceil \frac{1}{\epsilon} \right\rceil$. We remark that $k \geq 2$. Let us assume for a contradiction that $K_k \nsubseteq G$. Turáns Theorem together with Observation 1 imply that

$$(1-\epsilon)\frac{n^2}{2} \leq \|G\| \leq \text{ex}(K_n, K_k) = \|T(n, k-1)\| \leq \left(1 - \frac{1}{k-1}\right)\frac{n^2}{2}.$$

Thus, $\frac{1}{k-1} \leq \epsilon$ and $k \geq \frac{1}{\epsilon} + 1$, a contradiction. □

It is easy to observe that for $n \in \mathbb{N}$ an arbitrary non-empty graph H with chromatic number $\chi(H)$ can not be contained in $T(n, \chi(H) - 1)$. The Erdős, Stone, Simonovits theorem now states that the resulting lower bound for the extremal function of H is asymtotically sharp. Its standard proof can be found in Diestel [15].

Theorem 3 (Erdős, Stone, Simonovits)**.** For any non-empty graph $H :\ \text{ex}(n, H) = \left(\frac{\chi(H)-2}{\chi(H)-1} + o(1)\right)\frac{n^2}{2}\ (n \longrightarrow \infty)$.

In case that H is bipartite however the resulting bound $\text{ex}(n, H) = o\left(n^2\right)$ is not satisfactory since it does not give the exact order of magnitude of the extremal function. This case is often referred to as the *degenerate case.*

2.2.2 Degenerate case

The problem of determining the extremal function of bipartite graphs is significantly harder than for non-bipartite graphs. In most cases even the asymptotics are not known. In case that the forbidden subgraph is a biclique the problem of determining $\text{ex}^*(K_{n,n}, K_{s,t})$ is known as the *Zarankievicz problem.*
The next Lemma gives the classical bound on the asymmetrical extremal function of complete bipartite graphs.

Lemma 2 (Kővári, Sós, Turán [36])**.** Let $y_1, y_2, n_1, n_2 \in \mathbb{N}$ and $G = (Y_1 \cup Y_2,\ E)$ be a bipartite graph on partite sets of size $|Y_1| = n_1$ and $|Y_2| = n_2$ that does not contain a complete bipartite subgraph with y_1 vertices in Y_1 and y_2 vertices in Y_2, meaning $G \in \text{Free}^*(K_{n_1,n_2}, K_{y_1,y_2})$. Then

$$\|G\| \leq (y_2 - 1)^{\frac{1}{y_1}}(n_1 - y_1 + 1)(n_2)^{1-\frac{1}{y_1}} + (y_1 - 1)n_2.$$

By a simple calculation one can transform the bound to the following shape.

Observation 7. $\forall n_1, n_2, y_1, y_2 \in \mathbb{N} :\ \text{ex}^*(K_{n_1,n_2}, K_{y_1,y_2}) \leq \left((y_2)^{\frac{1}{y_1}} + \frac{y_1 (n_2)^{\frac{1}{y_1}}}{n_1}\right) n_1 (n_2)^{1-\frac{1}{y_1}}$.

There are also known lower bounds for the extremal function of special bicliques, based on so-called (projective) norm graph constructions that use certain system of equations over finite fields.

Theorem 4 (Alon Ronyai Szabo [3])**.** $\forall s \in \mathbb{N}$ with $s \geq 2$ $\exists c_s > 0$ $\forall t \in \mathbb{N}$ with $t > (s-1)!$ one has

$$\mathrm{ex}\,(n, K_{s,t}) = \left(\frac{c_s}{2} + o\,(1)\right)(t-1)^{\frac{1}{s}}\, n^{2-\frac{1}{s}} \ (n \longrightarrow \infty).$$

Theorem 5 (Füredi [22])**.** $\forall t \in \mathbb{N}:\ \mathrm{ex}\,(n, K_{2,t+1}) = \frac{\sqrt{t}}{2} n^{\frac{3}{2}} + O\left(n^{\frac{4}{3}}\right)\ (n \longrightarrow \infty)$.

2.2.3 The role of the host graph

The next Lemma gives a hint that we are fairly free in choosing the host graph when we are interested in determining the asymptotics of the extremal function in the general setting. Its proof is a simple sampling argument.

Lemma 3. Let $n \in \mathbb{N}$ and $G \subseteq K_n$ as well as $p := \frac{\|G\|}{\|K_n\|}$. Then for every graph H and every positive integer n:

$$p \cdot \mathrm{ex}\,(K_n, H) \leq \mathrm{ex}\,(G, H) \leq \mathrm{ex}\,(K_n, H)\,.$$

Proof of Lemma 3. Without loss of generality we may assume that $n \geq 2$, $|G| = n$ and further $V\,(G) = V\,(K_n) = [n]$ so the labeled copies of G in K_n correspond to the permutations $\mathscr{S}_n$. For an edge $e = \{u, v\} \in E\,(G)$ and $\sigma \in \mathscr{S}_n$ let us introduce the notation $\sigma(e) = \{\sigma(u), \sigma(v)\}$.
Consider a random labeled copy $\sigma \in \mathcal{U}\,(\mathscr{S}_n)$. Furthermore, let $G' \subseteq K_n$ be edgemaximal with respect to $H \nsubseteq G'$. Let us define a random graph

$$G'_\sigma := \left([n]\,,\ E\,(G') \cap \sigma\,(E\,(G))\right).$$

We remark that $E\,(G'_\sigma) = \left\{\, e \in E\,(G') \;\middle|\; \sigma^{-1}(e) \in E\,(G) \,\right\}$. Notice further that $\sigma \sim \sigma^{-1}$ and for all edges $e \in \binom{[n]}{2}:\ \sigma(e) \sim \mathcal{U}\left(\binom{[n]}{2}\right)$. Using this we calculate

$$\mathbb{E}\,[\|G'_\sigma\|] = \sum_{e \in E(G')} \mathbb{P}\left(\sigma^{-1}(e) \in E\,(G)\right) = \|G'\| \cdot \mathbb{P}\,(\sigma(\{1,2\}) \in E\,(G)) = \mathrm{ex}\,(K_n, H) \cdot \frac{\|G\|}{\binom{n}{2}} \geq p \cdot \mathrm{ex}\,(K_n, H)\,.$$

Hence, we find $\tau \in \mathscr{S}_n$ such that $\|G'_\tau\| \geq p \cdot \mathrm{ex}\,(K_n, H)$ and the first inequality follows by $H \nsubseteq G'_\tau$. The second inequality is trivial since $G \subseteq K_n$. ☐

In this thesis we mostly use balanced bipartite host graphs which fulfill the requirement of the last Observation with $p = \frac{1}{2}$.

Observation 8. For any bipartite graph H and $n \in \mathbb{N}$

$$\frac{\mathrm{ex}\,(K_n, H)}{2} \leq \mathrm{ex}\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}, H\right) \leq \mathrm{ex}\,(K_n, H)\,.$$

Proof of Observation 8. Notice $\|K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\| = \left\lfloor \frac{n^2}{4} \right\rfloor \geq \frac{n(n-1)}{4} = \frac{1}{2}\binom{n}{2}$, so the Observation is an immediate consequence of Lemma 3. ☐

Corollary 2 (Kővári, Sós, Turán on complete host)**.** $\forall s, t, n \in \mathbb{N}$ with $t \geq s$ and $n \geq 10:\ \mathrm{ex}\,(n, K_{s,t}) \leq t^{\frac{1}{s}} n^{2-\frac{1}{s}}$. Furthermore, there is a constant $C = C(s,t) > 0$ such that $\forall n \in \mathbb{N}:\ \mathrm{ex}\,(n, K_{s,t}) \leq Cn^{2-\frac{1}{s}}$.

Proof of Corollary 2. **case** $s = 1$. The statement is trivial.

case $s \geq 2$. In case $n \geq 10$ we use Observation 7 and 8 and simple calculations to bound

$$\begin{aligned}
\mathrm{ex}\left(K_n, K_{s,t}\right) &\leq 2\mathrm{ex}\left(K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}, K_{s,t}\right) \\
&\leq 2\left(t^{\frac{1}{s}} + \frac{t^{\frac{1}{s}} \lceil \frac{n}{2} \rceil^{\frac{1}{s}}}{\lfloor \frac{n}{2} \rfloor}\right) \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil^{1-\frac{1}{s}} \\
&\leq 2t^{\frac{1}{s}} \left(1 + 2\left(\frac{2}{n}\right)^{1-\frac{1}{s}}\right) \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil^{1-\frac{1}{s}} \\
&\leq 2t^{\frac{1}{s}} \left(1 + \frac{2}{5}\right) \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil^{1-\frac{1}{s}} \leq 2t^{\frac{1}{s}} \sqrt{2} \left\lfloor \frac{n^2}{4} \right\rfloor \left\lceil \frac{n}{2} \right\rceil^{-\frac{1}{s}} \leq t^{\frac{1}{s}} n^{2-\frac{1}{s}}.
\end{aligned}$$

Now if we set $C := \max\{10, t\}^{\frac{1}{s}}$ then trivially

$$\forall n \in [10]: \ \mathrm{ex}\left(n, K_{s,t}\right) \leq n^{\frac{1}{s}} n^{2-\frac{1}{s}} \leq C n^{2-\frac{1}{s}}$$

so the Claim holds for all $n \in \mathbb{N}$. □

We want to remark that in the induced case the simple sampling of Lemma 3 does not work anymore.

2.3 Some technies of Extremal Graph Theory

2.3.1 Dependent Random Choice

The following Lemma has many striking applications in Extremal Graph Theory. We are going to apply it in the proof of Theorem 6.

Lemma 4 (Dependent Random Choice [21])**.** Let $a, r, s, n \in \mathbb{N}$ and $G = ([n]\,,\ E)$ be a graph with $d := \mathrm{avdeg}\,(G)$. Then

$$\exists \tau \in \mathbb{N}: \ \frac{d^\tau}{n^{\tau-1}} - \binom{n}{s}\left(\frac{r}{n}\right)^\tau \geq a \implies \exists A \in \binom{[n]}{a}\ \forall U \in \binom{A}{s}: \ \deg_G(U) \geq r.$$

2.3.2 Hypergraph Removal lemma

One central tool we are going to use in the proof of our main Theorem 21 is the Hypergraph Removal lemma which we will state here. To state it in a convenient way we introduce some notation. For a hypergraph $\mathscr{H}$ and integers $d \leq q$ let us define the sets of edges, that's deletion make $\mathscr{H}$ free of copies of the d-uniform clique on q vertices.

Definition 34 (Removal edges)**.** $\mathrm{Rem}_q^{(d)}(\mathscr{H}) := \left\{ E' \subseteq E(\mathscr{H}) \;\middle|\; \binom{[q]}{d} \nsubseteq \mathscr{H} - E' \right\}$.

Furthermore, we introduce notation for the set of d-uniform q-cliques.

Definition 35 (Cliques)**.** $\mathscr{K}_q^{(d)}(\mathscr{H}) := \left\{ A \in \binom{V(\mathscr{H})}{q} \;\middle|\; \binom{A}{d} \subseteq E(\mathscr{H}) \right\}$.

The Hypergraph Removal lemma was proven independently by Nagl, Rödl, Schacht [39] and by Gowers [25].

Lemma 5 (Hypergraph Removal lemma)**.** $\forall \delta > 0,\ d, k \in \mathbb{N}\ \exists \epsilon > 0$ such that for every d-uniform hypergraph $\mathscr{H}$:

$$\min_{E' \in \mathrm{Rem}_q^{(d)}(\mathscr{H})} |E'| \geq \delta \binom{|\mathscr{H}|}{d} \implies |\mathscr{K}_q^{(d)}(\mathscr{H})| \geq \epsilon \binom{|H|}{q}.$$

2.3.3 Deletion method

Lower bounds on extremal functions often rely on the so-called *Deletion method*. For convenient application we might introduce some density ratio for graphs.

Definition 36. Let G be a graph with at least two edges. Define

$$\gamma(G) := \frac{|G|-2}{\|G\|-1}.$$

Let $\mathscr{G}$ be a graph property of graphs with at least two edges each. Then we define

$$\gamma(\mathscr{G}) := \sup_{G\in\mathscr{G}}\gamma(G).$$

The following Lemma can be found as Erdős-Rényi First Moment method, Theorem 2.26, in [24].

Lemma 6 (Deletion method [24])**.** Let $\mathscr{G}$ be a finite graph property of bipartite graphs with at least two edges each. Then

$$\mathrm{ex}(K_n, \mathscr{G}) = \Omega\left(n^{2-\mathscr{G}}\right).$$

Proof of Lemma 6. Fix $\alpha, \beta \in (0,1)$ and define $p_n := \beta n^{-\alpha}$, notice $p_n \in (0,1)$. We want to define G_n as the random graph on n vertices where each edge is sampled with probability p_n.
We can obtain a random graph $\tilde{G}_n$ that does not contain any graph of $\mathscr{G}$ as a subgraph by deleting an edge in every labeled copy of H for every $H \in \mathscr{G}$ in G_n according to some specific rule (e.g. for any labeled copy of H we delete the edge with the smallest rank in an arbitrary ordering of the edges).

$$\mathbb{E}\left[\|\tilde{G}_n\|\right] \geq \mathbb{E}\left[\|G_n\| - \sum_{H\in\mathscr{G}} |\mathrm{Isom}_{\mathrm{ind}}(H,\ G)|\right] \geq \binom{n}{2}p_n - \sum_{H\in\mathscr{G}} n^{|H|}p_n^{\|H\|} \geq \frac{\beta}{4}n^{2-\alpha} - \sum_{H\in\mathscr{G}} \beta^{\|H\|}n^{|H|-\alpha\|H\|}.$$

Let us asymtotically maximize this lower bound over $\alpha > 0$. This is achieved in case that the leading term of the subtrahend has exponent $2-\alpha$. Observe that for any $H \in \mathscr{G}$

$$2-\alpha = |H| - \alpha\|H\| \Longleftrightarrow \alpha = \frac{|H|-2}{\|H\|-1} = \gamma(H).$$

Thus, let us choose $\alpha := \gamma(\mathscr{G})$ and $\beta > 0$ small enough such that $\frac{\beta}{4} > \sum_{H\in\mathscr{G}} \beta^{\|H\|}$ which is possible since $\mathscr{G}$ is finite and the edgecount of any graph in $\mathscr{G}$ is at least two. Since for any $H \in \mathscr{G}$ we have that $|H| - \gamma(\mathscr{G})\|H\| \leq 2-\gamma(\mathscr{G})$ since $\frac{|H|-2}{\|H\|-1} \leq \gamma(\mathscr{G})$ it follows that

$$\mathbb{E}\left[\|\tilde{G}_n\|\right] \geq \left(\frac{\beta}{4} - \sum_{H\in\mathscr{G}} \beta^{\|H\|}\right) n^{2-\gamma(\mathscr{G})} = \Omega\left(n^{2-\gamma(\mathscr{G})}\right).$$ ☐

2.3.4 Independent sets in $K_{s,s}$-free graphs

For finding induced copies of bipartite graphs in a host it is essential to guarantee for large independent sets. When counting such copies in Theorem 19 we also need some counting results for small independent sets. In this section we want to draw the consequences of the following simple Observation.

Observation 9 (Union bound for neighborhoods)**.** Let G be a graph. For $r \in \mathbb{N}$ let us denote $V_{\mathrm{interlace}}(r) := \{\, v \in V(G) \mid \deg_G(v) \geq |G| - r \,\}$. Then

$$\forall r, s, t \in \mathbb{N} \text{ fulfilling } s \leq |V_{\mathrm{interlace}}(r)|,\ t \leq |G| - sr :\ K_{s,t} \subseteq G.$$

Proof of Observation 9. Let $r, s, t \in \mathbb{N}$ fulfilling the given requirements. Choose $A \in \binom{V_{\mathrm{interlace}}(r)}{s}$ and

$$B := V(G) \setminus \left(\bigcup_{a\in A} \left(V(G) \setminus N_G(a)\right)\right).$$

Then since $|B| \geq |G| - sr$ and $G[A, B]$ is complete bipartite we conclude that $K_{s,t} \subseteq G$. ☐

Definition 37 (Independent embeddings)**.** Let $l \in \mathbb{N}$ and G be a graph. We define

$$\mathscr{I}_l(G) := \left\{ (v_j)_{j\in[l]} \subseteq V(G) \;\middle|\; \{ v_j \mid j \in [l] \} \text{ independent set of order } l \right\}.$$

Lemma 7. Let $s, l \in \mathbb{N}$ and G be a $K_{s,s}$-free graph on at least $4s^{l-1}$ vertices. Then

$$|\mathscr{I}_l(G)| \geq 2^{-(l+1)} \left(\sqrt{s}\right)^{-l^2} |G|^l.$$

Proof of Lemma 7. First, we want to remark that we may assume that $s \geq 2$ since in case $s = 1$ the graph G is empty. We make use of the help function $\phi^1(n) := \left\lfloor \frac{n-s}{s} \right\rfloor$ $(n \in \mathbb{N})$ which we chain to $\phi^0 :=$ id and $\phi^k := \phi^{k-1} \circ \phi,\ k \in \mathbb{N}$. Let us show the following by induction on k.

$$(*)\ \forall k, n \in \mathbb{N},\ G \in \operatorname{Free}(n, K_{s,s}):\ |\mathscr{I}_k(G)| \geq \phi^{k-1}(n) \cdot \prod_{j\in[k-1]} \left(\phi^{j-1}(n) - s + 1\right).$$

base $k = 1$**.** This case is trivial since every single-vertex set is independent. Notice that we interpret the empty product to have value one.
step $k \geq 2$**.** Define $r := \left\lfloor \frac{n-s}{s} \right\rfloor = \phi(n)$ and recognize that in notation of Observation 9 we have $|V_{\text{interlace}}(r)| < s$ since otherwise we would have $K_{s,s} \subseteq G$. Observe the following inclusion.

$$\{ \{v\} \times \mathscr{I}_{k-1}(G[V(G) \setminus N_G(v)]) \mid v \in V(G) \setminus V_{\text{interlace}}(G) \} \subseteq \mathscr{I}_k(G).$$

Now induction yields that

$$\begin{aligned}\forall v \in V(G) \setminus V_{\text{interlace}}(G):\ |\mathscr{I}_{k-1}(G[V(G) \setminus N_G(v)])| &\geq \phi^{k-2}(r) \cdot \prod_{j\in[k-2]} \left(\phi^{j-1}(r) - s + 1\right) \\ &= \phi^{k-1}(n) \cdot \prod_{j\in[k-1]\setminus\{1\}} \left(\phi^{j-1}(n) - s + 1\right).\end{aligned}$$

where we used that by Definition $|V(G) \setminus N_G(v)| \geq r$. Now the step follows since

$$|V(G) \setminus V_{\text{interlace}}(G)| \geq \phi^0(n) - s + 1.$$

Furthermore, using that we assumed $s \geq 2$, we inductively see that for any $q \in [k]$

$$\phi^q(x) = \left\lfloor \frac{x}{s^q} - \sum_{0\leq j<q} \frac{1}{s^j} \right\rfloor = \left\lfloor \frac{x}{s^q} - \frac{1 - \left(\frac{1}{s}\right)^q}{1 - \frac{1}{s}} \right\rfloor \geq \frac{x}{s^q} - 3.$$

Using that $|G| \geq 4s^{k-1}$, which implies $\phi^{k-1}(|G|) \geq \frac{|G|}{s^{k-1}} - 3 \geq \frac{|G|}{4s^{k-1}}$ and $\forall 0 \leq j < k-1:\ \phi^j(|G|) - s \geq \frac{|G|}{2s^j}$, we conclude

$$\phi^{k-1}(|G|) \prod_{j\in[k-1]} \left(\phi^{j-1}(|G|) - s\right) \geq 2^{-(k+1)} s^{-\left(\sum_{j\in[k-1]} j\right)} |G| = 2^{-(k+1)} s^{\binom{k}{2}} |G|^k \geq 2^{-(k+1)} \left(\sqrt{s}\right)^{-k^2} |G|^k. \quad \square$$

A similar inductive argument yields the following.

Lemma 8 (Bonamy et al. [7])**.** Let $s, d, n \in \mathbb{N}$ and $G = ([n],\ E)$ be a graph with $K_{s,s} \nsubseteq G$. Then for any sequence of pairwise disjoint subsets $(V_j)_{j\in[d]} \subseteq \binom{[n]}{s^{d-1}}$ there is a sequence of independent vertices $(v_j)_{j\in[d]} \in \bigtimes_{j\in[d]} V_j$.

2.4 Induced Turán problem for non-bipartite graphs

Let $n \in \mathbb{N}$ and H, F be graphs where H is non-empty. It turns out that the only cases in which one not could easily determine the asymptotics of $\mathrm{ex}\,(K_n, \{F, H\text{-ind}\})$ are

(I) H and F are bipartite and H is non-empty.

(II) H is a biclique and F is non-bipartite.

Section 4 is determined to better understand the case I. We give a slight refinement of the result presented in [30].

Lemma 9 (Illingworth [30])**.** Let F, H be graphs where H is non-empty.

(i) In case that $\chi(F) \geq \chi(H) \geq 3$ and H is complete multipartite one has

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) = \left(\frac{\chi(H) - 2}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

(ii) In case that $\chi(H) > \chi(F) \geq 3$ one has

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) = \left(\frac{\chi(F) - 2}{\chi(F) - 1} + o(1)\right)\binom{n}{2}.$$

(iii) In case that F is bipartite and H is not one has

$$\frac{\mathrm{ex}\,(n, F)}{2} \leq \mathrm{ex}\,(n, \{F, H\text{-ind}\}) \leq \mathrm{ex}\,(n, F)\,.$$

(iv) In case that F is non-bipartite and H is not complete multipartite one has

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) = \left(\frac{\chi(F) - 2}{\chi(F) - 1} + o(1)\right)\binom{n}{2}.$$

Proof of Lemma 9. **Ad** (i): Let Γ be the possibly off diagonal $(|F|,\ |H|)$-Ramsey number, meaning that any subgraph of K_Γ either contains a clique of size $|F|$ or an independent set of size $|H|$. Let us define $T \coloneqq K_{\chi(H)}(\Gamma)$ to be the complete $\chi(H)$-partite graph in which all partition classes have size Γ. We observe that for any graph G that has T as a subgraph, either $\omega(G) \geq |F|$, implying that $F \subseteq H$, or $K_{\chi(H)}(|H|) \underset{ind}{\subseteq} G$. In the latter case $H \underset{ind}{\subseteq} G$. A contradiction. Thus, we have

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) \leq \mathrm{ex}\,(n, T)\,.$$

We obtain the claimed upper bound using the Erdős, Stone, Simonovits theorem 3 and the fact that $\chi(T) = \chi(H)$.

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) \leq \left(\frac{\chi(H) - 2}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

Regarding the lower bound we observe that $H, F \nsubseteq T(n, \chi(H) - 1)$. Thus, we conclude, using Observation 1

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) \geq \|T(n, \chi(H) - 1)\| \geq \left(\frac{\chi(H) - 2}{\chi(H) - 1}\right)\binom{n}{2}.$$

Ad (ii): We observe that $H \underset{ind}{\nsubseteq} T(n, \chi(F) - 1)$, which already shows the lower bound

$$\mathrm{ex}\,(n, \{F, H\text{-ind}\}) \geq \|T(n, \chi(F) - 1)\| \geq \left(\frac{\chi(F) - 2}{\chi(F) - 1}\right)\binom{n}{2},$$

where we again used Observation 1. For the upper bound we again use the Erdős, Stone, Simonovits theorem 3. We deduce with Observation 5

$$\text{ex}\left(n,\{F,H\text{-ind}\}\right)\leq \text{ex}\left(n,F\right)=\left(\frac{\chi\left(F\right)-2}{\chi\left(F\right)-1}+o\left(1\right)\right)\binom{n}{2}.$$

Ad (iii): We observe that Free $\left(K_{\lceil\frac{n}{2}\rceil,\lfloor\frac{n}{2}\rfloor},F\right)\subseteq\text{Free}\left(K_n,\{F,H\text{-ind}\}\right)$. Thus, with Observations 5 and 8 we deduce

$$\frac{\text{ex}\left(n,F\right)}{2}\leq \text{ex}\left(K_{\lceil\frac{n}{2}\rceil,\lfloor\frac{n}{2}\rfloor},F\right)\leq \text{ex}\left(K_n,\{F,H\text{-ind}\}\right)\leq \text{ex}\left(n,F\right).$$

Ad (iv): As in case (ii) we observe that $H\underset{ind}{\not\subseteq}T\left(n,\chi\left(F\right)-1\right)$. Indeed, if there were an induced copy of H in $T\left(n,\chi\left(F\right)-1\right)$, since H it is non-empty, it must contain vertices of different partition classes of $T\left(n,\chi\left(F\right)-1\right)$. Hence, H must be connected. Since H is not complete multipartite this implies that $K_2+K_1\underset{ind}{\subseteq}H$. However, this is a contradiction to the fact that $K_2+K_1\underset{ind}{\not\subseteq}T\left(n,\chi\left(F\right)-1\right)$. Now (iv) follows exactly as (ii). □

Lemma 10 (Induced Turán theorem)**.** Let $n,k\in\mathbb{N}$ and H be a graph. In case that either H is not complete multipartite or $\chi\left(H\right)\geq k$ it turns out that

$$\text{Ex}\left(n,\{K_k,H\text{-ind}\}\right)=\{T\left(n,k-1\right)\}.$$

Proof of Lemma 10. Turáns Theorem 2 states that $\text{Ex}\left(n,K_k\right)=\{T\left(n,k-1\right)\}$. Thus, the Claim follows from the fact that $H\underset{ind}{\not\subseteq}T\left(n,k-1\right)$, which we showed in the proof of the previous Lemma 9. □

Concerned about II we present the following Theorem.

Theorem 6 (Loh, Tait, Timmons, Zhou [37])**.** For any $r,s,t\in\mathbb{N}$ with $s\leq t$ and $r\geq 3$ we have

$$\text{ex}\left(n,\{K_r,K_{s,t}\text{-ind}\}\right)=O\left(n^{2-\frac{1}{s}}\right).$$

For illustrative purposes we want to give the simple proof sketched in [37] using the method of Dependent Random Choice, Lemma 4, that yields worse constants than the considered proof presented in the same paper.

Proof of Theorem 6. Let us define Γ_s,Γ_t as the possibly off-diagonal $(r,\ s)$-, $(r,\ t)$-Ramsey numbers respectively, meaning any subgraph of K_{Γ_s} either contains a clique of size r or an independent set of size s.
Let us assume for a contradiction that $\forall n\in\mathbb{N}:\ \text{ex}\left(n,\{K_r,K_{s,t}\text{-ind}\}\right)>\left(\frac{\left(\Gamma_s\right)^{\frac{1}{s}}}{2}+\frac{\Gamma_t e}{2s}\right)n^{2-\frac{1}{s}}$. Then for large enough $n\in\mathbb{N}$ there is a graph $G\in\text{Free}\left(n,\{K_r,K_{s,t}\text{-ind}\}\right)$ with $\text{avdeg}\left(G\right)\geq\left(\left(\Gamma_s\right)^{\frac{1}{s}}+\frac{\Gamma_t e}{s}\right)n^{1-\frac{1}{s}}$. We want to find a set $A\in\binom{V(G)}{\Gamma_s}$ such that $\forall U\in\binom{A}{s}:\ \deg_G\left(U\right)\geq\Gamma_r$.
When we have found such set A then by the Definition of Ramsey numbers either $\omega\left(G\right)\geq r$ or we can find $A'\in\binom{A}{s}$ that is independent. In the latter case, since $\deg_G\left(A'\right)\geq\Gamma_t$ either $\omega\left(G\right)\geq r$ or we find an independent set $B'\in\binom{N_G(A')}{t}$. In the latter case we have found an induced copy of $K_{s,t}$ in G. In all other cases we have found a copy of K_r in G. This is a contradiction to $G\in\text{Free}\left(n,\{K_r,K_{s,t}\text{-ind}\}\right)$.
We prove existence of the set A with help of Dependent Random Choice, Lemma 4. Let us check the condition of the Lemma with $\tau:=s$.

$$\frac{\text{avdeg}\left(G\right)^{\tau}}{n^{\tau-1}}-\binom{n}{s}\left(\frac{\Gamma_t}{n}\right)^{\tau}\geq\left(\left(\Gamma_s\right)^{\frac{1}{s}}+\frac{\Gamma_t e}{s}\right)^s n^{s\left(1-\frac{1}{s}\right)-(s-1)}-\left(\frac{ne}{s}\right)^s\left(\frac{\Gamma_t}{n}\right)^s=\left(\left(\Gamma_s\right)^{\frac{1}{s}}+\frac{\Gamma_t e}{s}\right)^s-\left(\frac{\Gamma_t e}{s}\right)^s\geq\Gamma_s.$$

□

2.5 Reduction lemma

In this section we present a Reduction lemma used by proofs of upper bounds for extremal functions. First we need to introduce some vocabulary regarding bipartite graphs. For this purpose let $G=\left(A\cup B,\ E\right)$ be a bipartite graph and $K > 0$.

Definition 38 (K-almost regularity)**.** We call G a K-almost regular graph when $\frac{\Delta(G)}{\delta(G)} \leq K$.

Lemma 11 (Reduction lemma)**.** $\forall \alpha \in (0,1)\,,\ \beta \in (0,\alpha)\,,\ C \in \mathbb{R}_+\ \exists N \in \mathbb{N}\ K \in \left(0, 4^{1+\frac{1}{\alpha-\beta}}\right),\ \tilde{C} \in \mathbb{R}_+\ \forall n \in \mathbb{N}$ with $\geq N$ and for all graphs $G = ([n]\,,\ E)$ with $\|G\| \geq Cn^{1+\alpha}$ we find an induced K-almost regular subgraph H with $|H| \geq n^\beta$ and $\|H\| \geq \tilde{C}|H|^{1+\alpha}$.

Our Lemma is based on a proof given by Conlon, Janzer and Lee for Lemma 2.2 in [10]. However, they themselves are referring to the origin of the Reduction lemma, namely Proposition 2.7 in [34]. We put some effort into making the constants more convenient to use. Note that in the original version the Claim does not include that the found subgraph is induced, which however is obvious from the construction.

Proof of Lemma 11. We give an algorithm to find the claimed induced subgraph.
First let us fix some constants and show that the choices fulfill all necessary technical inequalities. Consider the function

$$f : \left(0, \frac{\alpha}{1-\alpha}\right) \to \left(4^{\frac{1}{\alpha}}, \infty\right)$$

$$x \mapsto \exp\left(\frac{\ln(4)}{\alpha-(1-\alpha)x}\right) = \exp\left(\frac{\ln(4)}{(1-\alpha)\left(\frac{\alpha}{1-\alpha}-x\right)}\right)$$

Notice that f is a strictly growing function and $f(x) \longrightarrow \infty$ whenever $x \longrightarrow \frac{\alpha}{1-\alpha}$. Using $1-\alpha < 1-\beta$ we calculate

$$4^{\frac{1}{\alpha-\beta}} = \exp\left(\ln(4)\left(\frac{1}{\alpha-\beta}\right)\right) > \exp\left(\ln(4)\left(\frac{1-\beta}{(1-\beta)\alpha-(1-\alpha)\beta}\right)\right) = f(\frac{\beta}{1-\beta}).$$

By continuity, we can choose $\gamma \in \left(\frac{\beta}{1-\beta}, \frac{\alpha}{1-\alpha}\right)$ such that when we define $p := f(\gamma)$ then $p < 4^{\frac{1}{\alpha-\beta}}$.
Observe further that for this choice we have $\frac{\gamma}{\gamma+1} > \beta$ and $p^\alpha = \exp\left(\ln(4)\left(\frac{\frac{\alpha}{1-\alpha}}{\frac{\alpha}{1-\alpha}-\gamma}\right)\right) > 4$ and

$$\frac{p}{4} = \exp\left(\ln(4)\left(\frac{1}{\alpha-(1-\alpha)\gamma}-1\right)\right) = \exp\left(\ln(4)\left(\frac{(1-\alpha)(1+\gamma)}{\alpha-(1-\alpha)\gamma}\right)\right) = p^{(1-\alpha)(1+\gamma)}$$

Furthermore, we are able to choose $\rho \in \left(2p4^{-\left(1+\frac{1}{\alpha-\beta}\right)}, \frac{1}{2}\right)$. Then $K := \frac{2p}{\rho}$ fulfills $K < 4^{1+\frac{1}{\alpha-\beta}}$.

Now we want to describe the algorithm itself. We are going to construct a sequence of graphs starting with $G_0 := G$. Assume we are in step $s \geq 0$. Let us denote $n_s := |G_s|$.
Sort the vertices of G_s with respect to their degree in descending order and divide them into an equitable partition $(B_j)_{j\in[2p]}$ such that $\forall 1 \leq i < j \leq 2p$: $\min_{u\in B_i}\deg_{G_s}(u) \geq \max_{v\in B_j}\deg_{G_s}(v)$. By adjusting the distribution of vertices along the classes we may assume $|B_1| = \left\lfloor \frac{n_s}{2p} \right\rfloor + \mathbb{1}\left\{\frac{n_s}{2p} \notin \mathbb{N}\right\} \geq \frac{n_s}{2p}$. Let us consider the number of edges adjacent to vertices in the partition class containing the highest degree vertices which we denote by $\tilde{m}_s := |\{\, e \in E(G_s) \mid e \cap B_1 \neq \emptyset \,\}|$. We want to compare it with $m_s := \|G_s\|$.
case $\tilde{m}_s \leq \frac{m_s}{2}$**.** We know that not too many vertices have a too high degree. We can find our claimed induced subgraph inside of $G'_s := G_s - B_1$. Concerning the maximal degree in G'_s we know that

$$\Delta(G'_s) \leq \min_{b\in B_1}\deg_{G_s}(b) \leq \frac{2\tilde{m}_s}{|B_1|} \leq \frac{4p\tilde{m}_s}{n_s} \leq \frac{2pm_s}{n_s}.$$

Furthermore, we know that

$$\|G'_s\| = m_s - \tilde{m}_s \geq \frac{m_s}{2}.$$

Concerning the minimal degree we successively remove vertices v from G'_s whose degree inside the graph at the given stage is less than $\rho\frac{m_s}{n_s}$.
Define H to be the graph we have obtained when we can not continue with our procedure because there are no low degree vertices left. By $\rho < \frac{1}{2}$ observe that H is nontrivial.

$$\|H\| \geq \|G'_s\| - n_s\rho\frac{m_s}{n_s} \geq \frac{1-2\rho}{2}m_s$$

Using $\Delta(H) \leq \Delta(G'_s)$ and $\delta(H) \geq \rho\frac{m_s}{n_s}$ we conclude that H is K-almost regular.

$$\frac{\Delta(H)}{\delta(H)} \leq \frac{2pm_s}{n_s}\frac{n_s}{\rho m_s} = \frac{2p}{\rho} = K.$$

Finally, let us show that H has many vertices left.

$$|H| \geq \frac{2\|H\|}{\Delta(H)} \geq \frac{1-2\rho}{2}m_s\frac{n}{2pm_s} = \Omega(n_s).$$

case $\tilde{m}_s > \frac{m_s}{2}$. We want to repeat the case analysis on some induced subgraph $G_{s+1} \underset{ind}{\subseteq} G_s$. Since we want G_{s+1} to have many edges it sounds plausible to define $G_{s+1} \coloneqq G_s[B_1 \cup B_j]$ for some $2 \leq j \leq 2p$ maximizing the edges. The pigeonhole principle yields $2 \leq j \leq 2p$ such that

$$\|B_1, B_j\| \geq \frac{1}{2p-1}\|B_1, V(G_s) \setminus B_1\|$$

so with this choice for j we obtain

$$\begin{aligned}
\|G_s[B_1 \cup B_j]\| &\geq \|G_s[B_1]\| + \|B_1, B_j\| \\
&\geq \|G_s[B_1]\| + \frac{1}{2p-1}\|B_1, V(G_s) \setminus B_1\| \\
&\geq \frac{1}{2p}\left(\|G_s[B_1]\| + \|B_1, V(G_s) \setminus B_1\|\right) = \frac{\tilde{m}_s}{2p} > \frac{m_s}{4p}.
\end{aligned}$$

Now we repeat the case analysis on G_{s+1}. Assume that in step k the graph G_k fulfills the requirement for the first case the first time, i.e. we have found

$$G_k \underset{ind}{\subseteq} G_{k-1} \underset{ind}{\subseteq} \dots \underset{ind}{\subseteq} G_0 = G$$

such that for all $0 \leq s < k$

(i) $\|G_{s+1}\| > \frac{\|G_s\|}{4p}$.

(ii) $|G_{s+1}| \geq 2\left\lfloor\frac{|G_s|}{2p}\right\rfloor$.

(iii) $|G_{s+1}| \leq 2\left\lceil\frac{|G_s|}{2p}\right\rceil$.

From (i) we inductively conclude that $m_k \geq \frac{\|G\|}{(4p)^k}$ and using our assumptions on G as well as $p^\alpha > 4$ it follows that

$$m_k \geq \frac{C}{(4p)^k}n^{1+\alpha} > \frac{C}{(p^{1+\alpha})^k}n^{1+\alpha} = C\left(\frac{n}{p^k}\right)^{1+\alpha}.$$

From (ii) we inductively conclude that

$$|G_k| \geq 2\left\lfloor\frac{\left\lfloor\frac{|G_{k-2}|}{2p}\right\rfloor}{p}\right\rfloor = 2\left\lfloor\frac{|G_{k-2}|}{2p^2}\right\rfloor \geq 2\left\lfloor\frac{n}{2p^k}\right\rfloor \geq \frac{n}{p^k} - 2.$$

Analogously, from (iii) we deduce that

$$|G_k| \leq \frac{n}{p^k} + 2.$$

For any fixed $\eta > 0$ we have in case n is large enough that $(1+\eta)\left(\frac{n}{p^k}\right)^2 \geq \left(\frac{n}{p^k}+2\right)^2$. We deduce

$$(1+\eta)\left(\frac{n}{p^k}\right)^2 \geq \left(\frac{n}{p^k}+2\right)^2 \geq |G_k|^2 \geq m_k \geq \frac{C}{(4p)^k} n^{1+\alpha}.$$

Furthermore, using the identity $p^{(1-\alpha)(1+\gamma)} = \frac{p}{4}$ we calculate

$$p^{k(1-\alpha)(1+\gamma)} = \left(\frac{p}{4}\right)^k \leq \frac{1+\eta}{C} n^{1-\alpha}.$$

This however shows that $p^k = O\left(n^{\frac{1}{1+\gamma}}\right)$ which in turn implies $|G_k| \geq \frac{n}{p^k} - 4 = \Omega\left(n^{\frac{\gamma}{1+\gamma}}\right) = \omega\left(n^\beta\right)$. Using $m_k > C\left(\frac{n}{p^k}\right)^{1+\alpha}$ and $n_k \geq \frac{n}{p^k} + 4$ we bound

$$m_k > \frac{C}{(4p)^k} n^{1+\alpha} \geq \frac{C}{\left(p^{1+\alpha}\right)^k} n^{1+\alpha} = C\left(\frac{n}{p^k}\right)^{1+\alpha} = \Omega\left(n_k^{1+\alpha}\right).$$

Applying the arguments of the first case we find $\tilde{C} > 0$ and a K-almost regular subgraph $H \underset{ind}{\subseteq} G_k$ with $|H| = \Omega\left(|G_k|\right) = \omega\left(n^\beta\right)$ as well as $\|H\| \geq \tilde{C}|H|^{1+\alpha}$. □

3 Vapnik Chervonenkis dimension

In this section we are going to introduce an important complexity measure for set systems, the so-called *Vapnik Chervonenkis dimension* or short *VC dimension*. Its study will lead us to the celebrated *Packing lemma*, see section 3.4, that will be crucial in our counting framework for induced isomorphisms in section 4.3. Furthermore, it will be the main tool in the proof of the *Ultra Strong Regularity lemma* for graphs with bounded VC dimension in section 5.2, that we need in the proof of the Erdős-Hajnal conjecture for graphs with bounded VC dimension.

It turns out to be useful to apply the results developed in this section to the set of neighborhoods in a simple graph, see section 3.2. Many famous Conjectures of Extremal Graph Theory have recently been proven correct, when the set of neighborhoods of the vertices of the considered graphs have bounded VC dimension. For example Fox, Pach and Suk proved the Schur-Erdős Conjecture in this setting, see [19]. Motivated by the results in section 5 about the Erdős-Hajnal conjecture we are going to study hereditary graph properties of bounded VC dimension in section 3.2. We found a short proof for the fact, that hereditary graph properties of unbounded VC dimension contain either all the bipartite, all the co-bipartite or all the split graphs, see Theorem 9. Furthermore, we are going to introduce VC dimension for bipartite graphs, building the bridge from this section to our main result Theorem 16.

The complexity measure has been introduced originally by Vapnik and Chervonenkis in 1968, see [46] for a recent translation of the original paper in Russian [45]. The VC dimension is an important quantity in statistical learning theory, see [6], as well as in Computational Geometry. The latter is concerned with the VC dimension of natural occurring set systems such as intersection hypergraphs, see section 3.3. As an interesting case study we present the asymptotics of the VC dimension of the k-fold unions of halfspaces in section 3.3.1.

3.1 Definitions and Introduction

In this section let X be a possibly infinite set and $\mathscr{F} = (X,\ \mathscr{E})$ be a hypergraph on X. Notice that $\mathscr{E}$ as well as the edges themselves could be infinite. The following notions are often introduced on the set system $\mathscr{E}$ rather as on a hypegraph. We chose the latter option for clarity.

Definition 39 (Trace)**.** For $S \subseteq X$ let us introduce the notation $\mathscr{E} \cap S \coloneqq \{\, A \cap S \mid A \in \mathscr{E} \,\}$. With this we define the *trace of* S by

$$\mathscr{F}\big|_S \coloneqq (S,\ S \cap \mathscr{E})\,.$$

We point out the difference to $\mathscr{F}[S] = (S,\ \{\, A \in \mathscr{E} \mid A \subseteq S \,\})$.

Definition 40 (Shatter)**.** Given $S \subseteq X$ we say that $\mathscr{F}$ *shatters* S in case $\mathscr{E} \cap S = 2^S$.

Definition 41 (Shattered sets)**.** Let us introduce the notion $\text{Shatter}(\mathscr{F}) \coloneqq \{\, S \subseteq X \mid \mathscr{F} \textit{ shatters } S \,\}$.

Definition 42 (Shatter function)**.** For $z \in \mathbb{N}$ let us define $\pi_{\mathscr{F}}(z) \coloneqq \sup_{S \in \binom{X}{z}} |\mathscr{E} \cap S|$.

The Vapnik Chervonenkis dimension determines the overall local complexity of the hypergraph $\mathscr{F}$.

Definition 43 (Vapnik Chervonenkis dimension)**.** $\dim_{\text{VC}}(\mathscr{F}) \coloneqq \max_{S \in \text{Shatter}(\mathscr{F})} |S|$.

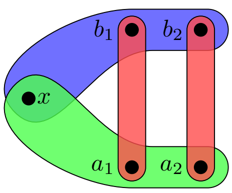

Figure 3: A small example hypergraph $\mathscr{F}$ of Vapnik Chervonenkis dimension 2.

Example 2 (VC dimension)**.** Let us define an exemplary hypergraph that is depicted in Figure 3.

$$\mathscr{F} := (\{x, a_1, a_2, b_1, b_2\}\,,\ \{\{x, a_1, a_2\}\,,\ \{x, b_1, b_2\}\,,\ \{a_1, b_1\}\,,\ \{a_2, b_2\}\})\,.$$

Let us list its shattered sets.

$$\text{Shatter}\,(\mathscr{F}) = \{\emptyset\} \cup \binom{V\,(\mathscr{F})}{1} \cup \left(\binom{V\,(\mathscr{F})}{2} \setminus \{\{a_1, b_2\}\,,\ \{b_1, a_2\}\}\right).$$

We conclude that $\dim_{\text{VC}}(\mathscr{F}) = 2$.

3.1.1 Sauer lemma

To get a feeling of the Vapnik Chervonenkis dimension we prove some standard results. The technique used in the proof of the following Lemma is taken from [26].

Lemma 12 (Pajors strengthening of the Sauer lemma)**.** $|\text{Shatter}\,(\mathscr{F})\,| \geq \|\mathscr{F}\|$.

Proof of Lemma 12. For sake of understanding the Lemma we introduce a shift function, operating on the edges. It will help us find many shattered sets. For fixed $x \in X$ the *shift operator of* x removes x from an edge if the resulting edge has not been in $\mathscr{F}$ before.

$$\begin{aligned} \text{shift}_{\mathscr{F}}^{(x)} :\ & \mathscr{E} \to 2^X \\ F \mapsto & \begin{cases} F \setminus \{x\} & F \setminus \{x\} \notin \mathscr{E} \\ F & F \setminus \{x\} \in \mathscr{E} \end{cases} \end{aligned}$$

Let us apply the shift operator of x on all edges of the hypergraph simultaneously to obtain $\mathscr{F}' := \left(X,\ \text{shift}_{\mathscr{F}}^{(x)}(\mathscr{E})\right)$. Notice that shifting has left the number of edges invariant: $\|\mathscr{F}'\| = \|\mathscr{F}\|$. Furthermore, it didn't produce any new shattered sets.

Claim 1. $\text{Shatter}\,(\mathscr{F}') \subseteq \text{Shatter}\,(\mathscr{F})$.

Proof of Claim 1. For sake of quickly verifying this Claim let $S \in \text{Shatter}\,(\mathscr{F}')$. We want to show that for any $U \subseteq S$ there is $H \in \mathscr{E}$ such that $U = H \cap S$. We know that there is $F' \in E\,(\mathscr{F}')$ such that $F' \cap S = U$ as well as $F \in \mathscr{E}$ such that $\text{shift}_{\mathscr{F}}^{(x)}(F) = F'$. In case $F' = F$ we can set $H = F$, and we are done. Otherwise, there is $x \in X$ such that $F = F' \,\dot\cup\, \{x\}$ and in case $x \notin S$ we can also set $H = F$. Otherwise, we know that $x \in S \setminus U$. There has to be $H' \in E\,(\mathscr{F}')$ such that $H' \cap S = U \cup \{x\}$. By Definition of the shift we know that $H := H' \setminus \{x\} \in \mathscr{E}$. Together with $H \cap S = U$ this yields Claim 1.
Note that the inclusion can happen to be a real one. Consider the hypergraph

$$\mathscr{F} := \left([4]\,,\ \left\{\ A \cup \{3, 4\}\ \middle|\ A \in 2^{\{1,2\}}\ \right\} \cup 2^{\{1,2\}}\right).$$

Then $\mathscr{F}' := \left([4]\,,\ \text{shift}_{\mathscr{F}}^{(3)}(\mathscr{F})\right)$ has edge set $E\,(\mathscr{F}') = \left\{\ A \cup \{4\}\ \middle|\ A \in 2^{\{1,2\}}\ \right\} \cup 2^{\{1,2\}}$. We observe that $\{1, 2, 3\}$ is shattered by $\mathscr{F}$ but not by $\mathscr{F}'$. □

By successively shifting with different $x \in X$ we will arrive at some hypergraph $\tilde{\mathscr{F}}$ that is invariant under any possible shift operation. This follows by the fact that every non-trivial shift on all edges decreases the sum of all edge sizes of the hypergraph.
Observe that $E\left(\tilde{\mathscr{F}}\right)$ is downwards closed, formally

$$\forall F \in E\left(\tilde{\mathscr{F}}\right)\ \forall x \in F :\ F \setminus \{x\} \in E\left(\tilde{\mathscr{F}}\right).$$

From a different perspective this means that every hyperedge of $\tilde{\mathscr{F}}$ is shattered by $\tilde{\mathscr{F}}$ which implies

$$E\left(\tilde{\mathscr{F}}\right) = \text{Shatter}\left(\tilde{\mathscr{F}}\right) \subseteq \text{Shatter}\left(\mathscr{F}\right).$$

The result now simply follows by $\|\tilde{\mathscr{F}}\| = \|\mathscr{F}\|$. □

We can use this result for bounding the shatter function of $\mathscr{F}$.

Corollary 3 (Sauer lemma)**.** $\forall z \in \mathbb{N}: \ \pi_{\mathscr{F}}(z) \leq \sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{F})} \binom{z}{j} \leq ez^{\dim_{\mathrm{VC}}(\mathscr{F})}$

Proof of Corollary 3. Let $z \in \mathbb{N}$ with $z \leq |\mathscr{F}|$ and $S \in \binom{X}{z}$. Notice $\dim_{\mathrm{VC}}\left(\mathscr{F}\big|_S\right) \leq \dim_{\mathrm{VC}}(\mathscr{F})$. Furthermore, by Lemma 12 we know that $\|\left(\mathscr{F}\big|_S\right)\| \leq |\text{Shatter}\left(\mathscr{F}\big|_S\right)|$. The first inequality now follows by the Observation

$$|\text{Shatter}\left(\mathscr{F}\big|_S\right)| \leq \sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{F}_S)} \binom{z}{j}.$$

Now we bound the sum of binomial coefficients as follows.

$$\sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{F})} \binom{z}{j} \leq \sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{F})} \frac{z^j}{j!} \leq z^{\dim_{\mathrm{VC}}(\mathscr{F})} \sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{F})} \frac{1}{j!} \leq z^{\dim_{\mathrm{VC}}(\mathscr{F})} \sum_{j \in \mathbb{N}_0} \frac{1}{j!} = ez^{\dim_{\mathrm{VC}}(\mathscr{F})},$$

where we used that $z \geq 1$ in the second inequality and the well known series representation of the Euler constant in the last equality. □

On the other hand a polynomial restriction on the shatter function is sufficient to bound the Vapnik Chervonenkis dimension.

Observation 10. In case that $\|\mathscr{F}\| \geq 2$ we have $\forall c > 0, \ d \in \mathbb{N}$

$$\left(\forall z \in \mathbb{N}: \ \pi_{\mathscr{F}}(z) \leq cz^d\right) \Longrightarrow \dim_{\mathrm{VC}}(\mathscr{F}) < 4d\log_2(cd).$$

Proof of Observation 10. Since $\|\mathscr{F}\| \geq 2$ we know that $\dim_{\mathrm{VC}}(\mathscr{F}) \geq 1$. By plugging in $z = 1$ we see that $c \geq 2$ and follow that $\log_2(4d) \leq \log_2\left((2d)^2\right) \leq 2\log_2(cd)$. Using this we see that a set of size $4d\log_2(cd)$ can not be *shattered* since

$$\pi_{\mathscr{F}}(4d\log_2(cd)) \leq c\left(4d\log_2(cd)\right)^d = 2^{\log_2(c) + d\log_2(4d) + d\log_2(\log_2(cd))} < 2^{4d\log_2(cd)}.$$ □

We can use the shift function defined in Lemma 12 to show a result about the edge density of induced subgraphs of the Boolean Hypercube. We are going to need this result in the proof of the important Packing lemma for hypergraphs in section 3.4. With help of the Boolean Hypercube, compare Definition 22, we introduce the so-called *Unit Distance Graph of* $\mathscr{F}$.

Definition 44 (Unit Distance Graph)**.** We want to define a graph on vertex set $\mathscr{E}$ where two vertices are adjacent if they differ in exactly one element. We may assume that for some $n \in \mathbb{N}: \ V(\mathscr{F}) = [n]$. We define

$$\mathbf{UD}(\mathscr{F}) := Q_n[\mathscr{E}] = \left(\mathscr{E}, \ \left\{ \{A, B\} \in \binom{\mathscr{E}}{2} \ \middle| \ |A \Delta B| = 1 \right\}\right).$$

Lemma 13 (Haussler [26])**.** $\frac{\|\mathbf{UD}(\mathscr{F})\|}{|\mathbf{UD}(\mathscr{F})|} \leq \dim_{\mathrm{VC}}(\mathscr{F})$.

Proof of Lemma 13. Let $x \in V(\mathscr{F})$ and $\mathscr{F}' := \left(X, \ \text{shift}_{\mathscr{F}}^{(x)}(\mathscr{E})\right)$ be the hypergraph obtained when we apply the shift operator on all the edges of $\mathscr{F}$ simultaneously.

Claim 2. $\|\mathbf{UD}(\mathscr{F})\| \leq \|\mathbf{UD}(\mathscr{F}')\|$.

Proof of Claim 2. Let us consider the mapping

$$\phi : E(\mathbf{UD}(\mathscr{F})) \to E(\mathbf{UD}(\mathscr{F}'))$$
$$\{A,B\} \mapsto \begin{cases} \{A,B\} & \{A\setminus\{x\}, B\setminus\{x\}\} \subseteq E(\mathscr{F}) \\ \{A\setminus\{x\}, B\setminus\{x\}\} & \text{otherwise} \end{cases}$$

This is well-defined. To show this let $\{A,B\} \in E(\mathbf{UD}(\mathscr{F}))$ and $\{Y,Y'\} = \phi(\{A,B\})$.
case $\{Y,Y'\} = \{A,B\}$. In this case it is clear that $\{Y,Y'\} \in E(\mathbf{UD}(\mathscr{F}'))$.
case $\{Y,Y'\} = \{A\setminus\{x\}, B\setminus\{x\}\}$. By the Definition of ϕ we know that $x \notin A\Delta B$ so

$$|(A\setminus\{x\})\,\Delta\,(B\setminus\{x\})| = 1.$$

Since by Definition of the shift operator $Y, Y' \in E(\mathscr{F}')$ it follows that $\{Y,Y'\} \in E(\mathbf{UD}(\mathscr{F}'))$.

To show the desired inequality it suffices to show that ϕ is injective. For this purpose let $\{A,B\}, \{A',B'\} \in E(\mathbf{UD}(\mathscr{F}))$ such that $\{Y,Y'\} := \phi(\{A,B\}) = \phi(\{A',B'\})$.
case $\{A,B\} = \{Y,Y'\}$. In this case we have that $\{Y\setminus\{x\}, Y'\setminus\{x\}\} \subseteq E(\mathscr{F})$. In case that $x \in Y\Delta Y'$ we may assume that $Y' = Y \cup \{x\}$, and it is easy to see that $\{A',B'\} = \{Y,Y'\}$. In case that $x \in Y \cap Y'$ we also know that $\{A',B'\} = \{Y,Y'\}$. Consider the case $x \in X \setminus (Y \cup Y')$. Let us assume for a contradiction that $\{A',B'\} \neq \{Y,Y'\}$. Then we may assume that $A' = Y \cup \{x\}$ and since $|A'\Delta B'| = 1$ we know that $B' = Y' \cup \{x\}$. However, since $\{A'\setminus\{x\}, B'\setminus\{x\}\} \subseteq E(\mathscr{F})$ we see that $\phi(\{A',B'\}) = \{A',B'\}$, a contradiction. Thus, also in this case $\{A',B'\} = \{Y,Y'\}$ and we are done.
case $\{A,B\} \neq \{Y,Y'\}$. In this case we have that $\{Y,Y'\} = \{A\setminus\{x\}, B\setminus\{x\}\}$, and we deduce that either $Y \notin \mathscr{F}$ or $Y' \notin \mathscr{F}$. Thus, we know that $\{A',B'\} \neq \{Y,Y'\}$ and the only way that $\{A',B'\}$ got mapped to $\{Y,Y'\}$ is $\{Y,Y'\} = \{A'\setminus\{x\}, B'\setminus\{x\}\}$. Thus, again $\{Y,Y'\} = \{A',B'\}$, which completes the proof of injectivity and of Claim 2. □

As in the proof of Lemma 12 we apply shifts with different $x \in X$ until the shifts corresponding to all vertices fix all hyperedges. Again we denote the resulting hypergraph by $\tilde{\mathscr{F}}$. Claim 2 and induction yield

$$\|\mathbf{UD}(\mathscr{F})\| \leq \|\mathbf{UD}\left(\tilde{\mathscr{F}}\right)\|.$$

Using the Observations we made in Lemma 12, namely $\max_{A\in\tilde{\mathscr{F}}}|A| = \dim_{\mathrm{VC}}\left(\tilde{\mathscr{F}}\right) \leq \dim_{\mathrm{VC}}(\mathscr{F})$ and $\|\mathscr{F}\| = \|\tilde{\mathscr{F}}\|$, we deduce

$$\|\mathbf{UD}\left(\tilde{\mathscr{F}}\right)\| = \sum_{A\in E\left(\tilde{\mathscr{F}}\right)} |\left(\{\, A\setminus\{x\} \mid x\in X \,\} \cap E\left(\tilde{\mathscr{F}}\right)\right)| \leq \sum_{A\in E\left(\tilde{\mathscr{F}}\right)} |A| \leq \|\tilde{\mathscr{F}}\| \cdot \dim_{\mathrm{VC}}(\mathscr{F}) = \|\mathscr{F}\| \cdot \dim_{\mathrm{VC}}(\mathscr{F})$$

This completes the proof of Lemma 13. □

3.1.2 Dual VC dimension

In the sequel we want to introduce duality to hypergraphs. For this purpose we define incidence sets of vertices. Let $\mathscr{F} = (X, \mathscr{E})$ be a hypergraph.

Definition 45 (Closed neighborhood)**.** For $x \in X$ let us define its incidence set $\mathscr{I}(x) := \{\, E \in \mathscr{E} \mid x \in E \,\}$.

Definition 46 (Duality)**.** Let us define the dual of $\mathscr{F}$ as $\mathscr{F}^* := (\mathscr{E}, \{\, \mathscr{I}(x) \mid x \in X \,\})$.

The best intuition about duality comes from the incidence graphs: $\mathrm{Incidence}(\mathscr{F}) = \mathrm{Incidence}(\mathscr{F}^*)$. Taking the dual of a hypergraph corresponds to flipping the partition classes in its incidence graph, compare with Figure 4.

Definition 47 (Dual Vapnik Chervonenkis dimension)**.** $\dim^*_{\mathrm{VC}}(\mathscr{F}) := \dim_{\mathrm{VC}}(\mathscr{F}^*)$.

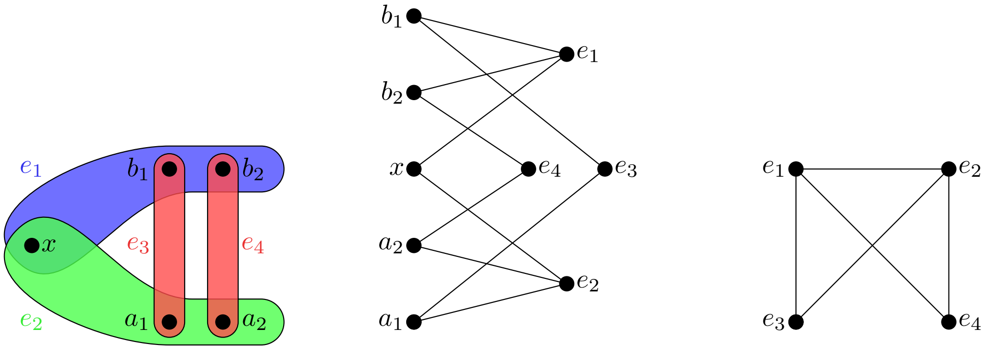

Figure 4: Two representations of the hypergraph $\mathscr{F}$ (left, middle) and the dual hypergraph $\mathscr{F}^*$ as graph (right).

One could ask if a high Vapnik Chervonenkis dimension implies a high dual Vapnik Chervonenkis dimension. The answer to this question is affirmative.

Lemma 14. $\dim^*_{\mathrm{VC}}(\mathscr{F}) \geq \lfloor \log_2 (\dim_{\mathrm{VC}}(\mathscr{F})) \rfloor$.

Proof of Lemma 14. Assume $\dim_{\mathrm{VC}}(\mathscr{F}) \geq 2^d$. Let $X \subseteq V(\mathscr{F})$ be a shattered set of size 2^d. We show that $\dim^*_{\mathrm{VC}}(\mathscr{F}) \geq d$ by finding an induced copy of $G_1 := \mathrm{Incidence}\left(2^{[d]}\right)$ inside $G_2 := \mathrm{Incidence}\left(2^X\right)$ such that the side of $2^{[d]}$ lies in X, see Figure 5.
Despite irritating notation it is targeted to identify $X = 2^{[d]}$ via an arbitrary embedding. For $j \in [d]$ consider the incident edges of j in $2^{[d]}$ which we denote by $D_j := \{ A \subseteq [d] \mid j \in A \}$. Consider $\mathscr{D} := \{ D_j \mid j \in [d] \}$ which we interpret as a subset of the power set of X. Then $G_2[\mathscr{D}, X] = G_1$ since for every $U \subseteq \mathscr{D}$, where $U = \{ D_j \mid j \in U' \}$ for some $U' \subseteq [d]$, we have that U' is an element in X whose neighborhood in $G_2[X, \mathscr{D}]$ is exactly U. □

Observe that other mappings from X to $2^{[d]}$ yield other induced copies of G_1 in G_2. In total, we can count $\frac{(2^d)!}{d!}$ many induced copies. Lemma 14 has the following instant Corollary.

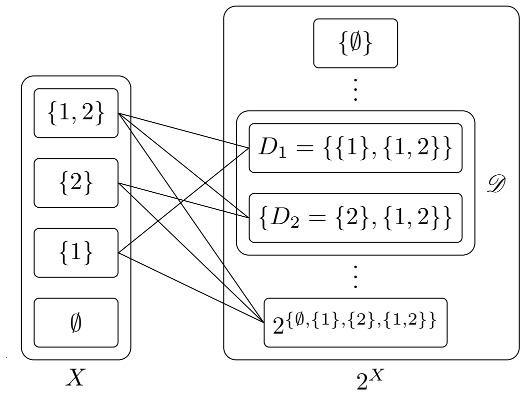

Figure 5: Visual proof of Lemma 14 in case $d = 2$.

Corollary 4. $\lfloor \log_2 (\dim^*_{\mathrm{VC}}(\mathscr{F})) \rfloor \leq \dim_{\mathrm{VC}}(\mathscr{F}) \leq 2^{\dim^*_{\mathrm{VC}}(\mathscr{F})}$

3.1.3 k-fold unions and VC dimension

In this section we study how set operations on the edges of a hypergraph change its VC dimension. Let $k \in \mathbb{N}$, $\Phi : \{0,1\}^k \longrightarrow \{0,1\}$ and $\mathscr{A}$ be a possibly infinite hypergraph.

Definition 48. For a sequence of hyperedges $(A_j)_{j\in[k]} \subseteq E(\mathscr{A})$ let us define

$$\Phi\left((A_j)_{j\in[k]}\right) := \left\{ v \in V(\mathscr{A}) \;\middle|\; \Phi\left((\mathbb{1}\{v \in A_j\})_{j\in[k]}\right) = 1 \right\}.$$

With this notation we define

$$\mathscr{A}^{\Phi} := \left\{ \Phi\left((A_j)_{j\in[k]}\right) \,\middle|\, (A_j)_{j\in[k]} \subseteq E(\mathscr{A}) \right\}.$$

Example 3. For a hypergraph $\mathscr{A}$ denote its closure under complements by $\mathscr{A}^c := \mathscr{A} \cup \mathscr{A}^{\neg}$. We observe $\dim_{\mathrm{VC}}(\mathscr{A}) \geq \left\lfloor \frac{\dim_{\mathrm{VC}}(\mathscr{A}^c)}{2} \right\rfloor$, which is sharp.

Proof of Example 3. Let us introduce $d := \dim_{\mathrm{VC}}(\mathscr{A})$ and $d^c := \dim_{\mathrm{VC}}(\mathscr{A}^c)$. Observe that $\pi_{\mathscr{A}^c}(z) \leq 2\pi_{\mathscr{A}}(z)$ for any $z \in \mathbb{N}$. Thus,

$$2^{d^c-1} = \frac{\pi_{\mathscr{A}^c}(d^c)}{2} \leq \pi_{\mathscr{A}}(d^c) \leq \sum_{0\leq j\leq d} \binom{d^c}{j}$$

where in the last inequality we used Corollary 3. Furthermore, by $\binom{q}{l} = \binom{q}{q-l}$ for any integers $l \leq q$ we have

$$\sum_{0\leq j<\frac{d^c}{2}} \binom{d^c}{j} = \sum_{\frac{d^c}{2}<j\leq d^c} \binom{d^c}{j}.$$

This argument shows that

$$2^{d^c-1} = \sum_{0\leq j<\frac{d^c}{2}} \binom{d^c}{j} + \frac{1}{2}\mathbb{1}\{d^c \text{ even}\}\binom{d^c}{\frac{d^c}{2}}.$$

Putting this together with the first inequality we obtain

$$\sum_{0\leq j<\frac{d^c}{2}} \binom{d^c}{j} + \frac{1}{2}\mathbb{1}\{d^c \text{ even}\}\binom{d^c}{\frac{d^c}{2}} \leq \sum_{0\leq j\leq d} \binom{d^c}{j}$$

which shows that $d \geq \left\lfloor \frac{d^c}{2} \right\rfloor$ as claimed.
For sharpness let $N \in \mathbb{N}$ and $d := \left\lfloor \frac{N}{2} \right\rfloor$. Consider the hypegraph

$$\mathscr{B} := \left([N],\ \{ B \subseteq [N] \mid |B| \leq d \}\right).$$

Obviously we have that $\dim_{\mathrm{VC}}(\mathscr{B}) = d$. Furthermore, we have that $E(\mathscr{B}^c) = 2^{[N]}$ so $\dim_{\mathrm{VC}}(\mathscr{B}^c) = N$. ☐

Definition 49 (k-fold union). $\mathscr{A}^{\cup,k} := \left(V(\mathscr{A}),\ \left\{ \bigcup_{j\in[k]} A_j \,\middle|\, (A_j)_{j\in[k]} \subseteq E(\mathscr{A}) \right\}\right).$

Obviously the k-fold union of a hypegraph $\mathscr{A}$ is of the form $\mathscr{A}^{\Phi}$ for $\Phi : \{0,1\}^k \longrightarrow \{0,1\}$ denoting the k-chaining of the or function.

Lemma 15. $\forall d \in \mathbb{N}\ \exists c_d > 0$ such that for every hypegraph $\mathscr{A}$ of VC dimension $\dim_{\mathrm{VC}}(\mathscr{A}) \leq d$ and any $k \in \mathbb{N}_{\geq 2}$ as well as any $\Phi:\ \{0,1\}^k \longrightarrow \{0,1\}$ one has

$$\dim_{\mathrm{VC}}\left(\mathscr{A}^{\Phi}\right) \leq c_d k \log_2(k).$$

A sketch of the proof of Lemma 15 can be found in [38].

Proof of Lemma 15. Let $S \subseteq V(\mathscr{A})$. For any sequence of edges $(A_j)_{j\in[k]} \subseteq E(\mathscr{A})$ we have that

$$\begin{aligned}\Phi\left((A_j)_{j\in[k]}\right) \cap S &= \left\{ v \in S \,\middle|\, \Phi\left((\mathbb{1}\{v \in A_j\})_{j\in[k]}\right) = 1 \right\} \\ &= \left\{ v \in S \,\middle|\, \Phi\left((\mathbb{1}\{v \in A_j \cap S\})_{j\in[k]}\right) = 1 \right\} \subseteq \Phi\left((A_j \cap S)_{j\in[k]}\right).\end{aligned}$$

This implies the inequality $\|\left(\mathscr{A}^{\Phi}\big|_S\right)\| \leq \|\left(\mathscr{A}\big|_S\right)\|^k$. Thus, we showed $\forall z \in \mathbb{N}:\ \pi_{\mathscr{A}^{\Phi}}(z) \leq \pi_{\mathscr{A}}(z)^k$.

Let us define $d_\Phi := \dim_{\mathrm{VC}}\left(\mathscr{A}^\Phi\right)$ and $d := \dim_{\mathrm{VC}}\left(\mathscr{A}\right)$. By choosing a large constant C we may assume that $d_\Phi \geq 4$. Using our Observation regarding the shatter function of $\mathscr{A}^\Phi$ and Corollary 3 we see that

$$2^{d_\Phi} = \pi_{\mathscr{A}^\Phi}\left(d_\Phi\right) \leq \left(\pi_{\mathscr{A}}\left(d_\Phi\right)\right)^k \leq \left(e \cdot \left(d_\Phi\right)^d\right)^k \leq \left(d_\Phi\right)^{(d+1)k},$$

where in the last inequality we used our assumption that $d_\Phi \geq 4 \geq e$. Set

$$\tilde{c}_d := \max\left\{ \frac{\dim_{\mathrm{VC}}\left(\mathscr{A}^\Phi\right)}{k\log_2\left(k\right)} \;\middle|\; k \in \left[(2d+3)^{(d+1)}\right],\ \Phi:\ \{0,1\}^k \longrightarrow \{0,1\} \right\}$$

and fix $c_d := \max\left\{2d+3, \tilde{c}_d\right\}$. Assume that $d_\Phi > c_d k\log_2\left(k\right)$. By Definition of c_d we know that $k > (2d+3)^{(d+1)}$. Let us define a real valued function

$$\begin{aligned} f : \mathbb{R}_+ &\to \mathbb{R}_+ \\ x &\mapsto \frac{\exp\left(\ln\left(2\right)x\right)}{x^{(d+1)k}} \end{aligned}$$

Taking the derivative and a simple calculation yield that f is growing on $[\frac{(d+1)k}{\ln(2)}, \infty)$. Since

$$d_\Phi \geq (2d+3)k\log_2\left(k\right) \geq \frac{(d+1)k}{\ln\left(2\right)}$$

we deduce

$$\frac{2^{((2d+3)k\log_2(k))}}{\left((2d+3)k\log_2\left(k\right)\right)^{(d+1)k}} = f\left((2d+3)k\log_2\left(k\right)\right) \leq f(d_\Phi) = 1.$$

This shows

$$k^{(2d+3)k} = 2^{(2d+3)k\log_2(k)} \leq \left((2d+3)k\log_2\left(k\right)\right)^{(d+1)k} \leq (2d+3)^{(d+1)k}k^{2(d+1)k}.$$

With this we deduce

$$k \leq (2d+3)^{(d+1)}.$$

However, this is a contradiction to our assumption and completes the proof of Theorem 15.

We remark that for any $y \in \mathbb{R}$:

$$y = \max\left\{ x \in \mathbb{R} \;\middle|\; 2^x \leq x^{(d+1)k} \right\} \Longrightarrow 2^y = y^{(d+1)k} \Longleftrightarrow \gamma\left(\frac{-\ln\left(2\right)y}{(d+1)k}\right) = \frac{-\ln\left(2\right)y}{(d+1)k}e^{\frac{-\ln(2)y}{(d+1)k}} = \frac{-\ln\left(2\right)}{(d+1)k}$$

where

$$\begin{aligned} \gamma : \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\mapsto ze^z \end{aligned}$$

The branches of the inverse relation of γ are called *Lambert W function*, which has been studied for centuries. Thus, one could optimize the constant c_d with knowledge about the Lambert W function. However, we decided that for our purposes here we do not need the best possible constants. □

3.2 VC dimension of graphs and hypergraphs

3.2.1 Various Definitions and Examples

In the sequel let $G = (V,\ E)$ be a graph. There are three types of VC dimensions for graphs. Goal of the Definition section is to make the distinction crystal clear.

Definition 50 (VC dimension of a graph)**.** Let us define the hypergraphs of the open and closed neighborhoods of G as

$$\mathscr{N}_{\text{open}}(G) := (V(G),\ \{\, N_G(v) \mid v \in V(G)\,\})\,.$$
$$\mathscr{N}_{\text{closed}}(G) := (V(G),\ \{\, N_G(v) \cup \{v\} \mid v \in V(G)\,\})\,.$$

With this let us define

$$\dim_{\text{VC}}(G\text{-open}) := \dim_{\text{VC}}(\mathscr{N}_{\text{open}}(G))\,.$$
$$\dim_{\text{VC}}(G\text{-closed}) := \dim_{\text{VC}}(\mathscr{N}_{\text{closed}}(G))\,.$$

Observation 11. $|\dim_{\text{VC}}(G\text{-open}) - \dim_{\text{VC}}(G\text{-closed})| \le 1$.

Proof of Observation 11. Let $X \in \text{Shatter}(\mathscr{N}_{\text{open}}(G))$ of maximal size, meaning $|X| = \dim_{\text{VC}}(G\text{-open})$. We have

$$\|\left(\mathscr{N}_{\text{closed}}(G)\big|_X\right)\| \ge 2^{|X|} - |X|$$

since the only vertices where the intersection of the open and closed neighborhood with X is different are the vertices in X. The Sauer lemma, Corollary 3, yields

$$2^{\dim_{\text{VC}}(G\text{-open})} - \dim_{\text{VC}}(G\text{-open}) \le \sum_{0 \le j \le \dim_{\text{VC}}(G\text{-closed})} \binom{\dim_{\text{VC}}(G\text{-open})}{j}.$$

This immediately implies that $\dim_{\text{VC}}(G\text{-closed}) \ge \dim_{\text{VC}}(G\text{-open}) - 1$.
The exact same argument for $X' \in \text{Shatter}(\mathscr{N}_{\text{closed}}(G))$ of maximal size shows

$$\dim_{\text{VC}}(G\text{-open}) \ge \dim_{\text{VC}}(G\text{-closed}) - 1.$$

Let us close with two minimal examples showing how the two VC dimensions can differ. Let us define

$$G_1 := K_1 + K_1.$$
$$G_2 := K_3 + K_1.$$

We observe that $\dim_{\text{VC}}(G_1\text{-open}) = 0$ but $\dim_{\text{VC}}(G_1\text{-closed}) = 1$. Furthermore, $\dim_{\text{VC}}(G_2\text{-open}) = 2$ but $\dim_{\text{VC}}(G_2\text{-closed}) = 1$. □

Definition 51 (Twins)**.** Two distinct vertices $a, b \in V(G)$ are called *twins* in case that they have the same neighborhood and *siblings* in case that they have the same closed neighborhood.

Observation 12. In case that G is twin-free one has $\mathscr{N}_{\text{open}}(G) = \mathscr{N}^*_{\text{open}}(G)$ and in case that G is sibling-free one has $\mathscr{N}_{\text{open}}(G) = \mathscr{N}^*_{\text{open}}(G)$. In any case

$$\dim_{\text{VC}}(\mathscr{N}^*_{\text{closed}}(G)) = \dim_{\text{VC}}(G\text{-closed})\,.$$
$$\dim_{\text{VC}}\left(\mathscr{N}^*_{\text{open}}(G)\right) = \dim_{\text{VC}}(G\text{-open})\,.$$

Proof of Observation 12. Let us first assume that G is twin-free. It suffices to show that the following is a surjective hypergraph isomorphism.

$$\Phi_{\text{open}}:\ V\left(\mathscr{N}^*_{\text{open}}(G)\right) \to V\left(\mathscr{N}_{\text{open}}(G)\right)$$
$$A \mapsto a,\ \text{where } A = N_G(a).$$

This is well-defined and bijective since G has no twins. Let us check, that it is a hypergraph homomorphism. For this purpose let $X = \{ A_j \mid j \in [l] \} \in E\left(\mathscr{N}^*_{\text{open}}(G)\right)$. By Definition for every $j \in [l]$ there is a unique $a_j \in V(G)$ such that $N_G(a_j) = A_j$. Moreover, there is $x \in V(G)$ such that $X = \mathscr{I}(x)$, see Definition 45. This just means, by Definition, that $N_G(x) = \{ a_j \mid j \in [l] \}$. This shows that

$$\Phi_{\text{open}}(X) := \{ \Phi_{\text{open}}(A_j) \mid j \in [l] \} = \{ a_j \mid j \in [l] \} \in E\left(\mathscr{N}_{\text{open}}(G)\right).$$

Let us now assume that G is sibling-free. Again we may show that the following is a surjective hypergraph isomorphism.

$$\Phi_{\text{closed}}:\ V\left(\mathscr{N}^*_{\text{closed}}(G)\right) \to V\left(\mathscr{N}_{\text{closed}}(G)\right)$$
$$A \mapsto a,\ \text{where } A = N_G(a) \cup \{a\}.$$

This is well-defined and bijective since G has no siblings. Let us check, that it is a hypergraph homomorphism. For this purpose let $X = \{ A_j \mid j \in [l] \} \in E\left(\mathscr{N}^*_{\text{closed}}(G)\right)$. By Definition for every $j \in [l]$ there is a unique $a_j \in V(G)$ such that $N_G(a_j) = A_j \cup \{a_j\}$. Moreover, there is $x \in V(G)$ such that $X = \mathscr{I}(x)$. This just means, by Definition, that $N_G(x) \cup \{x\} = \{ a_j \mid j \in [l] \}$. This shows that

$$\Phi_{\text{closed}}(X) := \{ \Phi_{\text{closed}}(A_j) \mid j \in [l] \} = \{ a_j \mid j \in [l] \} \in E\left(\mathscr{N}_{\text{closed}}(G)\right).$$

Now that we have shown the first two Claims let us show the Claims about the open and closed dual VC dimension. Let $a, b \in V(G)$ be twins. Observe that for any $S \in \text{Shatter}\left(\mathscr{N}_{\text{open}}(G)\right)$ one has $|S \cap \{a, b\}| \leq 1$. We deduce

$$\text{Shatter}\left(\mathscr{N}_{\text{open}}(G)\right) = \text{Shatter}\left(\mathscr{N}_{\text{open}}(G-a)\right) \cup \{ (S \setminus \{b\}) \cup \{a\} \mid S \in \text{Shatter}\left(\mathscr{N}_{\text{open}}(G-a)\right) \text{ with } b \in S \}.$$

This shows

$$\dim_{\text{VC}}(G\text{-open}) = \dim_{\text{VC}}((G-a)\text{-open}).$$

Claim 3. $\forall S \in \text{Shatter}\left(\mathscr{N}^*_{\text{open}}(G)\right):\ S' := \{ N \setminus \{a\} \mid N \in S \} \in \text{Shatter}\left(\mathscr{N}^*_{\text{open}}(G-a)\right)$ and $|S| = |S'|$.

Proof of Claim 3. By Definition there are vertices $V_S \subseteq V(G)$ such that $S = \{ N_G(v) \mid v \in V_S \}$ and for any $A \subseteq V_S$ there is $u_A \in V(G)$ such that $N_{V_S}(u_A) = A$. Here we may assume that $u_A \neq a$. Indeed in case that $u_A = a$ we can simply choose $u_A = b$ instead. Furthermore, we observe $|V_S \cap \{ a, b \}| \leq 1$ and by replacing a with b if necessary, we may assume that $a \notin V_S$. This shows that $S' = \{ N_{G-a}(v) \mid v \in V_S \}$ and for any $A' \subseteq V_S$ there is $u_{A'} \in V(G) \setminus \{a\}$ such that $N_{V_S}(u_{A'}) = A'$. However, we just showed that $S' \in \text{Shatter}\left(\mathscr{N}^*_{\text{open}}(G-a)\right)$. Let $N_1, N_2 \in S$ be two distinct neighborhoods. Assume for a contradiction that $N_1 \setminus \{a\} = N_2 \setminus \{a\}$. It follows that $N_1 \Delta N_2 = \{a\}$ and we may assume that $N_1 = N_2 \cup \{a\}$. Since a and b are twins it follows that $b \in N_1$. Thus, we know that $b \in N_2$ and again the twin property yields that $a \in N_1$. A contradiction.
The contradiction argument showed that $|S| = |S'|$. This closes the proof of Claim 3. □

With Claim 3 we deduce that

$$\dim_{\text{VC}}\left(\mathscr{N}^*_{\text{open}}(G)\right) = \dim_{\text{VC}}\left(\mathscr{N}^*_{\text{open}}(G-a)\right).$$

Thus, we may assume that G is twin-free and the Claim follows by the fact that $\mathscr{N}_{\text{open}}(G)$ is isomorphic with its dual. An analogous proof shows the Claim for the closed VC dimension. This closes the proof of Observation 12. □

Definition 52 (VC dimension of graph properties)**.** Let $\mathscr{C}$ be a graph property. Let us define

$$\dim_{\mathrm{VC}}(\mathscr{C}) := \dim_{\mathrm{VC}}(\mathscr{C}\text{-open}) := \max_{G\in\mathscr{C}}\dim_{\mathrm{VC}}(G\text{-open})\,.$$
$$\dim_{\mathrm{VC}}(\mathscr{C}\text{-closed}) := \max_{G\in\mathscr{C}}\dim_{\mathrm{VC}}(G\text{-closed})\,.$$

For illustrating purposes let us study the VC dimension of some graph properties.

Definition 53. We call a graph $G = ([n]\,,\ E)$ a permutation graph if there is $\phi \in \mathscr{S}_n$ such that

$$E = \left\{ \{u, v\} \in \binom{[n]}{2} \,\middle|\, u < v \text{ and } \phi(u) < \phi(v) \right\}.$$

We denote the graph property of permutation graphs by $\mathscr{C}_{\text{permutation}}$.

Example 4. $\dim_{\mathrm{VC}}(\mathscr{C}_{\text{permutation}}\text{-open}) = \dim_{\mathrm{VC}}(\mathscr{C}_{\text{permutation}}\text{-closed}) = 2$.

Proof of Example 4. Let us first show the upper bounds. For this purpose let G be a permutation graph and let $\emptyset \neq X \subseteq V(G)$ be shattered by its closed neighborhoods and Y be shattered by its neighborhoods. Assume $|X| \geq 3$, meaning there are $x_1, x_2, x_3 \in X$ with $x_1 < x_2 < x_3$. Let $i_1, i_2, i_3 \in [3]$ such that $\phi(x_{i_1}) < \phi(x_{i_2}) < \phi(x_{i_3})$. Let $v \in V(G)$ with $N_X(v) = \{x_{i_1}, x_{i_3}\}$. This implies that $v \leq \min\{x_{i_1}, x_{i_3}\}$ and $\phi(v) \leq \min\{\phi(x_{i_1}), \phi(x_{i_3})\}$. Thus, we know that $\phi(v) \leq \phi(x_{i_2})$ and since v and x_{i_2} are not adjacent we conclude that $v > x_{i_2}$. We deduce that $x_{i_2} = x_1$. However, this means that any vertex adjacent to x_1 is also adjacent to x_{i_3}, a contradiction. We remark that exactly the same contradiction arises for Y in case that we assume $|Y| \geq 3$.
Regarding the lower bound we check that the permutation graph corresponding to the permutation $(2, 4, 5, 1, 3, 6)$ has both open and closed VC dimension 2. Indeed, the set $\{3, 5\}$ is the largest set that gets shattered by both the open and closed neighborhoods. □

Since the following is only an example we are not going to rigorously introduce planar graphs at this point. We refer to the corresponding section in Diestel [15].

Example 5. Let us denote the graph property of planar graphs by $\mathscr{C}_{\text{planar}}$.

$$\dim_{\mathrm{VC}}(\mathscr{C}_{\text{planar}}\text{-open}) = 3.$$
$$\dim_{\mathrm{VC}}(\mathscr{C}_{\text{planar}}\text{-closed}) = 4.$$

Proof of Example 5. **Ad open VC dimension**. Let $G \in \mathscr{C}_{\text{planar}}$ and assume that there would be a set $X = \{x_1, x_2, x_3, x_4\} \in \binom{V(G)}{4}$ that is shattered by the neighborhoods. The celebrated Theorem of Kuratowski states that in G there is no subdivision of K_5.
For $u, w \in X$ let us write $u \sim w$ in case that there is $b \in V(G) \setminus X$ such that $N_X(b) = \{u, w\}$ and $u \nsim w$ otherwise. We observe that for any two distinct vertices $u, w \in X$ either $u \sim w$ or $\exists z \in X$ such that $N_X(z) = \{u, w\}$. Since there is $a \in N_G(X)$ with $a \notin X$ we know that X can not be the branching vertices of K_4, which implies that there are two distinct and non-adjacent vertices $u, w \in X$ such that $u \nsim w$. Thus, there is $z \in X$ with $N_X(z) = \{u, w\}$. We may assume that $u = x_1,\ w = x_3,\ z = x_2$. Since know that $\{x_1, x_3\}, \{x_2, x_4\} \notin E(G)$ we must have that $x_1 \sim x_4$ and $x_3 \sim x_4$, meaning there are $v_{\{1,4\}}, v_{\{3,4\}} \in V(G) \setminus X$ with $N_X(v_{\{1,4\}}) = \{x_1, x_4\}$ and $N_X(v_{\{3,4\}}) = \{x_3, x_4\}$.
Thus, we know that X lies on a common circle $C = (x_1, x_2, x_3, v_{\{3,4\}}, x_4, v_{\{1,4\}}, v_1)$. There have to be $b, c \in V(G) \setminus V(C)$ such that $N_X(b) = \{x_1, x_2, x_3\}$ and $N_X(c) = \{x_2, x_3, x_4\}$. In a planar embedding of G we observe that x has to lie "inside" of C and y "outside" of C or the other way around. In both cases there is no way for a to send all its edges towards X, a contradiction. Since this is a marginal example we do not formulate our argument in a more rigorous manner, the following sketch should be enough to reveal our simple Observation.

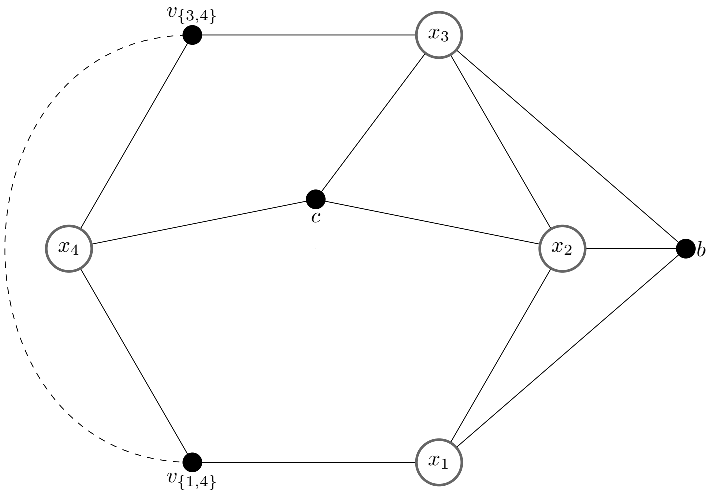

Figure 6: The setting of X.

On the other hand there is a planar graph G_{open} of open VC dimension 3, see Figure 7. We define it by

$$G_{\text{open}} \coloneqq \left(\{x\} \cup \{\, v_j, l_j \mid j \in [3]\,\},\ \left\{\, \{v_j, l_j\}, \{x, v_j\}, \left\{v_j, v_{((j \bmod 3)+1)}\right\} \,\middle|\, j \in [3] \,\right\}\right).$$

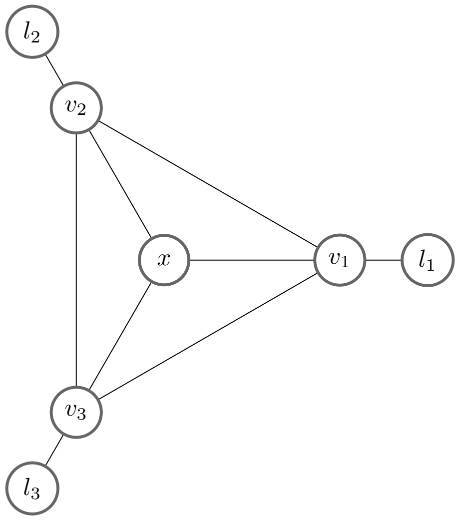

Figure 7: A planar embedding of G_{open}.

It is obvious from Figure 7 that G_{open} is planar. Since it is not enlightening to check that $\{\, v_j \mid j \in [3]\,\}$ is the (unique) largest set that is shattered by the open neighborhoods we leave this as an exercise for the reader.
Ad closed VC dimension. Let $G \in \mathscr{C}_{\text{planar}}$ and assume that there is a set $X = \{\, x_j \mid j \in [5]\,\} \in \binom{X}{5}$ that is shattered by the closed neighborhoods. For any $\{i, j\} \in \binom{[5]}{2}$ one either has that $N_X(x_i) = \{x_j\}$ or $N_X(x_j) = \{x_i\}$ or $x_i \sim x_j$, see for the proof of the open VC dimension. However, this shows that X are the branching vertices of a subdivision of K_5 inside G, a contradiction.
On the other hand there is a planar graph G_{closed} of closed VC dimension 4, see Figure 8, where we did not render $\left\{\, v_{\{j\}} \,\middle|\, j \in [4]\,\right\}$. Using $\mathscr{J} \coloneqq \binom{[4]}{1} \cup \binom{[4]}{2} \cup \{\{1,2,3\}, \{1,3,4\}\}$ we define it by

$$G_{\text{closed}} \coloneqq \left(\{\, x_j \mid j \in [k]\,\} \cup \{\, v_J \mid J \in \mathscr{J}\,\},\ \{\{x_2, x_4\}\} \cup \left\{\, \left\{x_j, x_{(j \bmod 4)+1}\right\} \,\middle|\, j \in [4]\,\right\} \cup \{\, \{x_j, v_J\} \mid J \in \mathscr{J},\ j \in J\,\}\right).$$

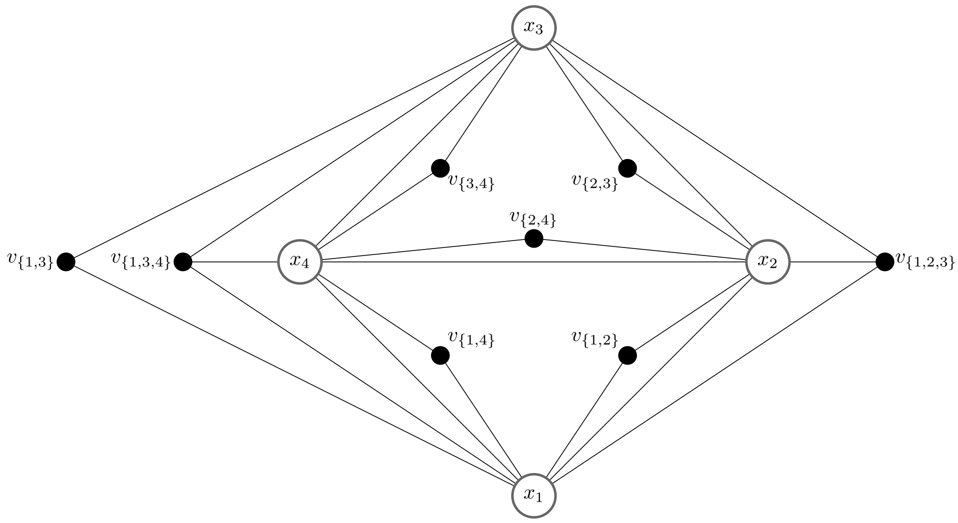

Figure 8: A planar embedding of $G_{\text{closed}} - \{ v_{\{j\}} \mid j \in [4] \}$.

It is obvious from Figure 8 that G_{closed} is planar. Since it is not enlightening to check that $\{ x_j \mid j \in [4] \}$ is the unique largest set that is shattered by the closed neighborhoods we leave this as an exercise for the reader. □

To give the result of Janzer and Pohoata, Theorem 15, as in the original paper let us define the asymmetric VC dimension of a bipartite graph.

Definition 54 (Asymmetric VC dimension)**.** Let $H = (A \cup B,\ F)$ be a bipartite graph. Then we define

$$\dim_{\text{VC}}(H, A) := \dim_{\text{VC}}((A,\ \{ N_A(b) \mid b \in B \}))\,.$$

Our main result, Theorem 17, is a strengthening of the following result.

Theorem 7 (Janzer and Pohoata [32])**.** $\forall s, d \in \mathbb{N}$ with $d \geq 3$ $\forall C > 0$ $\exists N \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n \geq N$ and any bipartite graph $G = (A \cup B,\ E)$ with $|A| = |B| = \left\lfloor \frac{n}{2} \right\rfloor$ that fulfills $K_{s,s} \notin G$ and $\dim_{\text{VC}}(G, A) \leq d$ one finds $\|G\| \leq Cn^{2-\frac{1}{d}}$.

Notice that in case $d = 2$ the bound $o\left(n^{\frac{3}{2}}\right)$ could not hold since $K_{2,2} \subseteq \text{Incidence}\left(2^{[3]}\right)$ and by Lemma 5 we have $\text{ex}(n, K_{2,2}) = \Omega\left(n^{\frac{3}{2}}\right)$.

We also want to remark that the result of Janzer and Pohoata was a strengthening of the following result by Fox, Pach, Sheffer, Suk and Zahl.

Theorem 8 (Fox, Pach, Sheffer, Suk and Zahl [18])**.** $\forall d, t \in \mathbb{N}$ with $t \geq d \geq 3$ $\forall c > 0$ $\exists C > 0$ such that for any bipartite graph $H = (A \cup B,\ E)$ with $a := |A|$ and $b := |B|$ such that the hypergraph $\mathscr{F} := (A,\ \{ N_A(b) \mid b \in B \})$ fulfills $\forall z \in \mathbb{N} :\ \pi_{\mathscr{F}}(z) \leq cz^d$ as well as $K_{t,t} \nsubseteq G$ one has $\|G\| \leq C\left(ab^{1-\frac{1}{d}} + b\right)$.

3.2.2 VC dimension of hypergraphs

When dealing with Ultra Strong Regularity in section 5.2 we are going to need a generalization of the notion of VC dimension of the neighborhoods to hypergraphs. This section is rather technical and needs to use heavy notation. It is recommended to read it only in context of Ultra Strong Regularity.
In the sequel let $\mathscr{H}$ be a hypergraph. First let us introduce the notion of neighborhood to hypergraphs.

Definition 55 (Neighborhood in hypergraphs)**.** For $U \subseteq V(\mathscr{H})$ let us define

$$N_{\mathscr{H}}(U) := \{ W \subseteq V(\mathscr{H}) \setminus U \mid U \cup W \in E(\mathscr{H}) \}\,,$$

where we also use the notation $N_{\mathscr{H}}(v) = N_{\mathscr{H}}(\{v\})$ for $v \in V(\mathscr{H})$. We want to remark the difference to Definition 45 and also remark the ambiguity with the Definition of common neighborhood in graphs.

Definition 56 (VC dimension for hypergraphs)**.** Let us define the hypergraph of the neighborhoods of $\mathscr{H}$ as

$$\mathscr{N}_{\mathscr{H}} := \left(2^{V(\mathscr{H})}, \left\{ N_{\mathscr{H}}(U) \,\middle|\, U \subseteq 2^{V(\mathscr{H})} \right\}\right).$$

In case that, for some $k \in \mathbb{N}$, $\mathscr{H}$ is k-uniform we define for $j \in [k]$

$$\dim_{\mathrm{VC}}^{(j)}(\mathscr{H}) := \dim_{\mathrm{VC}}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right).$$

$$\dim_{\mathrm{VC}}^{(j)*}(\mathscr{H}) := \dim_{\mathrm{VC}}\left(\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right)^*\right).$$

In case of uniformity we want to remark that

$$E\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right) = \left\{ N_{\mathscr{H}}(U) \,\middle|\, U \in \binom{V(G)}{k-j} \right\}.$$

Furthermore, in case $k = 2$ the Definition is equivalent to open VC dimension in 50.

Definition 57 (Twins in hypergraphs)**.** For $j \in [k]$ we refer to distinct sets $U_1, U_2 \in \binom{V(\mathscr{H})}{j}$ as *j-twins* in case that they have the same neighborhood $N_{\mathscr{H}}(U_1) = N_{\mathscr{H}}(U_2)$. In case $j = 1$ we simply say that distinct vertices $a, b \in V(\mathscr{H})$ are *twins* in case that $N_{\mathscr{H}}(a) = N_{\mathscr{H}}(b)$.

Observation 13 (Dual VC dimension for hypergraphs)**.** Let $k \in \mathbb{N}$ and $\mathscr{H}$ be a k-uniform hypergraph as well as $j \in [k]$. In case that $\mathscr{H}$ is $(k-j)$-twin-free we have that

$$\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right)^* = \mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-j}\right].$$

In every case we have

$$\dim_{\mathrm{VC}}^{(k-1)*}(\mathscr{H}) = \dim_{\mathrm{VC}}^{(1)}(\mathscr{H}).$$

Proof of Observation 13. The proof of Observation 13 follows the proof of Observation 12. Let us first assume that $\mathscr{H}$ is $(k-j)$-twin-free. We may show that the following is a surjective hypergraph isomorphism.

$$\Phi:\ V\left(\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right)^*\right) \to \binom{V(\mathscr{H})}{k-j}$$
$$N \mapsto U, \text{ where } N = N_{\mathscr{H}}(U).$$

Just for comprehension we remark that $\mathrm{N} \subseteq \binom{V(\mathscr{H})}{j}$. The mapping Φ is well-defined and bijective since $\mathscr{H}$ has no $(k-j)$-twins. Let us check, that it is a hypergraph homomorphism. For this purpose let

$$X = \{ N_i \mid i \in [w] \} \in E\left(\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right)^*\right).$$

By Definition for every $i \in [w]$ there is a unique $U_i \in \binom{V(\mathscr{H})}{k-j}$ such that $N_{\mathscr{H}}(U_j) = N_i$. Moreover, there is $A \in \binom{V(\mathscr{H})}{j}$ such that $X = \mathscr{I}(A) = \left\{ B \in E\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{j}\right]\right) \,\middle|\, A \in B \right\}$, see Definition 45. This just means, by Definition, that $N_{\mathscr{H}}(A) = \{ U_j \mid j \in [l] \}$, i.e.

$$\Phi(X) := \{ \Phi(N_j) \mid j \in [l] \} = \{ U_j \mid j \in [l] \} \in E\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-j}\right]\right).$$

Now as in Observation 12 we would like to show that when determining the VC dimensions $\dim_{\mathrm{VC}}^{(k-1)*}(\mathscr{H})$ and $\dim_{\mathrm{VC}}^{(1)}(\mathscr{H})$ we might assume that $\mathscr{H}$ is twin-free in the meaning of Definition 57. Let $a, b \in V(\mathscr{H})$ be twins. First, we remark that there is no hyperedge $e \in E(\mathscr{H})$ with $a, b \in e$ since otherwise $e \setminus \{a\} \in N_{\mathscr{H}}(a) \setminus N_{\mathscr{H}}(b)$.

Let us define a swap operation that swaps b out with a.

$$\mathrm{swap}_{a,b} : \; 2^{V(G)} \to 2^{V(G)}$$
$$A \mapsto \begin{cases} A & b \notin A \\ (A \setminus \{b\}) \cup \{a\} & b \in A, a \notin A. \\ \bot & \text{otherwise} \end{cases}$$

Furthermore, for $S \subseteq 2^{V(G)}$ we define the image $\mathrm{swap}_{a,b}(S) := \{ \mathrm{swap}_{a,b}(A) \mid A \in S \}$.

Claim 4. $\forall S \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]\right) : \; \mathrm{swap}_{a,b}(S) \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}-b}\left[\binom{V(\mathscr{H}-b)}{k-1}\right]\right)$ and $|S| = |\mathrm{swap}_{a,b}(S)|$.

We remark that, as in case of graphs, we defined $\mathscr{H} - b = \mathscr{H}\left[V(\mathscr{H}) \setminus \{b\}\right]$.

Proof of Claim 4. Let $S \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]\right)$ and $S' := \mathrm{swap}_{a,b}(S)$. S' is well-defined since no hyperedge in $\mathscr{H}$ can contain both a and b. For any $A' \subseteq S'$ there is $A \subseteq S$ with $A' = \mathrm{swap}_{a,b}(A)$. Furthermore, there is $v \in V(\mathscr{H})$ such that $N_{\mathscr{H}}(v) \cap S = A$. In case that there is $e \in A$ with $b \in e$ then by the twin property of a and b the $(k-1)$-set $e' := \mathrm{swap}_{a,b}(e)$ fulfills $e' \cup \{v\} \in E(\mathscr{H})$. However, we know that $e' \notin S$ since $N_{\mathscr{H}}(e) = N_{\mathscr{H}}(e')$. This argument shows that $|S| = |S'|$ as well as $N_{\mathscr{H}}(v) \cap S' = A'$. Since for any $e \in S'$ we have $b \notin e$ we conclude that $S' \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]\right)$. This closes the proof of Claim 4. ☐

With Claim 4 we deduce that

$$\dim_{\mathrm{VC}}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]\right) = \dim_{\mathrm{VC}}\left(\mathscr{N}_{\mathscr{H}-b}\left[\binom{V(\mathscr{H}-b)}{k-1}\right]\right).$$

Let us define the set of all $(k-1)$-sets that contain b and are themselves part of an edge.

$$E_b := \left\{ e \in \binom{A}{k-1} \;\middle|\; A \in E(\mathscr{H}),\; b \in e \right\}.$$

Claim 5. $\forall S \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]^*\right) : \; S' := \{ N \setminus E_b \mid N \in S \} \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}-b}\left[\binom{V(\mathscr{H}-b)}{k-1}\right]^*\right)$ and $|S| = |S'|$.

Proof of Claim 5. There are vertices V_S such that $S = \{ N_{\mathscr{H}}(v) \mid v \in V_S \}$. Furthermore, for any $Y \subseteq S$ there is $e_Y \in \binom{V(\mathscr{H})}{k-1}$ such that $Y = \mathscr{I}(e_Y) \cap S$. We know that there is $V_Y \subseteq V_S$ such that $Y = \{ N_{\mathscr{H}}(v) \mid v \in V_Y \}$. By the Definitions $V_Y = N_{\mathscr{H}}(e_Y) \cap V_S$, where we identify vertices with the one-element sets containing them. For this reason we know that $|V_S \cap \{a, b\}| \leq 1$ and by replacing b with a we may assume that $b \notin V_S$.
Furthermore, we know may assume that $b \notin e_Y$. Indeed, in case that $b \in e_Y$ we know that $a \notin e_Y$. Let us define $e'_Y := (e_Y \setminus \{b\}) \cup \{a\}$. Since by the twin property of a and b we have $N_{\mathscr{H}}(e_Y) = N_{\mathscr{H}}(e'_Y)$ we can simply take e'_Y instead of e_Y.
At this point it is clear that $S' \in \mathrm{Shatter}\left(\mathscr{N}_{\mathscr{H}-b}\left[\binom{V(\mathscr{H}-b)}{k-1}\right]^*\right)$.
Let us argue that $|S| = |S'|$. Assume otherwise, this is there are distinct $N_1, N_2 \in S$ such that $N_1 \setminus E_b = N_2 \setminus E_b$. Then, without loss of generality, there is $e \in E_b \cap (N_1 \setminus N_2)$. Let $v_1, v_2 \in V(G)$ such that $N_1 = N_{\mathscr{H}}(v_1)$ and $N_2 = N_{\mathscr{H}}(v_2)$. Since $e \cup \{v_1\} \in E(\mathscr{H})$ and $b \in e$ we know that $a \notin e'$. Let us define $e' := (e \setminus \{b\}) \cup \{a\}$. By the twin property we know that e' fulfills $e' \cup \{v_1\} \in E(\mathscr{H})$. Since $e' \notin E_b$ we know that $e' \in N_2$. However, again by the twin property it follows that $e \in N_2$, a contradiction. This closes the proof of Claim 5. ☐

With Claim 5 we deduce that

$$\dim_{\mathrm{VC}}\left(\mathscr{N}_{\mathscr{H}}\left[\binom{V(\mathscr{H})}{k-1}\right]^*\right) = \dim_{\mathrm{VC}}\left(\mathscr{N}_{\mathscr{H}-b}\left[\binom{V(\mathscr{H}-b)}{k-1}\right]^*\right).$$

Thus, deleting b from $\mathscr{H}$ does not change the two VC dimensions considered. The equality $\dim_{\mathrm{VC}}^{(k-1)*}(\mathscr{H}) = \dim_{\mathrm{VC}}^{(1)}(\mathscr{H})$ now follows from the first Claim. ☐

3.2.3 VC dimension of hereditary graph properties

Finally, in this section we want to draw the connection to the induced Turán problem.

Observation 14. For any graph H with $H \underset{ind}{\not\subseteq} P_4$ one has $\dim_{VC}(\text{Free}(H\text{-ind})\text{-open}) = \infty$. We remark that the requirement on H states that $H \notin \mathscr{P} := \{K_1, K_2, 2K_1, P_3, K_2 + K_1, P_4\}$.

Proof of Observation 14. Let $d \in \mathbb{N}$. We may find a graph $G \in \text{Free}(H\text{-ind})$ such that $\dim_{VC}(G\text{-open}) \geq d$.
case H is not bipartite. In this case we can simply choose $G = \text{Incidence}\left(2^{[d]}\right)$. Obviously $H \underset{ind}{\not\subseteq} G$.
case H is bipartite but $H \notin \mathscr{P} \cup \{C_4, 2K_2\}$. In this case we can take $G = \text{Incidence}\left(2^{[d]}\right) + \binom{[d]}{2} + \binom{2^{[d]}}{2}$ where we simply filled the partition classes of the incidence graph with all edges inside.
Notice H has independence number at least three. If $|H| \geq 5$ this follows by bipartitness and the pigeon whole principle. If $H \in \{3K_1, 4K_1, P_3 + K_1\}$ this follows by studying the specific graphs. Now if there were an induced copy of H in G then at least two of three independent vertices would lie in the same partition class of the incidence graph in G, a contradiction.
case $H \in \{C_4, 2K_2\}$. In this case we can take $G = \text{Incidence}\left(2^{[d]}\right) + \binom{2^{[d]}}{2}$ where we simply filled the bigger partition class of the incidence graph with all possible edges. We remark that $\overline{C_4} = 2K_2$. Notice that H has the property that every three of its vertices span an edge and a non-edge. Thus, since $A := [d]$ is independent in G and $B := 2^{[d]}$ is a clique we may assume that exactly two vertices $\{a_1, a_2\}$ of H lie in A and two vertices $\{b_1, b_2\}$ of H lie in B. However, as one can easily check, H has the property that either $\{a_1, a_2\}, \{b_1, b_2\} \in E(H)$ or $\{a_1, a_2\}, \{b_1, b_2\} \notin E(H)$. But in either case there is a contradiction to A independent or to the fact that B is a clique. □

Observation 15. $\dim_{VC}(\text{Free}(P_4\text{-ind})\text{-open}) = \dim_{VC}(\text{Free}(P_3\text{-ind})\text{-open}) = 2$.

Proof of Observation 15. First let us show, that any graph G that doesn not contain P_4 as an induced subgraph has VC dimension at most 2. Notice that this implies that every graph, that does not contain P_3 as an induced subgraph also has VC dimension at most 2.
Let us assume for a contradiction there would be $X = \{x_1, x_2, x_3\} \in \binom{V(G)}{3}$ that is shattered by the neighborhoods. For $I \subseteq \{1, 2, 3\}$ there is $y_I \in V(G)$ such that $N_X(y_I) = \{\, x_i \mid i \in I \,\}$.
case There are distinct $i, j \in [3]$ such that x_i, x_j are non-adjacent. Let $z \in [3] \setminus \{i, j\}$. Assume for a contradiction that x_z is adjacent to both x_i and x_j. In this case we would have $G\left[\left\{y_{\{i\}}, x_i, x_z, x_j\right\}\right] = P_4$, a contradiction.
Assume for a contradiction that x_z is adjacent to x_i but not to x_j. In this case we would have $G\left[\left\{x_z, x_i, y_{\{i,j\}}, x_j\right\}\right] = P_4$, a contradiction.
The previous two contradiction arguments showed that x_i, x_j and x_z are pairwise non-adjacent.
Assume for a contradiction that $y_{\{i,j\}}$ and $y_{\{i,z\}}$ are adjacent. In this case we would have $G\left[\left\{x_z, y_{\{z,i\}}, y_{\{i,j\}}, x_j\right\}\right] = P_4$, a contradiction.
Thus, we know that $y_{\{i,z\}}$ is non-adjacent to $y_{\{i,j\}}$. However, in this case $G\left[\left\{y_{\{i,j\}}, x_i, y_{\{i,z\}}, x_z\right\}\right] = P_4$, a contradiction.
case G contains all edges between x_1, x_2, x_3. Assume for a contradiction that $y_{\{2\}}$ and $y_{\{3\}}$ are non-adjacent. In this case we would have $G\left[\left\{y_{\{2\}}, x_2, x_3, y_{\{3\}}\right\}\right] = P_4$, a contradiction. Thus, we know that $y_{\{2\}}$ and $y_{\{3\}}$ are adjacent. However, in this situation $G\left[\left\{y_{\{2\}}, y_{\{3\}}, x_3, x_1\right\}\right] = P_4$, a contradiction.
Since we arrived at a contradiction in both cases we have shown that $\dim_{VC}(G) \leq 2$.

On the other hand, there is a graph that does not contain P_3 as an induced subgraph but has VC dimension 2. We give the minimal example $G = K_3 + K_1$ where any pair of vertices of the triangle is shattered by the neighborhoods. □

Observation 16. $\dim_{VC}(\text{Free}((K_2 + K_1)\text{-ind})\text{-open}) = 1$.

Proof of Observation 16. Let us assume for a contradiction that there is a graph G not containing $H := K_2 + K_1$ as an induced subgraph but having VC dimension at least two, meaning that there is a set $X = \{x_1, x_2\} \in \binom{V(G)}{2}$ that is shattered by the neighborhoods. Again for $I \subseteq \{1, 2\}$ there is $y_I \in V(G)$ such that $N_X(y_I) =$

$\{ x_i \mid i \in I \}$. In case $\{x_1, x_2\} \notin E(G)$ we find an induced copy of H on $\{x_1, y_{\{1\}}, x_2\}$, a contradiction. In case $\{x_1, x_2\} \in E(G)$ we find an induced copy of H on $\{x_1, x_2, y_\emptyset\}$, a contradiction.
On the other hand, there is a graph, that does not contain $K_2 + K_1$ but has VC dimension 1. We give the minimal example $G = P_3$ where the set of cardinality one containing any of the two endpoints of the path is shattered by the neighborhoods. □

Since in the proof of the Erdős-Hajnal conjecture the graph property Free (P_4-ind) is going to play an important role we take the time to present some structural result about it. Let us introduce a graph property that turns out to be Free (P_4-ind).

Definition 58 (Cograph)**.** We define the graph property $\mathscr{C}_{\text{Cograph}}$ of Cographs inductively by

(a) $K_1 \in \mathscr{C}_{\text{Cograph}}$.

(b) $\forall G, H \in \mathscr{C}_{\text{Cograph}} : \; G + H \in \mathscr{C}_{\text{Cograph}}$.

(c) $\forall G, H \in \mathscr{C}_{\text{Cograph}} : \; G \times H \in \mathscr{C}_{\text{Cograph}}$.

The following two Observations are an immediate consequence of the Definition of Cographs.

Observation 17. $\mathscr{C}_{\text{Cograph}}$ is closed under taking the graph complement.

Observation 18. $\mathscr{C}_{\text{Cograph}}$ is a hereditary graph property.

Observation 19. $\forall G \in \mathscr{C}_{\text{Cograph}}$: either G or $\overline{G}$ is connected.

Observation 20. $\forall G \in \mathscr{C}_{\text{Cograph}} : \; \alpha \vee \omega(G) \geq \sqrt{|G|}$.

Proof of Observation 20. Let us show the following by induction on $|G|$.

$$(*) \; \forall G \in \mathscr{C}_{\text{Cograph}} : \; \omega(G) \cdot \alpha(G) \geq |G|.$$

Then the Claim of Observation 20 is an immediate consequence.
base $|G| = 1$**.** The Claim is trivial.
step $|G| \geq 2$**.** By Definition of Cographs there are graphs G_1, G_2 containing at least one vertex each such that $G \in \{G_1 + G_2, G_1 \times G_2\}$.
case $G = G_1 + G_2$. We have $\alpha(G) = \alpha(G_1) + \alpha(G_2)$ as well as $\omega(G) = \max\{\omega(G_1), \omega(G_2)\}$. Thus,

$$\omega(G)\,\alpha(G) \geq \max\{\omega(G_1), \omega(G_2)\}(\alpha(G_1) + \alpha(G_2)) \geq \omega(G_1)\,\alpha(G_1) + \omega(G_2)\,\alpha(G_2) \geq |G_1| + |G_2| = |G|,$$

where we used induction in the last inequality.
case $G = G_1 \times G_2$. We remark that $\overline{G} = \overline{G_1} + \overline{G_2}$. Thus, we deduce with Observation 17 and the previous case that

$$\omega(G)\,\alpha(G) = \omega\left(\overline{G}\right)\alpha\left(\overline{G}\right) \geq |\overline{G}| = |G|.$$ □

Lemma 16. $\mathscr{C}_{\text{Cograph}} = \text{Free}(P_4\text{-ind})$.

Proof of Lemma 16. "$\subseteq$": Let $G \in \mathscr{C}_{\text{Cograph}}$ and assume for a contradiction that $P_4 \underset{ind}{\subseteq} G$. By Observation 18 it follows that $P_4 \in \mathscr{C}_{\text{Cograph}}$. However, since $\overline{P_4} = P_4$ and P_4 is connected, according to Observation 19, P_4 can not be a Cograph. A contradiction.
"$\supseteq$": Let us show the following by induction on n.

$$\forall n \in \mathbb{N}, \; G \in \text{Free}(n, P_4\text{-ind}) : \; G \in \mathscr{C}_{\text{Cograph}}.$$

base $n = 1$**.** This case is trivial.
step $n \geq 2$**.** Let G be a graph on more than one vertex that does not contain P_4 as an induced subgraph. By induction, we know that every proper induced subgraph of G is a Cograph. We may show that G is a Cograph.

Assume for a contradiction that G and $\overline{G}$ are connected. Fix $v \in V(G)$. Since $\overline{G}$ is connected there is some $u \in V(G) \setminus N_G(v)$. Let G', G'' be two connectivity components of $G - \{v\}$, we may assume that $u \in V(G')$. Assume that $G' \neq G''$.
Since G is connected we know that v has some neighbor in G', and we remark

$$\emptyset \neq N_{V(G')}(v) \neq V(G').$$

Since G' is connected there is $u' \in V(G') \setminus N(v)$ and $a' \in N_{V(G')}(v)$ such that a' and u' are adjacent. Furthermore, there is $a'' \in N_{V(G'')}(v)$. We arrive at the contradiction that $G[\{u', a', v, a''\}] = P_4$.
The contradiction argument showed that there is only one connectivity component in $G - v$, which implies $G' = G - v$. Since G' is a Cograph and connected we know that $\overline{G'}$ is disconnected. Let G'_1, G'_2 be the graph complements of two connectivity components in $\overline{G'}$. We observe that $G[V(G'_1), V(G'_2)]$ is complete bipartite. Since we assumed that G is connected we know that v is adjacent to some vertex w in G'. We may assume $w \in V(G'_1)$.
We know that $N_{V(G'_1)}(v) \neq V(G'_1)$ since otherwise $\overline{G}$ would be disconnected since $\overline{G'_1}$ would be its own connectivity component, a contradiction. Similarly, $N_{V(G'_2)}(v) \neq V(G'_1)$ and we can choose $b \in V(G'_2) \setminus N_{V(G'_2)}(v)$.
Let us partition $A := V(G'_1) \cap N_{V(G'_1)}(v)$ and $B = V(G'_1) \setminus N_{V(G'_1)}(v)$. Since $A \neq \emptyset \neq B$ and $\overline{G'_1}$ is connected we know that there are two non-adjacent vertices $a \in A$ and $b \in B$. We arrive at the contradiction that $G[\{v, a, b, b\}] = P_4$.
The contradiction argument showed that either G or $\overline{G}$ is disconnected. Notice that *(iii)* also holds for $\overline{G}$. Thus, induction yields that all components of either G or $\overline{G}$ are Cographs. By the inductive construction of Cographs this yields the Claim. □

In contrast to Observation 14 forbidding a bipartite graph as a biinduced subgraph bounds the VC dimension.

Observation 21. For any bipartite graph H one has $\dim_{\mathrm{VC}}(\mathrm{Free}(H\text{-biind})) < \infty$.

Proof of Observation 21. Let $H = (A \cup B,\ E)$ be a bipartite graph. We know that there is $t \in \mathbb{N}$ such that $H \underset{ind}{\subseteq} \mathrm{Incidence}\left(2^{[t]}\right)$.
Fix $d' \in \mathbb{N}$ and consider a graph G with $\dim_{\mathrm{VC}}(G) \geq d'$. Then there is a set $X \in \binom{V(G)}{d'}$ that is shattered by the neighborhoods, implying that there is a set $Y \subseteq V(G) \setminus X$ of at least $y := 2^{d'} - d'$ vertices such that $(N_X(v))_{v \in Y}$ are pairwise distinct. Consider the hypergraph $\mathscr{X} := (X,\ \{\ N_X(v) \mid v \in Y\ \})$. By the previous we know that $\|\mathscr{X}\| = y$. Now an application of the Sauer lemma, Corollary 3, yields that

$$y = \|\mathscr{X}\| \leq \pi_{\mathscr{X}}(d') \leq ed'^{\dim_{\mathrm{VC}}(\mathscr{X})}.$$

Using $d' \geq 2$ i.e. $y \geq 2^{d'-1}$, we conclude

$$\dim_{\mathrm{VC}}(\mathscr{X}) \geq \frac{\log_2(y) - \log_2(e)}{\log_2(d')} \geq \frac{d' - 1 - \log_2(e)}{\log_2(d')}.$$

Thus, if we choose d' large enough we have that $\dim_{\mathrm{VC}}(\mathscr{X}) \geq t$, which in turn yields that $H \underset{biind}{\subseteq} G$, a contradiction. □

Our next Theorem characterizes hereditary graph properties of unbounded VC dimension. First, we need some Definitions.

Definition 59 (Split graph)**.** We call a graph G a split graph if we can partition its vertices into a clique and an independent set.

Definition 60 (Co-bipartite graph)**.** We call a graph co-bipartite if its graph complement is bipartite.

Theorem 9 (Hereditary graph properties of unbounded VC dimension, Bousquet et al. [9])**.** Let $\mathscr{C}$ be a graph property such that $\sup_{G \in \mathscr{C}} \dim_{\mathrm{VC}}(G) = \infty$. Then $\mathscr{C}$ contains either all split graphs, or all co-bipartite graphs, or all bipartite graphs.

We present a proof orienting at the proof given in [9] but getting rid of two intermediate steps, significantly shortening the proof.

Proof of Theorem 9. Assume there would be a bipartite graph $G_1 = (A_1 \mathbin{\dot\cup} B_1,\ E_1)$, a co-bipartite graph $G_2 = (A_2 \mathbin{\dot\cup} B_2,\ E_2)$, where A_2, B_2 are cliques, and a split graph $G_3 = (A_3 \mathbin{\dot\cup} B_3,\ E_3)$, where A_3 is a clique and B_3 is an independent set, such that $G_1, G_2, G_3 \notin \mathscr{C}$. There is $d' \in \mathbb{N}$ such that $H := \text{Incidence}\left(2^{[d']}\right)$ fulfills the following properties, where we denote the partition classes of H by $X := [d']$ and $Y := 2^{[d']}$.

(i) There is a copy of G_1 in H where either the vertices corresponding to A_1 lie in X and the vertices corresponding to B_1 lie in Y or the other way around.

(ii) There is a copy of G_2 in H where either the vertices corresponding to A_2 lie in X and the vertices corresponding to B_2 lie in Y or the other way around.

(iii) There are two copies of G_3 in H. One, where the vertices corresponding to A_3 lie in X and the vertices corresponding to B_2 lie in Y and one other copy where it is exactly the other way around.

Let $C > 0$ be the constant given by Theorem 23 when applied to G_1. Let us fix $d := \left\lceil 2^{\frac{(2\log_2(d'))^2}{C^3}} + 1 \right\rceil$.
Assume for a contradiction that there is $G \in \mathscr{C}$ with $\dim_{\mathrm{VC}}(G) \geq d$. Then there is $X \in \binom{V(G)}{d}$ that gets shattered by the neighborhoods. Now we know that there are at least $y := 2^d - d$ vertices $Y \subseteq V(G) \setminus X$ such that $(N_X(v))_{v \in Y}$ are pairwise distinct.
We further know that $G_1 \underset{ind}{\nsubseteq} G[Y]$ so an application of Theorem 23 yields a homogeneous set $Y' \subseteq Y$ of size y' where $y' \geq 2^{C\sqrt{\log_2(y)}}$. Since $d \geq 2$ we know that $d \leq 2^{d-1}$ so $y \geq 2^{d-1}$, and we deduce that $y' \geq 2^{C\sqrt{d-1}}$.
An application of the Sauer lemma, Corollary 3, on the hypergraph $\mathscr{X} := (X,\ \{\ N_X(y) \mid y \in Y\ \})$ and basic algebra yield

$$|\mathscr{X}| = \pi_{\mathscr{X}}(d) \leq \sum_{0 \leq j \leq \dim_{\mathrm{VC}}(\mathscr{X})} \binom{d}{j} \leq (d+1)^{\dim_{\mathrm{VC}}(\mathscr{X})}$$

With $|\mathscr{X}| = y'$ we deduce that $2^{C\sqrt{d-1}} \leq (d+1)^{\dim_{\mathrm{VC}}(\mathscr{X})}$ which yields $x' := \dim_{\mathrm{VC}}(\mathscr{X}) \geq \frac{C\sqrt{d-1}}{\log_2(d+1)}$.
By Definition of VC dimension we find $X' \in \binom{X}{x}$ that is shattered by the neighborhoods of the vertices in Y'.
Again we know that $G_1 \underset{ind}{\nsubseteq} G[X']$ so another application of Theorem 23 yields a homogeneous set $X'' \subseteq X'$ of size x'' where $x'' \geq 2^{C\sqrt{\log_2(x')}}$. Observe that for large enough d one has $\log_2(\log_2(d+1)) \leq C\frac{\log_2(d-1)}{4}$. With this we calculate, where in the last inequality we plug in the Definition of d

$$x'' \geq 2^{C\sqrt{\log_2(x')}} \geq 2^{C\sqrt{\frac{C}{2}\log_2(d-1) - \log_2(\log_2(d+1))}} \geq 2^{\frac{C^{\left(\frac{3}{2}\right)}}{2}\sqrt{\log_2(d-1)}} \geq d'.$$

Finally, we observe that in all cases X'' independent or clique, Y' independent or clique using (i), (ii) and (iii) we find an induced copy of G_1, G_2 or G_3 in G, a contradiction. □

Observation 22. Let $k \in \mathbb{N}_{\geq 2}$ and H be a bipartite graph. Then there exists $d \in \mathbb{N}$ such that for any $n \in \mathbb{N}$

$$\forall G \in \text{Free}(K_n, \{K_k, H\text{-ind}\}) :\ \dim_{\mathrm{VC}}(G) \leq d.$$

Proof of Observation 22. Observe that $\bigcup_{n \in \mathbb{N}} \text{Free}(K_n, \{K_k, H\text{-ind}\})$ is a hereditary graph property by Definition. However, it does not contain all split graphs, nor all the co-bipartite graphs, nor all the bipartite graphs. Thus, Theorem 9 yields the Claim. □

3.3 VC dimension and intersection hypergraphs

As an interesting case study in this section we consider the VC dimension of naturally occurring geometric set systems.

Definition 61. Let $\mathscr{A} \subset 2^{\mathbb{R}^d}$. For finitely many points $P \subseteq \mathbb{R}^d$ we define the hypergraph

$$\mathscr{H}_{\mathscr{A}}(P) := (P, \{ E \subset P \mid \exists A \in \mathscr{A} : A \cap P = E \}).$$

With this notion in mind we define the VC dimension of $\mathscr{A}$ as the supremum of the VC dimensions of all such hypergraphs.

$$\begin{aligned} \text{Shatter}(\mathscr{A}) &:= \{ P \subseteq \mathbb{R}^d \text{ finite} \mid \mathscr{H}_{\mathscr{A}}(P) = 2^P \}. \\ \dim_{\text{VC}}(\mathscr{A}) &:= \sup\nolimits_{P \in \texttt{Shatter}(\mathscr{A})} |P|. \end{aligned}$$

Here we are going to consider the cases that $\mathscr{A}$ are all closed or open unit-balls, halfspaces or axis parallelogram boxes. Considering the VC dimension Hellys theorem will prove itself a useful tool. To state it correctly we need the notion of affine independence.

Definition 62. Let $m \in \mathbb{N}$ and $X = \{ x_j \mid j \in [m] \} \subseteq \mathbb{R}^d$. We say X is *affinely dependent* in case that $\exists \alpha \in \mathbb{R}^m \setminus \{0\}$ such that

$$\sum_{j \in [m]} \alpha_j = 0 \text{ and } \sum_{j \in [m]} \alpha_j x_j = 0.$$

Otherwise, X is called *affinely independent.*

Observation 23. A maximal set of affinely independent points in $\mathbb{R}^d$ contains $d+1$ points.

Proof of Observation 23. Notice that $X = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^d$ is affinely independent if and only if

$$\left\{ \begin{bmatrix} x_1 \\ \\ 1 \end{bmatrix}, \ldots, \begin{bmatrix} x_m \\ \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^{d+1}$$

is linearly independent. This already proves the upper bound. The lower bound can be seen by considering the independent set $\{(e_1 + e_{d+1}), \ldots, (e_d + e_{d+1}), e_{d+1}\}$ in $\mathbb{R}^{d+1}$. ☐

Definition 63. Define barycentrics$(m) := \left\{ \alpha \in [0,1]^m \;\middle|\; \sum_{j \in [m]} \alpha_j = 1 \right\}$. Let $X = \{ x_j \mid j \in [m] \} \subseteq \mathbb{R}^d$ a finite set of points. Let us define the convex hull of X as

$$\text{conv}(X) := \left\{ \sum_{j \in [m]} \alpha_j x_j \;\middle|\; \alpha \in \text{barycentrics}(m) \right\}.$$

A proof of the following Theorem can also be found in a textbook on Convex Geometry by Hug and Weil [28].

Lemma 17 (Hellys theorem)**.** Let $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^d$ be an affinely independent set. Then there is a partition $A \cup B = [m]$ such that $\text{conv}(\{ x_i \} i \in A) \cap \text{conv}(x_j | j \in A) \neq \emptyset$.

Proof of Lemma 17. By Definition of affinely dependence there is $\alpha \in \mathbb{R}^m \setminus \{0\}$ such that $\sum_{j \in [m]} \alpha_j = 0$ and $\sum_{j \in [m]} \alpha_j x_j = 0$. Define $A := \{ j \in [m] \mid \alpha_j \geq 0 \}$ and $B := [m] \setminus A$. Then $\sum_{j \in A} \alpha_j = -\sum_{j \in B} \alpha_j$. Furthermore, $\alpha^+ := \frac{(\alpha_j \mathbb{1}\{j \in A\})_{j \in [m]}}{\sum_{j \in A} \alpha_j}$ and $\alpha^- := \frac{(\alpha_j \mathbb{1}\{j \in B\})_{j \in [m]}}{\sum_{j \in B} \alpha_j}$ are well-defined vectors in $\mathbb{R}^m$ since the denominators are positive. Moreover, $\alpha^+, \alpha^- \in$ barycentrics(m) and $\sum_{j \in [m]} \alpha_j^+ x_j = \sum_{j \in [m]} \alpha_j^- x_j$ lies in the intersection of the convex hulls as claimed. ☐

Definition 64 (Geometric objects)**.** Let $x, y \in \mathbb{R}^d$. We write $x \leq y$ in case that $\forall j \in [d] : x_j \leq y_j$. Similarly, we write $x < y$ in case that $\forall j \in [d] : x_j < y_j$. Furthermore, in case $x \leq y$ we denote the *open box* or *interval* spanned by x, y as

$$(x, y) := \{ z \in \mathbb{R}^d \mid x < z < y \} .$$

Moreover, we denote the set of all open d-dimensional boxes as

$$\mathscr{B}_d := \{ (a, b) \mid a, b \in \mathbb{R}^d,\ a \leq b \} .$$

For $u \in \mathbb{R}^d,\ r \in \mathbb{R}_+$ denote with $B_r(u)$ the open ball in $\mathbb{R}^d$ with center u and radius r. Furthermore, we denote all the d-dimensional open balls by

$$\mathscr{A}_d := \{ B_r(u) \mid r > 0,\ u \in \mathbb{R}^d \} .$$

Lastly for some normal vector $n \in \mathbb{R}^d$ with $|n|_2 = 1$ and some shift constant $s \in \mathbb{R}$ we want to define the open halfspace

$$H_s(n) := \{ x \in \mathbb{R}^d \mid \langle x, n \rangle > s \} .$$

With this we define the set of all d-dimensional open halfspaces by

$$\mathscr{O}_d := \{ H_s(n) \mid n \in \mathbb{R}^d \text{ with } |n|_2 = 1,\ s \in \mathbb{R} \} .$$

The following construction of a regular polyeder will provide some simple example point set.

Observation 24 (Existence of regular convex polytopes)**.** For $d \in \mathbb{N}_0$ there is a point set $P_d \in \binom{\mathbb{R}^d}{d+1}$ such that

(i) $\forall p \neq p' \in P_d :\ |p - p'|_2 = 1.$

(ii) $\exists x_d \in \mathbb{R}^d,\ \rho_d \in [0, 1) :\ \forall p \in P_d :\ |p - x_d|_2 = \rho_d.$

Proof of Observation 24. We prove the Claim by induction on d.
base: In case $d = 0$ we choose $P_0 := \mathbb{R}^0,\ \rho_0 := 0$ and x_0 to be the single element in $\mathbb{R}^0$, which we want to identify with 0. In case $d = 1$ choose $P_1 := \{0, 1\}$ and $x_1 := \frac{1}{2},\ \rho_1 := \frac{1}{2}$.
step: Let $d \in \mathbb{N}$ and assume the Claim holds for d.
Observe that by Pythagoras any $(x_d, h) \in x_d \times \mathbb{R}$ has distance exactly $\sqrt{\rho_d^2 + h^2}$ to any point in $P_d \times \{0\}$. By choosing $h_{d+1} := \sqrt{1 - \rho_d^2}$, which is well-defined since $\rho_d < 1$, we see that if we choose $p_{d+1} := (x_d, h_{d+1})$ then all points in $P_{d+1} = P_d \times \{0\} \cup \{p_{d+1}\}$ have pairwise distance 1.
To complete the induction step we need to find $x_{d+1} \in \mathbb{R}^{d+1}$ and $\rho_{d+1} \in [0, 1)$ such that all points in P_{d+1} have distance ρ_{d+1} to x_{d+1}. Choose $x_{d+1} := (x_d, \frac{2h_{d+1}^2 - 1}{2h_{d+1}})$ and $\rho_{d+1} = \frac{1}{2h_{d+1}}$. Then we have

$$|x_{d+1} - p_{d+1}|_2 = h_{d+1} - \frac{2h_{d+1}^2 - 1}{2h_{d+1}} = \rho_{d+1}.$$

Moreover, for any $p \in P_d \times \{0\}$, using Pythagoras and $\rho_d^2 = 1 - h_{d+1}^2$, we calculate

$$|x_{d+1} - p|_2 = \sqrt{\rho_d^2 + \left(\frac{1 - 2h_{d+1}^2}{2h_{d+1}} \right)^2} = \sqrt{1 - h_{d+1}^2 + \frac{1 - 4h_{d+1}^2 + 4h_{d+1}^4}{4h_{d+1}^2}} = \sqrt{\frac{1}{4h_{d+1}^2}} = \rho_{d+1}. \qquad \square$$

Lemma 18. Let $d \in \mathbb{N}$ and let $\mathscr{A}_d := \{ B_r(x) \mid x \in \mathbb{R}^d,\ r \in \mathbb{R}_+ \}$ denote all open balls in $\mathbb{R}^d$. Then

$$\dim_{\mathrm{VC}}(\mathscr{A}_d) = \dim_{\mathrm{VC}}(\mathscr{O}_d) = d + 1.$$

Proof of Lemma 18. Let us first prove the upper bounds. For this let $X \in \binom{\mathbb{R}^d}{d+2}$. We may show that none of $\mathscr{A}_d, \mathscr{O}_d$ shatters X. By Observation 23 we know that X is affinely dependent so Helly's Theorem yields a partition $X_1 \cup X_2 = X$ and a point $y \in \mathrm{conv}(X_1) \cap \mathrm{conv}(X_2)$.

First we want to show that $\mathscr{O}_d$ does not shatter X. Every open half space A in $\mathscr{O}_d$ that contains X_1 also contains the point y by convexity. Assume that A does not contain any point of X_2. Again by convexity of the complement of A this would imply that $\operatorname{conv}(X_2)$ lies in the complement of A, a contradiction to $y \in X_2$. Thus, there is no open halfspace $A \in \mathscr{O}_d$ such that $X \cap A = X_1$.
Now we want to prove that $\mathscr{A}_d$ can not shatter X. For this assume, it would be possible. Specifically there are open balls B_1, B_2 such that $B_1 \cap X = X_1$ and $B_2 \cap X = X_2$. Since y lies in the inner of both balls, the balls are intersecting and there is a unique hyperplane P that contains the intersection of the borders of B_1 and B_2. Notice now that $X_1 \subseteq B_1 \setminus B_2$ and $X_2 \subseteq B_2 \setminus B_1$ implies that X_1, X_2 lie in the open halfspaces separated by P (observe that X_1, X_2 cannot intersect P because $P \cap B_1 \subseteq B_2$ and $P \cap B_2 \subseteq B_1$). This however is a contradiction to the existence of y.
For the lower bound consider the regular polyeder P_d guaranteed by the previous Observation 24. Let $X \subseteq P_d$ and define $d' := |X| - 1$. We want to find a hyperplane that separates the points X and $P_d \setminus X$. In case that $X = \emptyset$ it is simple to see that we can choose an open halfspace that contains none of the points in P_{d+1}. Thus, we may assume that $d' \geq 0$. Since by (i) all pairs of points in P_d have the same distance and $P_{d'} \subseteq P_d$ by construction there is a sequence of flips and rotations that maps X to $P_{d'}$. We do not prove this standard result from geometry since its technicalities here do not yield any more insights. Thus, by symmetry we may assume that $X = P_{d'}$. Let us choose the normal vector

$$n := \left(\underbrace{0, \ldots, 0}_{d'}, \underbrace{1, \ldots, 1}_{d-d'} \right).$$

Then by the construction of the regular polyeder we know

$$\forall x \in P_{d'} : \ \langle x, n \rangle = 0 \text{ and}$$
$$\forall x \in P_d \setminus P_{d'} : \ \langle x, n \rangle > 0.$$

Thus, we find $s > 0$ such that

$$H_s(n) \cap P_d = P_d \setminus P_{d'}.$$

This already proves that P_d is shattered by $\mathscr{O}_d$. In the region of P_d we can approximate the open halfspaces by very large open balls, each of the same radius. This argument shows that also $\mathscr{A}_d$ shatters P_d. □

3.3.1 VC dimension of the k-fold union of halfspaces

For the lower bound of the VC dimension of the k-fold union of halfspaces we need the following Lemma which we prove at the end of this subsection.

Lemma 19 (Pumpkin lemma, Kupavskii, Nabil, Pach [35])**.** For $n, d \in \mathbb{N}_{\geq 2}$ define $K := (d-1)(n+1)\,2^{n-2}$. Let us remind of the notation for d-dimensional boxes in Definition 64. We can construct a set of K many open d-dimensional boxes $\mathscr{A} \in \binom{\mathscr{B}_d}{K}$ such that $\forall S \subseteq \mathscr{A}\ \exists \mathcal{Q}(S) \in \binom{\mathbb{R}^d}{2^{n-1}}$

(i) $\forall\,(a,\ b) \in S : \ |\,(a,\ b) \cap \mathcal{Q}(S)| = 1$

(ii) $\forall\,(a,\ b) \in \mathscr{A} \setminus \mathcal{S} : \ (a,\ b) \cap \mathcal{Q}(S) = \emptyset$

Lemma 20 (Csikós, Mustafa, Kupavskii [14])**.** Let $d, k \in \mathbb{N}$ and let $\mathscr{O}_d^{\cup,k}$ denote the k-fold union of the open halfspaces in $\mathbb{R}^d$. Then there is a constant $c := c(d)$ such that

$$\forall k \in \mathbb{N} : \ \frac{1}{c} dk\log_2(k) \leq \dim_{\mathrm{VC}}\left(\mathscr{O}_d^{\cup,k}\right) \leq cdk\log_2(k).$$

We present a refinement of the proof given in [14], where we spell out some of the steps.

Proof of Lemma 20. The upper bound follows by Lemma 15 and 18.

In order to prove the lower bound fix $d, k \in \mathbb{N}$ and define $d' := \left\lfloor \frac{d}{2} \right\rfloor$, $n := \lfloor \log_2(k) \rfloor + 1$ as well as $K := (d'-1)(n+1)2^{n-2}$. The Pumpkin lemma 19 yields a set of K many d'-dimensional axis-parallel boxes $\mathscr{B} \in \binom{\mathscr{B}_{d'}}{K}$, such that $\forall S \subseteq \mathscr{B}$ there is a set of 2^{n-1} many points $\mathcal{Q}(S) \in \binom{\mathbb{R}^{d'}}{2^{n-1}}$ fulfilling the properties $(i), (ii)$ stated in the Lemma.

By Definition, we have $2^{n-1} \leq k$. For $S \subseteq \mathscr{B}$ let us arbitrarily double points of $\mathcal{Q}(S)$ such that $|\mathcal{Q}(S)| = k$. We want to shift the boxes $\mathscr{B}$ to obtain boxes $\mathscr{B}'$ such that for any two corner points p_1, p_2 of boxes in $\mathscr{B}$ and the corresponding corner points p_1', p_2' of the shifted boxes we have

(Shift_1) $p_1 < p_2 \Longrightarrow p_1' < p_2'$.

(Shift_2) $p_1 = p_2 \Longrightarrow p_1' < p_2'$ or $p_2' < p_1'$.

(Shift_3) $p_1' \in Q_{\text{corners}}$.

Here we define grid points $Q_{\text{corners}} := \left\{ (d')^{2j} \;\middle|\; j \in \mathbb{N} \right\}^{d'}$. Let us denote $\mathscr{B} = \left\{ \left(a^{(j)}, b^{(j)}\right) \mid j \in [K] \right\}$ and $\mathscr{B}' = \left\{ \left(a'^{(j)}, b'^{(j)}\right) \mid j \in [K] \right\}$ where $a^{(j)}, b^{(j)}$ got shifted to $a'^{(j)}, b'^{(j)}$ respectively for $j \in [K]$.
Note that for every $S \subseteq [K]$ we can also shift $\mathcal{Q}(S) = \left\{ q^{(i)} \;\middle|\; i \in [k] \right\}$ and obtain points $\mathcal{Q}(S)' = \left\{ q'^{(i)} \;\middle|\; i \in [k] \right\}$ such that for every $j \in [K]$ and $i \in [k]$

(Shift_4) $a^{(j)} \leq q^{(i)} \leq b^{(j)} \Longleftrightarrow a'^{(j)} < q'^{(i)} < b'^{(j)}$.

(Shift_5) $q^{(i)} \in Q_{\text{seeds}}$.

Here we define grid points $Q_{\text{seeds}} := \left\{ (d')^{2j+1} \;\middle|\; j \in \mathbb{N} \right\}^{d'}$. Furthermore, we want to define the mappings

$$\pi_{d'} : \left(\mathbb{R}_+^{d'}\right)^2 \to \mathbb{R}_+^{2d'} \qquad\qquad \gamma_{d'} : \mathbb{R}_+^{d'} \to \mathbb{R}_+^{2d'}$$

$$(a,b) \mapsto \begin{bmatrix} a_1 \\ \frac{1}{b_1} \\ \vdots \\ a_d \\ \frac{1}{b_d} \end{bmatrix} \qquad\qquad s \mapsto \begin{bmatrix} \frac{1}{s_1} \\ s_1 \\ \vdots \\ \frac{1}{s_d} \\ s_d \end{bmatrix}$$

Now for every $j \in [K]$, $i \in [k]$ we have (let us abbreviate $a'^{(j)}, b'^{(j)}, q'^{(i)}$ by a', b', q' respectively)

$$q^{(i)} \in \left(a^{(j)}, b^{(j)}\right) \Longleftrightarrow a'^{(j)} < q'^{(i)} < b'^{(j)} \Longleftrightarrow \begin{bmatrix} \frac{a_1'}{q_1'} \\ \frac{q_1'}{b_1'} \\ \vdots \\ \frac{a_{d'}'}{q_{d'}'} \\ \frac{q_{d'}'}{b_{d'}'} \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \Longleftrightarrow \sum_{j \in [d']} \left(\frac{a_j'}{q_j'} + \frac{q_j'}{b_j'} \right) < 3,$$

where in the first equivalence we used (Shift_4) and in the second equivalence we used the positivity of all shifted coordinates. The third equivalence follows by the fact that for every $x \in Q_{\text{corners}}$ and $y \in Q_{\text{seeds}}$ we have the following equivalences.

$$x < y \Longleftrightarrow \frac{x}{y} \leq \frac{1}{d'}.$$
$$x > y \Longleftrightarrow \frac{x}{y} \geq d'.$$

Furthermore, we may assume that $d' > 2$.

Now we observe

$$\sum_{j\in[d']}\left(\frac{a'_j}{q'_j}+\frac{q'_j}{b'_j}\right)=\langle\begin{bmatrix}a'_1\\ \frac{1}{b'_1}\\ \vdots\\ a'_{d'}\\ \frac{1}{b'_{d'}}\end{bmatrix},\begin{bmatrix}\frac{1}{q'_1}\\ q'_1\\ \vdots\\ \frac{1}{q'_{d'}}\\ q'_{d'}\end{bmatrix}\rangle=\langle\pi_{d'}(a',b'),\gamma_{d'}(q)\rangle.$$

We conclude

$$q^{(i)}\in\left(a^{(j)},b^{(j)}\right)\iff\langle\pi_{d'}(a',b'),\gamma_{d'}(q)\rangle<3.$$

Observe that the second condition just states that $\pi_{d'}(a'^{(j)},b'^{(j)})$ lies in some open halfspace in $\mathbb{R}^{2d'}$ that is determined by $\gamma_{d'}(q^{(i)})$.
This leads to the insight that $X:=\left\{\ \pi_{d'}(a'^{(j)},b'^{(j)})\ \middle|\ j\in[K]\ \right\}$ is shattered by the k-fold union of halfspaces in $\mathbb{R}^{2d'}$. Namely, for every $S\subseteq[K]$ we have found a set of k halfspaces in $\mathbb{R}^{2d'}$ such that for every $j\in[K]$: $\pi_{d'}(a'^{(j)},b'^{(j)})$ lies in the union of the halfspaces if and only if $j\in S$. Notice that when we ignore the point doubling at the start of the proof by property (i) of the pumpkin construction we even know that every point in $\left\{\ \pi_{d'}(a'^{(j)},b'^{(j)})\ \middle|\ j\in[K]\ \right\}$ lies in exactly one of the found hyperplanes. Note further that $|S|=K$ since $\pi_{d'}$ is injective, so the number of boxes does not change at shifting.
By $X\in\text{Shatter}\left(\mathscr{O}_d^{\cup,k}\right)$ we have that

$$\dim_{\text{VC}}\left(\mathscr{O}_d^{\cup,k}\right)\geq\dim_{\text{VC}}\left(\mathscr{O}_{2d'}^{\cup,k}\right)\geq K=(d'-1)(n+1)2^{n-2}\geq\frac{d}{4}\log_2(k)\,2^{\lfloor\log_2(k)\rfloor-1}\geq\frac{d}{16}k\log_2(k)\,.\qquad\square$$

Proof of Lemma 19. For $s\in\mathbb{N}_0$ and $(l_j)_{j\in[s]}\in\{0,1\}^s$ let us define

$$\alpha((l_j)_{j\in[s]}):=\sum_{j\in[s]}l_j2^{-j}$$
$$\beta((l_j)_{j\in[s]}):=\sum_{j\in[s]}l_j2^{-j}+2^{-s}$$

Here we interpret $\epsilon:=(l_j)_{j\in[0]}$ as the empty word and interpret empty sums to have the value zero. This just means that $\alpha(\epsilon)=0$ and $\beta(\epsilon)=1$.
Now for $t\in[n]$, $i\in[d-1]$ and $X\in\{0,1\}^{n-t}$, $Y\in\{0,1\}^t$ let us define the open box

$$\mathscr{I}_i(X,Y):=\left(\begin{pmatrix}0\\ \vdots\\ \alpha(X)\\ \alpha(Y)\\ \vdots\\ 0\end{pmatrix},\begin{pmatrix}1\\ \vdots\\ \beta(X)\\ \beta(Y)\\ \vdots\\ 1\end{pmatrix}\right)$$

where the entries $\alpha(X),\alpha(Y)$ and $\beta(X),\beta(Y)$ are at position i and $i+1$ respectively.

Using these definitions we can construct the set of boxes by

$$\begin{aligned}
\mathscr{A}_i(t) &:= \left\{ \mathscr{I}_i(X,Y) \;\middle|\; X \in \{0,1\}^{n-t-1} \times \{1\},\ Y \in \{0,1\}^{t-1} \times \{1\} \right\} && t \in [n-1],\ i \in [d-1]. \\
\mathscr{A}_i(t) &:= \left\{ \mathscr{I}_i(\epsilon,Y) \;\middle|\; Y \in \{0,1\}^{n-1} \times \{1\} \right\} && t = n,\ i \in [d-1]. \\
\mathscr{A}_i &:= \bigcup_{t\in[n]} \mathscr{A}_i(t) && i \in [d-1]. \\
\mathscr{A} &:= \bigcup_{i\in[d-1]} \mathscr{A}_i(t).
\end{aligned}$$

In Figure 9 the blue areas represent the boxes in $\mathscr{A}_i(t)$, $t \in [n]$. We check

$$|\mathscr{A}| = (d-1)\left((n-1)\,2^{n-2} + 2^{n-1}\right) = (d-1)\,(n+1)\,2^{n-2}.$$

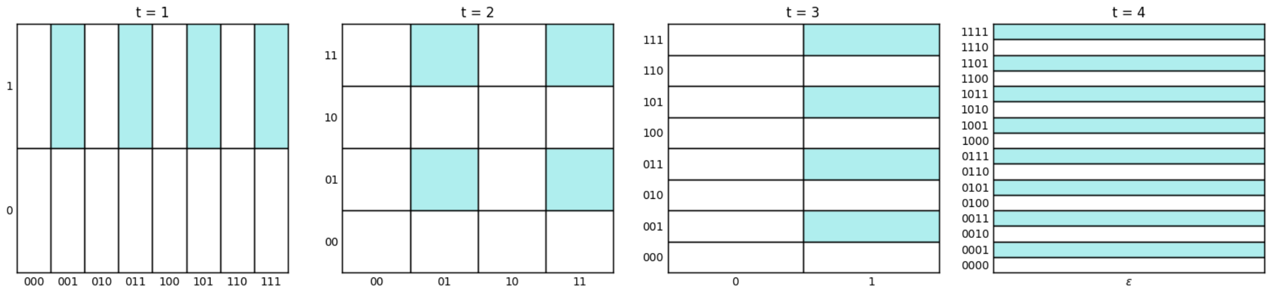

Figure 9: The sets $\mathscr{A}_i(t)$ for arbitrary i, $n = 4$ and $t \in [n]$

Now given a set of pumpkin pieces $S \subseteq \mathscr{A}$ we want to find a set of points $\mathcal{Q}(S)$ with the required properties. Let us call it set of seeds.

We will identify each point $p \in [0,1]^d$ in the set of seeds with its sequence $\left(\tilde{p}^{(i)}\right)_{i\in[d]} \subseteq \{0,1\}^n$ of the first n decimal places of the binary representations of its coordinates. Notice that we may restrict ourselves in picking the seeds in such a manner that they sit exactly in the middle of the boxes

$$\left\{ (a,b) \in \mathscr{B}_d \;\middle|\; \forall i \in [d]:\ a_i \in \left\{ \frac{z}{2^n} \;\middle|\; 0 \le z < 2^n \right\} \text{ and } b_i = a_i + 2^{-n} \right\}.$$

This just means

$$\forall i \in [d]:\ p_i = \alpha(\tilde{p}^{(i)}) + 2^{-(n+1)}.$$

We will construct the seeds in such a way that for any $i \in [d]$ the set of sequences of the first $n-1$ decimal places of the binary representations of the i-th coordinates of all the seeds cover $\{0,1\}^{n-1}$, formally

$$(\star)\ \forall i \in [d]:\ \left\{ \tilde{p}^{(i)} \;\middle|\; p \in \mathcal{Q}(S) \right\}\Big|_{[n-1]} = \{0,1\}^{n-1}.$$

Note that this condition does not restrict us in how we choose the n-th decimal places of the binary representations of the coordinates. Furthermore, for every $i \in [d]$ and $s \in \{0,1\}^{n-1}$ we know that there is a unique $p \in \mathcal{Q}(S)$ such that $\tilde{p}^{(i)} \in \{s\} \times \{0,1\}$.

To determine $\mathcal{Q}(S)$ it is enough to find surjective mappings $\phi_i : \{0,1\}^{n-1} \longrightarrow \{0,1\}^n$ such that for every $p \in \mathcal{Q}(S)$ we have that $\phi_i\left(\tilde{p}^{(i)}\big|_{[n-1]}\right) = \tilde{p}_{i+1}$ for every $i \in [d-1]$. The idea of this approach is that for the i-th mapping we only need to consider how the seeds are placed with respect to the boxes $S \cap \mathscr{A}_i$.

Let us consider a visualizing example construction of a mapping ϕ_i in Figure 10. As in Figure 9 we are assuming $n = 4$. In the top row the darkest fields are representing the boxes in $S \cap \mathscr{A}_i(t)$ $(t \in [4])$. The plot in the bottom right corner represents the mapping ϕ_i where its domain is aligned with the x-axis and its image is aligned with the y-axis. The conditions on ϕ_i transferred on the red fields of the grid in the bottom right plot are

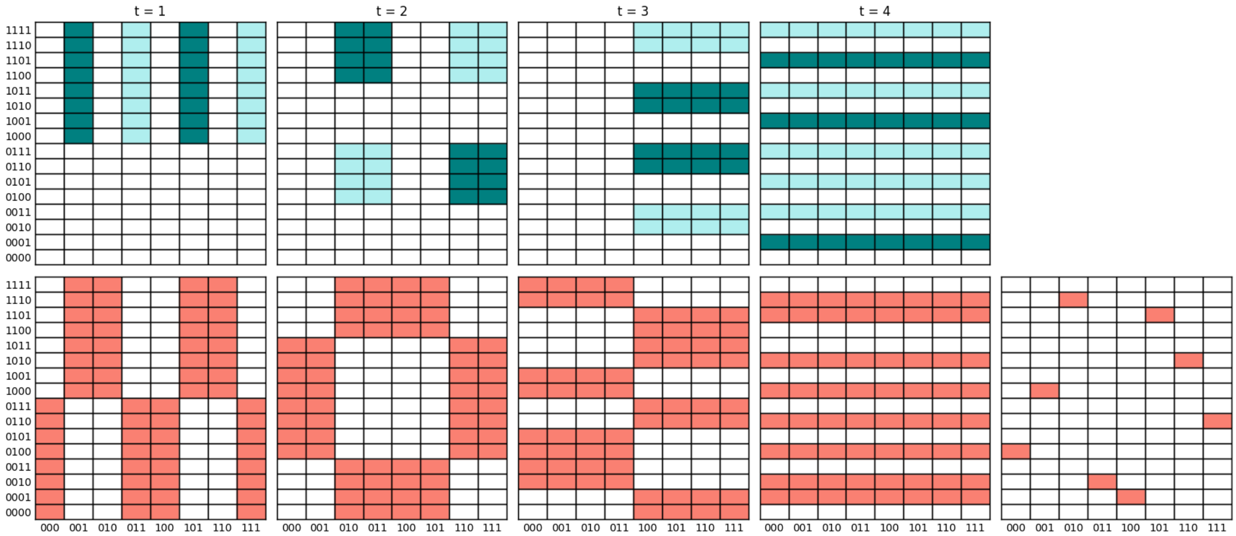

Figure 10: Construction of a mapping ϕ_i

(1) every column of the grid contains exactly one red field.

(2) for every $s \in \{0,1\}^3$ every pair of neighbored rows of the grid labeled with $\{s\} \times \{0,1\}$ contains exactly one red field.

(3) for every $t \in [4]$ and every box in $S \cap \mathscr{A}_i(t)$ exactly one of the fields corresponding to the box is filled red.

(4) for every $t \in [4]$ and every box in $\mathscr{A}_i(t)$ we have that if one of the fields corresponding to the box is filled red, then the box is in S.

We constructed the bottom right plot by intersecting the red areas of the other four bottom plots. In the t-th bottom plot for $X \in \{0,1\}^{n-t}$ and $Y \in \{0,1\}^t$ we filled the fields corresponding to the box $\mathscr{I}_i(X,Y)$ if and only if

$$t = n \text{ and } \mathbb{1}\left\{\mathscr{I}_i(\epsilon, (Y_1, \dots, Y_{(t-1)}, 1)) \in S\right\} = Y_t$$

or

$$t < n \text{ and } \mathbb{1}\left\{\mathscr{I}_i((X_1, \dots, X_{(n-t-1)}, 1), (Y_1, \dots, Y_{(t-1)}, 1)) \in S\right\} = \mathbb{1}\left\{X_{n-t} + Y_t \neq 1\right\}.$$

Observe that when intersecting the first bottom plot with the second, then intersect this intersection with the third bottom plot and so on in every intersection step the number of red fields in every column of the grid exactly halves, so there is exactly one red field in every column of the total intersection.
With this insights let us formally define ϕ_i. Consider $x \in \{0,1\}^{n-1}$ where we think of $x = \tilde{p}^{(i)}\big|_{[n-1]}$ for some seed p. We define the decimal places of $\phi_i(x)$ iteratively. For $0 \le j \le n-2$

$$(\phi_i(x))_{(1+j)} := \begin{cases} x_{(n-1-j)} & \mathscr{I}_i\left((x|_{[n-2-j]}, 1), (\phi_i(x)|_{[j]}, 1)\right) \in S \cap \mathscr{A}_{j+1} \\ 1 - x_{(n-1-j)} & \text{otherwise} \end{cases}$$

$$(\phi_i(x))_n := \begin{cases} 1 & \mathscr{I}_i\left(\epsilon, (\phi_i(x)|_{[n-1]}, 1)\right) \in S \cap \mathscr{A}_n \\ 0 & \text{otherwise} \end{cases}$$

Claim 6. $\forall 0 \le j \le n-2,\ \forall x \in \{0,1\}^{n-2-j} :\ \phi_i\left(\{x\} \times \{0,1\}^{j+1}\right)\Big|_{[j+1]} = \{0,1\}^{j+1}$.

Notice that for $j = n-2$ this simply states $(\star)$.

Proof of Claim 6. We do induction on j.
base $j = 0$**.** Let $x \in \{0,1\}^{n-2}$. Then for any $y \in \{0,1\}$

$$\phi_i(x,y)_1 = \begin{cases} y & \mathscr{I}_i\big((x_1,\ldots,x_{(n-2)},1),(1)\big) \in S \\ 1-y & \text{otherwise} \end{cases}$$

Thus, we deduce $\phi_i(x,\bullet)_1 \in \{\bullet, 1-\bullet\}$ which yields the Claim.
step $0 \le j < n-2$**.** Fix $x \in \{0,1\}^{n-2-(j+1)}$. Induction yields

$$\forall y \in \{0,1\}: \ \phi_i\left(\{(x,y)\} \times \{0,1\}^{j+1}\right)\Big|_{[j+1]} = \{0,1\}^{j+1}. \tag{1}$$

Let $y_1, y_2 \in \{0,1\}$ and $x'_1, x'_2 \in \{0,1\}^{j+1}$ such that $\phi_i(x,y_1,x'_1)\big|_{[j+2]} = \phi_i(x,y_2,x'_2)\big|_{[j+2]}$. Notice that by Definition

$$\phi_i(x,y_1,x'_1)_{(1+(j+1))} = \begin{cases} y & \mathscr{I}_i\left((x,1),(\phi_i(x,y_1,x'_1)\big|_{[j+1]},1)\right) \in S \\ 1-y & \text{otherwise} \end{cases}$$

Since $\phi_i(x,y_1,x'_1)\big|_{[j+1]} = \phi_i(x,y_2,x'_2)\big|_{[j+1]}$ and $\phi_i(x,y_1,x'_1)_{(1+(j+1))} = \phi_i(x,y_2,x'_2)_{(1+(j+1))}$ we conclude that $y_1 = y_2$. However, now (1) yields that also $x_1 = x_2$.
This shows injectivity of the mapping $\phi_i(x,\bullet)\big|_{[j+2]}$ and an argument about cardinality of the image of this mapping yields that it is also surjective. This proves the step. ☐

Claim 7. $\forall i \in [d-1],\ t \in [n],\ A = \mathscr{I}_i(X,Y) \in \mathscr{A}_i(t)$ we have

(i') $A \in S \Longrightarrow \exists! X' \in \{0,1\}^{t-1}: \ \phi_i(X,X') \in \left(\{Y\} \times \{0,1\}^{n-t}\right)$

(ii') $A \notin S \Longrightarrow \phi_i\left(X \times \{0,1\}^{t-1}\right) \cap \left(\{Y\} \times \{0,1\}^{n-t}\right) = \emptyset.$

Proof of Claim 7. Notice that X has length $n-t$ whereas Y has length t.
case $t = n$**.** Observe that $X = \epsilon$. Claim 6 with $j = n-2$ states that $\phi_i\left(\{0,1\}^{n-1}\right)\Big|_{[n-1]} = \{0,1\}^{n-1}$. Thus, there is a unique $X' \in \{0,1\}^{n-1}$ such that $\phi_i(X') = Y\big|_{[n-1]}$.
Following the Definition we have

$$(\phi_i(X'))_n := \begin{cases} 1 & \mathscr{I}_i\left(\epsilon,(Y\big|_{[n-1]},1)\right) \in S \\ 0 & \text{otherwise} \end{cases}$$

Note that since $\mathscr{I}_i(X,Y) \in \mathscr{A}_i(n)$ we have that $Y_n = 1$. Hence,

$$Y \in \phi_i\left(\{0,1\}^{n-1}\right) \Longleftrightarrow \phi_i(X') = Y \Longleftrightarrow \phi_i(X')_n = 1 \Longleftrightarrow \mathscr{I}_i(X,Y) \in S$$

which proves the Claim since in case that $\mathscr{I}_i(X,Y) \in S$ the solution X' is unique.
case $1 < t < n$**.** Claim 6 with $j = t-2$ states that $\phi_i\left(\{X\} \times \{0,1\}^{t-1}\right)\Big|_{[t-1]} = \{0,1\}^{t-1}$. Hence, there is a unique $X' \in \{0,1\}^{t-1}$ such that $\phi_i(X,X')\big|_{[t-1]} = Y\big|_{[t-1]}$. Following the Definition with $j = t-1$ we have

$$\begin{aligned}(\phi_i(X,X'))_t &= \begin{cases} X_{(n-t)} & \mathscr{I}_i\big((X_1,\ldots,X_{(n-t-1)},1),(Y_1,\ldots,Y_{t-1},1)\big) \in S \\ 1-X_{(n-t)} & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \mathscr{I}_i(X,Y) \in S \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

where again $X_{(n-t)} = 1 = Y_t$ follows by $\mathscr{I}_i(X,Y) \in \mathscr{A}_i(t)$.

We conclude $Y \in \phi_i\left(\{X\} \times \{0,1\}^{t-1}\right)\Big|_{[t]} \iff \phi_i\left(X, X'\right)\Big|_{[t]} = Y$

$$\iff \left(\phi_i\left(X, X'\right)\right)_t = Y_t$$
$$\iff \left(\phi_i\left(X, X'\right)\right)_t = 1$$
$$\iff \mathscr{I}_i\left(X, Y\right) \in S.$$

which proves the Claim since again in case that $\mathscr{I}_i\left(X, Y\right) \in S$ the solution X' is unique.
case $t = 1$**.** Following the Definition with $j = 0$ we have

$$\left(\phi_i\left(X\right)\right)_1 = \begin{cases} 1 & \mathscr{I}_i\left(X, Y\right) \in S \\ 0 & \text{otherwise} \end{cases}$$

where we used $X_{(n-1)} = 1 = Y_1$ as in the former case. Analogously we conclude

$$\phi_i\left(X\right) \in \left(\{Y\} \times \{0,1\}^{n-1}\right) \iff \left(\phi_i\left(X\right)\right)_1 = 1 \iff \mathscr{I}_i\left(X, Y\right) \in S,$$

which proofs the Claim since in case that $\mathscr{I}_i\left(X, Y\right) \in S$ the solution $X' = \epsilon$ is unique. □

Now we are able to us close the proof. For $S \subseteq \mathscr{A}$ we defined the set of seeds $\mathcal{Q}(S)$ by iteratively defining the binary representation of the coordinates of $\tilde{p}$ for every $p \in \mathcal{Q}(S)$. Note that by $(\star)$ we know that $|\mathcal{Q}(S)| = 2^{n-1}$. Furthermore, the conditions (i) on (ii) follow by the conditions (i') and (ii') in Claim 7. □

3.4 Generalised δ-packings

Let $\mathscr{F} := (X,\ \mathscr{E})$ be a hypergraph. Unlike as in the other sections here we explicitly allow double hyperedges, meaning $\mathscr{E}$ is a multiset. We remark that we keep the multiplicity of the edges when taking traces and also respect the multiplicity of the edges when determining the shatter function. We want to introduce a generalization of the Hamming distance between edges of $\mathscr{F}$. Let $k \in \mathbb{N}$, $\left(A_j\right)_{j \in [k]} \subseteq \mathscr{E}$.

Definition 65. $\Delta\left(\left(A_j\right)_{j\in[k]}\right) := \left(\bigcup\limits_{j\in[k]} A_j\right) \setminus \left(\bigcap\limits_{j\in[k]} A_j\right).$

Definition 66. $\text{disparity}\left(\left(A_j\right)_{j\in[k]}\right) := \left|\Delta\left(\left(A_j\right)_{j\in[k]}\right)\right|.$

We remark that $0 \leq \text{disparity}\left(\left(A_j\right)_{j\in[k]}\right) \leq |\mathscr{H}|$. The disparity is a metric in the following sense.

Observation 25. For any $l_A, l_B, l_C \in \mathbb{N}$ and $\left(A_j\right)_{j\in[l_A]},\ \left(B_i\right)_{i\in[l_B]},\ \left(C_j\right)_{j\in[l_C]} \subseteq \mathscr{E}$

(M1) $\text{disparity}\left(\left(A_j\right)_{j\in[l_A]}\right) = 0 \iff \forall i, j \in [l_A]:\ A_i = A_j.$

(M2) $\forall \sigma \in \mathscr{S}_{l_A}:\ \text{disparity}\left(\left(A_j\right)_{j\in[l_A]}\right) = \text{disparity}\left(\left(A_{\sigma(j)}\right)_{j\in[l_A]}\right).$

(M3) $\text{disparity}\left(\left(A_j\right)_{j\in[l_A]}, \left(C_j\right)_{j\in[l_C]}\right) \leq \text{disparity}\left(\left(A_j\right)_{j\in[l_A]}, \left(B_j\right)_{j\in[l_B]}\right) + \text{disparity}\left(\left(B_j\right)_{j\in[l_B]}, \left(C_j\right)_{j\in[l_C]}\right).$

Proof of Observation 25. (M1) follows since the disparity is zero if and only if $\bigcup\limits_{j\in[l_A]} A_j = \bigcap\limits_{j\in[l_A]} A_j$, which again is equivalent to the right-hand side. (M2) follows directly from the symmetry of the union and intersection. To show (M3) let us denote $X := A_\cap \cap B_\cap \cap C_\cap$ and

$$A_\cup := \bigcup_{j\in[l_A]} A_j, \qquad A_\cap := \bigcap_{j\in[l_A]} A_j,$$
$$B_\cup := \bigcup_{j\in[l_B]} B_j, \qquad B_\cap := \bigcap_{j\in[l_B]} B_j,$$
$$C_\cup := \bigcup_{j\in[l_C]} C_j, \qquad C_\cap := \bigcap_{j\in[l_C]} C_j.$$

We observe

$$\begin{aligned}\text{disparity}\left((A_j)_{j\in[l_A]},(C_j)_{j\in[l_C]}\right) &= |\,(A_\cup\cup C_\cup)\setminus(A_\cap\cap C_\cap)\,| \\ &\le |\,(A_\cup\cup B_\cup)\setminus(A_\cap\cap B_\cap)\,| + |\,(B_\cup\cup C_\cup)\setminus(B_\cap\cap C_\cap)\,| \\ &= \text{disparity}\left((A_j)_{j\in[l_A]},(B_j)_{j\in[l_B]}\right) + \text{disparity}\left((B_j)_{j\in[l_B]},(C_j)_{j\in[l_C]}\right).\end{aligned}$$

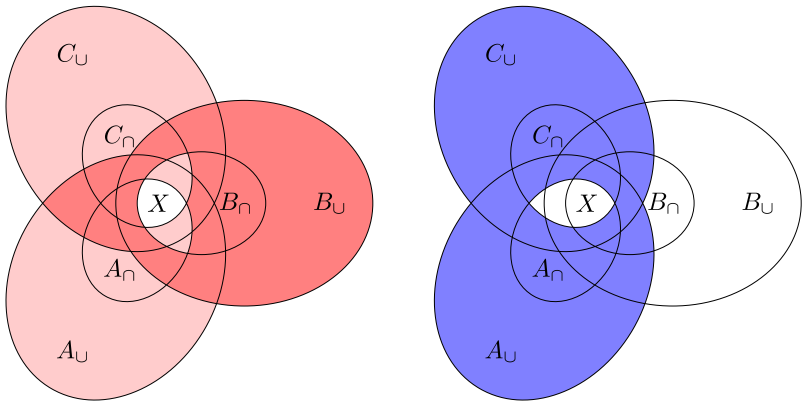

Figure 11: Venn Diagrams to prove (M3).

Since splitting the sum into its different terms is not enlightening we give a visual proof instead. In Figure 11 on the left-hand side we have depicted the term

$$|\,(A_\cup\cup B_\cup)\setminus(A_\cap\cap B_\cap)\,| + |\,(B_\cup\cup C_\cup)\setminus(B_\cap\cap C_\cap)\,|.$$

The elements in the dark red areas are counted twice, the elements in the bright red areas are counted once. On the right-hand side we have depicted the term

$$|\,(A_\cup\cup C_\cup)\setminus(A_\cap\cap C_\cap)\,|.$$

The elements in the dark blue areas are counted once. Considering Figure 11 it is easy to check that the claimed inequality certainly holds. □

We draw some consequence of the triangle inequality (M3) that we use in the proof of the Packing lemma, Theorem 10.

Observation 26. Let $k,t\in\mathbb{N}_{\ge 2}$ and T be a tree on vertices $[t]$ as well as $(A_j)_{j\in[k]}\subseteq\mathscr{F}$ as well as $(\mathscr{I}_i)_{i\in[t]}\subseteq[k]$ index sets that cover $[k]$, formally $[k]=\bigcup_{i\in[t]}\mathscr{I}_i$. Then

$$\text{disparity}\left((A_j)_{j\in[k]}\right)\le\sum_{\{i_1,i_2\}\in E(T)}\text{disparity}\left((A_j)_{j\in\mathscr{I}_{i_1}\cup\mathscr{I}_{i_2}}\right).$$

Proof of Observation 26. We prove the Claim by induction on t.
base $t=2$**.** This case is trivial.
base $t\ge 3$**.** Let $v\in[t]$ be a leaf of T and $u\in[t]$ be its unique neighbor. Let us define $\mathscr{I}:=\bigcup_{i\in[t]\setminus\{v\}}\mathscr{I}_i$. We

apply (M3) and induction to obtain

$$\begin{aligned}\text{disparity}\left((A_j)_{j\in[k]}\right) &\leq \text{disparity}\left((A_j)_{j\in\mathscr{J}_v},(A_j)_{j\in\mathscr{J}_u}\right)+\text{disparity}\left((A_j)_{j\in\mathscr{J}}\right)\\ &\leq \text{disparity}\left((A_j)_{j\in\mathscr{J}_v},(A_j)_{j\in\mathscr{J}_u}\right)+\sum_{\{i_1,i_2\}\in E(T-v)}\text{disparity}\left((A_j)_{j\in\mathscr{J}_{i_1}\cup\mathscr{J}_{i_2}}\right)\\ &=\sum_{\{i_1,i_2\}\in E(T)}\text{disparity}\left((A_j)_{j\in\mathscr{J}_{i_1}\cup\mathscr{J}_{i_2}}\right).\end{aligned}$$ □

Let us define a generalization of a δ-packing.

Definition 67 ((k,δ) - separated)**.** For $k,d\in\mathbb{N}$ we say $\mathscr{F}$ is (k,δ)-separated when

$$\forall\{\,A_j\mid j\in[k]\,\}\in\binom{\mathscr{E}}{k}:\ \text{disparity}\left((A_j)_{j\in[k]}\right)\geq\delta.$$

In this case we also say that $\mathscr{F}$ is a (k,δ)-packing. In case that $k=2$ we simply say $\mathscr{F}$ is δ-separated or $\mathscr{F}$ is a δ-packing.

We are going to present a proof of a Packing lemma from Fox et al. It is going to be the main tool in the proof of the Ultra Strong Regularity lemma, Theorem 27. Furthermore, we apply it in its full generality in our counting framework for induced isomorphisms in Theorem 19.

Theorem 10 (Packing lemma, Fox, Pach, Sheffer, Suk, Zahl [18])**.** $\forall k,d\in\mathbb{N}\ \forall c>0\ \exists C>0$ such that $\forall\delta,m\in\mathbb{N}$ and all (k,δ)-separated hypergraphs $\mathscr{F}=([m]\,,\ \mathscr{E})$ fulfilling $\forall z\in\mathbb{N}:\ \pi_{\mathscr{F}}(z)\leq cz^d$ one finds

$$\|\mathscr{F}\|\leq C\left(\frac{m}{\delta}\right)^d.$$

Proof of Theorem 10. Fix $C>0$. Assume $\exists m\in\mathbb{N}$ and $\mathscr{F}\subset 2^{[m]}$ fulfilling the requirements but

$$\|\mathscr{F}\|>C\left(\frac{m}{\delta}\right)^d.$$

The polynomial bound on the shatter function and Lemma 10 give us

$$\dim_{\mathrm{VC}}(\mathscr{F})\leq d_0:=4d\log_2(cd)\,.$$

We might have chosen $C\geq 2c\left(16d_0k^2\right)^d$. By the polynomial restriction on the shatter function it follows that

$$\|\mathscr{F}\|\leq cm^d.$$

Putting this together with our assumption on $\|\mathscr{F}\|$ and the choice of C we see deduce

$$2c\left(16d_0k^2\right)^d\left(\frac{m}{\delta}\right)^d<cm^d.$$

calculation yields

$$\delta>16d_0k^2.$$

Let us fix $s:=\left\lceil 8d_0(k-1)\cdot\frac{m}{\delta}\right\rceil$. By the previous we have $s<m$. Observe that by the polynomial restriction on the shatter function as well as our choice of s and C for any $S'\in\binom{[m]}{s-1}$

$$(2k+1)\cdot|\mathscr{E}\cap S'|\leq(2k+1)\cdot\pi_{\mathscr{H}}(s)\leq(2k+1)\cdot c\cdot\left\lfloor\frac{8d_0(k-1)m}{\delta}\right\rfloor^d\leq\frac{C}{2}\cdot\left(\frac{m}{\delta}\right)^d. \tag{2}$$

Seeking some contradiction we are going to double count certain sums of weights we are going to define now.

Given $S \subseteq [m]$ and $A \subseteq S$ let us define the class of edges that look like A when restricted to S.

$$\mathscr{E}(S, A) := \{ F \in \mathscr{E} \mid F \cap S = A \}.$$

Let us interpret the count of these edges as a weight $w_S(A) := |\mathscr{E}(S, A)|$. We will lift this weight to the edges of $\mathbf{UD}\left(\mathscr{F}\big|_S\right)$ by defining for $A, B \in \mathscr{E} \cap S$ with $|(A \Delta B)| = 1$

$$w_S(\{A, B\}) := \min\{w_S(A), w_S(B)\}.$$

Now we can define the weight of S simply by the sum of all edge weights of $\mathbf{UD}\left(\mathscr{F}\big|_S\right)$.

$$w(S) := \sum_{e \in E\left(\mathbf{UD}\left(\mathscr{F}\big|_S\right)\right)} w_S(e).$$

The idea of the whole proof is to double count $Z := \sum_{S \in \binom{[m]}{s}} w(S)$.

Regarding the upper bound consider some $S \in \binom{[m]}{s}$. Let us show that $\mathbf{UD}\left(\mathscr{F}\big|_S\right)$ is $(2d_0)$-degenerative. Since $\forall S' \subseteq S$ we have $\dim_{\mathrm{VC}}\left(\mathscr{F}\big|_{S'}\right) \leq d_0$ Lemma 13 yields that

$$\|\mathbf{UD}\left(\mathscr{F}\big|_{S'}\right)\| \leq d_0 \cdot |\mathbf{UD}\left(\mathscr{F}\big|_{S'}\right)|,$$

and we always find a vertex in S' of degree less or equal $2d_0$ in $\mathbf{UD}\left(\mathscr{F}\big|_{S'}\right)$ by the pigeonhole principle.
Thus, we can conclude by iteratively deleting vertices of minimal degree in S and adding the weights of the adjacent edges to the overall weight sum that

$$w(S) \leq \sum_{A \in \mathscr{E} \cap S} 2d_0 \cdot w_S(A) = 2d_0 \cdot \|\mathscr{F}\|,$$

where we used that the weight of an edge is at most the weight of any vertex adjacent to it.
Thus, we obtain the upper bound

$$Z \leq \binom{m}{s} \cdot 2d_0 \cdot \|\mathscr{F}\|.$$

In the sequel we are going to prove the lower bound

$$Z \geq \frac{\delta}{2(k-1)} \binom{m}{s-1} \left(\|\mathscr{F}\| - \frac{C}{2} \left(\frac{m}{\delta}\right)^d \right).$$

Those inequalities together with the Definition of s yield

$$\|\mathscr{F}\| \geq \frac{s\delta}{4md_0(k-1)} \left(\|\mathscr{F}\| - \frac{C}{2} \left(\frac{m}{\delta}\right)^d \right) \geq 2 \left(\|\mathscr{F}\| - \frac{C}{2} \left(\frac{m}{\delta}\right)^d \right).$$

Using this we calculate

$$\|\mathscr{F}\| \leq C \left(\frac{m}{\delta}\right)^d$$

which is a contradiction to our assumption and thereby proves Theorem 10.

It is left to show the lower bound of the double counting argument. We are going to use the following Claim.

Claim 8. Let $S' \subseteq [m]$ and $B \subseteq S'$. Then

$$\sum_{A_1, A_2 \in \mathscr{E}(S', B)} |A_1 \Delta A_2| \geq \frac{\delta \cdot w_{S'}(B)(w_{S'}(B) - 2k + 1)}{2(k-1)}.$$

Proof of Claim 8. For demonstrating purposes let us define a help graph

$$H := \left(\mathscr{E}(S', B), \ \left\{ \{A_1, A_2\} \subseteq \mathscr{E}(S', B) \ \middle| \ |A_1 \Delta A_2| \geq \frac{\delta}{k-1} \right\} \right).$$

We remark that by Definition $|H| = w_{S'}(B)$. Assume for a contradiction that $\|\overline{H}\| \geq (k-1)|H|$. It is easy to see that in this case there would be a subgraph $\overline{H}' \subseteq \overline{H}$ with minimal degree $k-1$. Thus, one could greedily find a tree $T \subseteq \overline{H}'$ of order k. Then it follows with Observation 26, where we choose the index sets containing one index each, that

$$\text{disparity}\left(V\left(T\right)\right) \leq \sum_{\{A,B\} \in E(T)} \text{disparity}\left(A, B\right) < (k-1)\frac{\delta}{k-1} = \delta.$$

However, this is a contradiction to the assumption that $\mathscr{F}$ is (k, δ)-separated. Thus, we know that

$$\|H\| = \binom{|H|}{2} - \|\overline{H}\| > \binom{|H|}{2} - (k-1)|H| = \frac{|H|(|H| - 2k + 1)}{2}.$$

We deduce

$$\sum_{A_1, A_2 \in \mathscr{E}(S', B)} |A_1 \Delta A_2| \geq \frac{\delta}{k-1}\|H\| \geq \frac{\delta \cdot w_{S'}(B)(w_{S'}(B) - 2k + 1)}{2(k-1)},$$

which completes the proof of Claim 8. □

For clarity, we chose to format the proof of the lower bound for Z as a long sequence of commented inequalities.

First we measure the weight of each $S \in \binom{[m]}{s}$ by adding the weight of any edge of $\mathbf{UD}\left(\mathscr{F}\big|_S\right)$ from the perspective of the vertex with less cardinality.

$$Z = \sum_{S \in \binom{[m]}{s}} w(S) = \sum_{S \in \binom{[m]}{s}} \left(\sum_{x \in S} \left(\sum_{\substack{B \in (\mathscr{E} \cap S) \\ x \notin B}} \min\left\{w_S(B), w_S(B \cup \{x\})\right\} \right) \right)$$

Then we use that for any non-negative reals a, b one has $\min\{a, b\} \geq \mathbb{1}\{a + b \neq 0\} \frac{a \cdot b}{a+b} = \frac{a \cdot b}{a+b}$ where for notational simplicity here we interpret $\frac{0}{0} := 0$.

$$\geq \sum_{S \in \binom{[m]}{s}} \left(\sum_{x \in S} \left(\sum_{B \subseteq S \setminus \{x\}} \frac{w_S(B) \cdot w_S(B \cup \{x\})}{w_S(B) + w_S(B \cup \{x\})} \right) \right)$$

Observe that for any $B \subseteq S \setminus \{x\}$ the number of hyperedges in $\mathscr{F}$ whose intersection with S is either B or $B \cup \{x\}$ equals the number of hyperedges whose intersection with $S \setminus \{x\}$ is B.

$$= \sum_{S \in \binom{[m]}{s}} \left(\sum_{x \in S} \left(\sum_{B \subseteq S \setminus \{x\}} \frac{w_S(B) \cdot w_S(B \cup \{x\})}{w_{(S \setminus \{x\})}(B)} \right) \right)$$

Let us reorder the sum and substitute $S' = S \setminus \{x\}$.

$$= \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \frac{1}{w_{S'}(B)} \left(\sum_{x \in [m] \setminus S'} w_{(S' \cup \{x\})}(B) \cdot w_{(S' \cup \{x\})}(B \cup \{x\}) \right) \right)$$

Count the weights with sums of indicators. Observe that a hyperedge of $\mathscr{F}$ intersects with $S' \cup \{x\}$ in B if and only if it intersects with S' in B, and it does not contain x. Furthermore, a hyperedge of $\mathscr{F}$ intersects with $S' \cup \{x\}$ in $B \cup \{x\}$ if and only if it intersects with S' in B, and it contains x.

$$= \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \frac{1}{w_{S'}(B)} \left(\sum_{x \in [m] \setminus S'} \left(\sum_{A_1 \in \mathscr{E}(S',B)} \mathbb{1}\{x \notin A_1\} \left(\sum_{A_2 \in \mathscr{E}(S',B)} \mathbb{1}\{x \in A_2\} \right) \right) \right) \right)$$

Change the order of summation and use the identity $\mathbb{1}\{x \notin A_1\} \mathbb{1}\{x \in A_2\} = \mathbb{1}\{x \in A_1 \Delta A_2\}$.

$$= \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \frac{1}{w_{S'}(B)} \left(\sum_{A_1 \in \mathscr{E}(S',B)} \left(\sum_{A_2 \in \mathscr{E}(S',B)} \left(\sum_{x \in [m] \setminus S'} \mathbb{1}\{x \in A_1 \Delta A_2\} \right) \right) \right) \right)$$

Evaluating the innermost sum. Observe that for $A_1, A_2 \in \mathscr{E}(S', B)$ no vertex in S' lies in their symmetric difference.

$$= \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \frac{1}{w_{S'}(B)} \left(\sum_{A_1, A_2 \in \mathscr{F}_{S'}(B)} |A_1 \Delta A_2| \right) \right)$$

Now we apply Claim 8 to the innermost sum.

$$\begin{aligned}
&\geq \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \frac{1}{w_{S'}(B)} \left(\frac{\delta \cdot w_{S'}(B)(w_{S'}(B) - 2k + 1)}{2(k-1)} \right) \right) \\
&= \frac{\delta}{2(k-1)} \sum_{S' \in \binom{[m]}{s-1}} \left(\sum_{B \subseteq S'} \mathbb{1}\{w_{S'}(B) \neq 0\} (w_{S'}(B) - 2k + 1) \right) \\
&= \frac{\delta}{2(k-1)} \sum_{S' \in \binom{[m]}{s-1}} \left(\left(\sum_{B \subseteq S'} w_{S'}(B) \right) - (2k+1)|\mathscr{E} \cap S'| \right)
\end{aligned}$$

The sum of all weights of subsets of S' is simply the number of hyperedges in $\mathscr{F}$. Finally, we apply (2).

$$\geq \frac{\delta}{2(k-1)} \binom{m}{s-1} \left(\|\mathscr{F}\| - \frac{C}{2} \left(\frac{m}{\delta} \right)^d \right).$$

This closes the proof of Theorem 10. □

4 Main results

In this section we present and prove our two main results about the extremal properties of bipartite graphs H fulfilling certain degree conditions. Namely, for parameter $d \in \mathbb{N}$, we require that H has a partite set A where any vertex $a \in A$ either has full degree or degree at most d. In Theorem 16, in case that H is $K_{d,d}$-free, we show $\mathrm{ex}\left(n, \{K_{s,s}, H\text{-biind}\}\right) = o\left(n^{2-\frac{1}{d}}\right)$, where $s \in \mathbb{N}$ is an arbitrary integer. Theorem 19 provides a counting result for the number of induced labeled copies of H in a host graph $G \in \mathrm{Free}\left(K_{d+1,d+1}\right)$, in case that $\|G\| \geq C|G|^{2-\frac{1}{d}}$ for some constant $C = C(H, d)$.
Section 4.1 gives an introduction to the problem and surveys related results. The proof of Theorem 16 is presented in section 4.2, the proof of Theorem 19 can be found in section 4.3.

4.1 Introduction to the extremal properties of the hedgehog

The idea of bounding the extremal function of bipartite graphs in case that they fulfill certain degree conditions goes back to Füredi.

Theorem 11 (Füredi [23])**.** $\forall d \in \mathbb{N}$ and any bipartite graph $H = (A \cup B,\ F)$ that fulfills $\max_{b\in B}\deg_A(b) \leq d$ one finds

$$\mathrm{ex}\left(n, H\right) = O\left(n^{2-\frac{1}{d}}\right).$$

We remark that this result was later reproved with help of the Dependent Random Choice technique by Alon, Krivelevich and Sudakov in [2]. Some algebraic constructions for the Zarankievicz problem show that Theorem 11 is tight, see Lemma 4.
Conlon and Lee were able to improve the exponent in the bound of Theorem 11 in case $d = 2$ and $K_{2,2} \nsubseteq H$. To properly state their result we first need the Definition of subdivision.

Definition 68 (Subdivision)**.** Let G be a graph. Then we define its *subdivision* as

$$G' \coloneqq \mathrm{Incidence}\left(E\left(G\right)\right).$$

This means that every edge in G got replaced by a path on three vertices in G'. Furthermore, for $t \in \mathbb{N}$ we want to define the special subdivision

$$H_t \coloneqq \left(K_t\right)' = \mathrm{Incidence}\left(\binom{t}{2}\right).$$

In the notion of Definition 20 this is just a $(k, 2, 1)$-hedgehog.

Observation 27. Let $H = (A \cup B,\ F)$ be a bipartite graph that fulfills $\max_{b\in B}\deg_A(b) = 2$ and $K_{2,2} \nsubseteq H$. Then there is $t \in \mathbb{N}$ such that

$$H \underset{ind}{\subseteq}^{*} H_t.$$

Theorem 12 (Conlon and Lee [11])**.** $\forall t \in \mathbb{N}:\ \mathrm{ex}\left(n, H_t\right) = O\left(n^{\frac{3}{2}-\frac{1}{6^t}}\right)$.

Theorem 12 got improved by Janzer who simplified their arguments and showed the following

Theorem 13 (Janzer [31])**.** $\forall t \in \mathbb{N}:\ \mathrm{ex}\left(n, H_t\right) = O\left(n^{\frac{3}{2}-\frac{1}{4t-6}}\right)$.

Theorem 13 is tight for $t = 3$ since in this case $H_3 = C_6$ and $\mathrm{ex}\left(n, C_6\right) = \Theta\left(n^{\frac{4}{3}}\right)$, see [8] for the upper bound and [5] for the lower bound. Conlon and Lee conjectured that the equivalent should hold in the general case.

Conjecture 1. $\forall d \in \mathbb{N}$ and any bipartite graph $H = (A \cup B,\ F)$ that fulfills $\max_{b\in B}\deg_A(b) \leq d$ and $K_{d,d} \nsubseteq H$ one finds some $\delta > 0$ such that $\mathrm{ex}\left(n, H\right) = O\left(n^{2-\frac{1}{d}-\delta}\right)$.

A further step towards this Conjecture was made in 2019 by Sudakov and Tomon.

Theorem 14 (Sudakov and Tomon [44])**.** For every integer $d \geq 2$ and every bipartite graph H such that $K_{d,d} \nsubseteq H$ and in one partite set every vertex has degree at most d one finds $\mathrm{ex}(n, H) = o\left(n^{2-\frac{1}{d}}\right)$.

Their proof uses the Hypergraph Removal lemma, Theorem 5. A similar approach was taken by Janzer and Pohoata in the induced setting. They showed a similar upper bound for the number of edges of a bipartite graph that does not contain $K_{t,t}$ as a subgraph and has bounded VC dimension. For the Definition of this notion we refer the reader to section 3, especially to Definition 54 and the Theorem 7, which states their result as in the original paper. For our convenience we restate the result at this point.

Theorem 15 (Janzer and Pohoata [32])**.** $\forall d, t \in \mathbb{N}:\ \mathrm{ex}^*\left(K_{n,n}, \left\{K_{t,t}, \mathrm{Incidence}\left(2^{[d+1]}\right)\text{-ind}\right\}\right) = o\left(n^{2-\frac{1}{d}}\right)$.

We were able to merge the ideas of Theorem 14 and 15 to obtain an even stronger statement. Our main result, that we published in [4], states that

Theorem 16 (Main result)**.** Let $d, t \in \mathbb{N},\ d \geq 2$ and $H = (A \cup B,\ F)$ be a bipartite graph such that $\forall b \in B:$ $\deg_A(b) \in [d] \cup \{|A|, 0\}$ as well as $K_{d,d} \nsubseteq H$. Then

$$\mathrm{ex}^*\left(K_{n,n}, \{K_{t,t}, H\text{-ind}\}\right) = o\left(n^{2-\frac{1}{d}}\right).$$

This implies

$$\mathrm{ex}\left(n, \{K_{t,t}, H\text{-biind}\}\right) = o\left(n^{2-\frac{1}{d}}\right).$$

We present the proof of Theorem 16 at the end of section 4.2.
We remark that Theorem 16 implies Theorem 15 since for $d \in \mathbb{N}_{\geq 3}$ one finds that $\mathrm{Incidence}\left(2^{[d+1]}\right)$ is $K_{d,d}$-free and fulfills the degree restriction with one complete vertex. We remark that for $d \in \{1, 2\}$ one finds $K_{d,d} \subseteq \mathrm{Incidence}\left(2^{[d+1]}\right)$.
Furthermore, Theorem 16 implies Theorem 14 as can be seen using Observation 5 and 8.
The following class of bipartite graphs will turn out useful in order to prove our main result, Theorem 16.

Definition 69. Let $d, k \in \mathbb{N},\ r \in \mathbb{N}_0$ with $k \geq d \geq r + 2$. Define $W(k, d, r)$ as the bipartite graph with left-hand side of size k such that all d-sets on the left-hand side have exactly $(d - r - 1)$ common neighbors and additionally there are r complete vertices on the right-hand side. Formally let L, X be disjoint sets with $|L| = k,\ |X| = r$. Then we define $Y := \left\{ (i, A) \,\middle|\, i \in [d - r - 1],\ A \in \binom{L}{d} \right\}$ and

$$W(k, d, r) := \left(L \cup (X \cup Y),\ \{\, \{v, x\} \mid v \in L, x \in X \,\} \cup \{\, \{l, (i, A)\} \mid l \in A,\ (i, A) \in Y \,\}\right).$$

In reference to the $(k, d, d - r - 1)$-*hedgehog* in $W(k, d, r)$ between L and Y we want to call L the *body* of $W(k, d, r)$. See Figure 12 for a rendering of $W(6, 5, 1)$.

In order to prove our main result, Theorem 16, we need the following.

Theorem 17. Let r, d, k, t be non-negative integers fulfilling $k \geq d \geq r + 2$. Then

$$\mathrm{ex}^*\left(K_{n,n}, \{K_{t,t}, W(k, d, r)\text{-ind}\}\right) = o\left(n^{2-\frac{1}{d}}\right).$$

Notice that by Observation 5 and Kővári, Sós, Turán we have $\mathrm{ex}^*\left(K_{n,n}, \{K_{t,t}, W(k, d, r)\text{-ind}\}\right) \leq \mathrm{ex}^*\left(K_{n,n}, K_{t,t}\right) = O\left(n^{2-\frac{1}{t}}\right)$ so the case $t < d$ is not interesting.
We present the proof of Theorem 17 in section 4.2. Observe that Theorem 17 is sharp in the following sense.

Observation 28. For every $\epsilon > 0$ and non-negative integers r, d, t with $d \geq r + 2$ and $t \geq 2d - 1$ there is some $k \in \mathbb{N}$ such that

$$\mathrm{ex}^*\left(K_{n,n}, \{K_{t,t}, W(k, d, r)\text{-ind}\}\right) = \Omega\left(n^{2-\frac{1}{d}-\epsilon}\right).$$

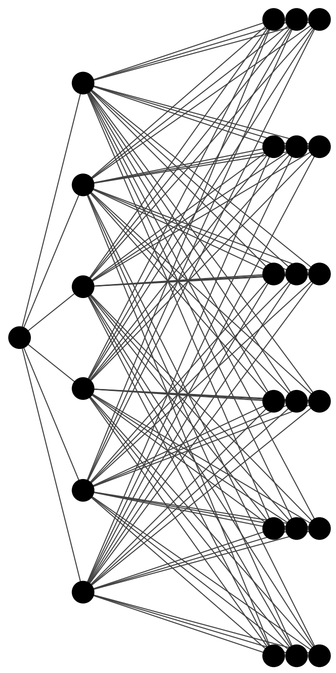

Figure 12: A rendering of $W(6,5,1)$.

Proof of Observation 28. First, we observe that

$$\begin{aligned}\mathrm{ex}^*\left(K_{n,n},\{K_{t,t},W(k,d,r)\text{-ind}\}\right) &\geq \mathrm{ex}\left(K_{n,n},\{K_{t,t},W(k,d,r)\text{-ind}\}\right)\\ &\geq \mathrm{ex}\left(K_n,\{K_{t,t},W(k,d,r)\text{-ind}\}\right)\\ &\geq \mathrm{ex}\left(n,\{K_{t,t},W(k,d,r)\}\right).\end{aligned}$$

Let us show $\mathrm{ex}\left(n,\{K_{t,t},W(k,d,r)\}\right)=\Omega\left(n^{2-\frac{1}{d}-\epsilon}\right)$ with the Deletion method, Lemma 6. We calculate

$$\gamma\left(K_{t,t}\right)=\frac{|K_{t,t}|-2}{\|K_{t,t}\|-1}=\frac{2(t-1)}{t^2-1}=\frac{2}{t+1}.$$

Hence, $\gamma\left(K_{t,t}\right)<\frac{1}{d}+\epsilon$ in case that $t\geq 2d-1$. On the other hand

$$\begin{aligned}\gamma\left(W(k,d,r)\right)=\frac{|W(k,d,r)|-2}{\|W(k,d,r)\|-1}&=\frac{(d-r-1)\binom{k}{d}+r+k-2}{d(d-r-1)\binom{k}{d}+rk-1}\\ &=\frac{1}{d}\left(1+\frac{r+k-2-\frac{rk-1}{d}}{(d-r-1)\binom{k}{d}+\frac{rk-1}{d}}\right)\leq\frac{1}{d}\left(1+\frac{2k}{\binom{k}{d}}\right)=\frac{1}{d}+o\left(1\right)\ \ (k\longrightarrow\infty).\ \square\end{aligned}$$

In order to provide some intuition about Conjecture 1 now we give two Observations.

Observation 29. For $d\in\mathbb{N}_{\geq 2}$ there is a sequence of $K_{d,d}$-free bipartite graphs $\left(H_s\right)_{s\in\mathbb{N}_{\geq 2}}$ and

$$\epsilon_s=O\left(s^{-\frac{d-1}{d+1}}\right)=o\left(1\right)\ \ (s\longrightarrow\infty)$$

such that for every $s\in\mathbb{N}$

$$\mathrm{ex}\left(n,H_s\right)=\Omega\left(n^{2-\epsilon_s}\right)\ \ (n\longrightarrow\infty).$$

Proof of Observation 29. An application of the Deletion method, Lemma 6, and Observation 8 shows that there is a sequence of $K_{d,d}$-free bipartite graphs $(H_s)_{s\in\mathbb{N}}$ on parts of size s such that

$$\|H_s\| = \Omega\left(s^{2-\gamma(K_{d,d})}\right) = \Omega\left(s^{2-\frac{2}{d+1}}\right) = \Omega\left(s^{\frac{2d}{d+1}}\right).$$

Now the Deletion method yields the Claim since $\gamma(H_s) = \frac{|H_s|-2}{\|H_s\|-1} = O\left(\frac{|H_s|}{\|H_s\|}\right) = O\left(s^{1-\frac{2d}{d+1}}\right) = O\left(s^{-\frac{d-1}{d+1}}\right)$. □

Observation 30. For $d \in \mathbb{N}_{\geq 2}$ there is a $K_{d,d}$-free bipartite graph H with maximal degree $d+1$ in one partition class such that $\operatorname{ex}(n, H) = \Omega\left(n^{2-\frac{1}{d}}\right)$ $(n \longrightarrow \infty)$.

Proof of Observation 30. **case** $d \geq 4$**.** Consider the $(d+1)$-uniform complete multipartite hypergraph $K := K_{d+1}^{(d+1)}(d-1)$ on $d+1$ many partition classes of size $d-1$ each. We want to define $H := \operatorname{Incidence}(K)$. Notice that $K_{d,d} \nsubseteq H$ since for all $A' \in \binom{V(K)}{d}$ either $\deg_H(A') = 0$ in case that A' contains two vertices from the same partition class or $\deg_H(A') = d-1$ otherwise. Observe that all vertices in H corresponding to the hyperedges in K have degree $d+1$. Now the Deletion method, Lemma 6, yields the Claim since for $d \geq 4$

$$\gamma(H) = \frac{|H|-2}{\|H\|-1} = \frac{(d+1)(d-1)+(d-1)^{d+1}-2}{(d+1)(d-1)^{d+1}-1} \leq \frac{(d+1)(d-1)+(d-1)^{d+1}}{(d+1)(d-1)^{d+1}} = \frac{1}{d+1} + \frac{1}{(d-1)^d} < \frac{1}{d}.$$

case $d = 3$**.** We want to make use of a $3-(8,4,1)$-design meaning a 4-uniform hypergraph on 8 vertices such that for every triple of vertices there is exactly one edge containing the triple. To prove existence we give a construction.

$$\begin{aligned}
A_i &:= \{\, a_{i,j} \mid j \in [4] \,\} \ \text{ for } i \in [2]\,. \\
E_{2,2} &:= \left\{ \{a_{1,x}, a_{1,y}, a_{2,x}, a_{2,y}\} \;\middle|\; \{x,y\} \in \binom{[4]}{2} \right\}. \\
E_{3,1} &:= \{\, \{a_{1,x}, a_{1,y}, a_{1,z}, a_{2,q}\} \mid q \in [4]\,,\ \{x,y,z\} = [4]\setminus\{q\} \,\}\,. \\
E_{1,3} &:= \{\, \{a_{1,q}, a_{2,x}, a_{2,y}, a_{2,z}\} \mid q \in [4]\,,\ \{x,y,z\} = [4]\setminus\{q\} \,\}\,. \\
\mathscr{H} &:= (A_1 \cupdot A_2,\ E_{2,2} \cupdot E_{3,1} \cupdot E_{1,3})\,.
\end{aligned}$$

Let us check that $\mathscr{H}$ is a $3-(8,4,1)$-design as claimed. For this purpose consider an arbitrary triple $\{x,y,z\} \in \binom{V(\mathscr{H})}{3}$. In case that $\{x,y,z\}$ is contained in either A_1 or A_2 there is exactly one edge containing it in $E_{3,1}$ or $E_{1,3}$ respectively. In the other case we may assume that $\{x,y\} \subset A_1$ and $z \in A_2$ so there are $\tilde{x} \neq \tilde{y}, \tilde{z} \in [4]$ such that $x = a_{1,\tilde{x}},\ y = a_{1,\tilde{y}},\ z = a_{2,\tilde{z}}$. In case $\tilde{z} \in \{\tilde{x}, \tilde{y}\}$ there is exactly one edge in $E_{2,2}$ containing $\{x,y,z\}$, otherwise there is exactly one edge in $E_{3,1}$ containing it.
Now consider the graph $H := \operatorname{Incidence}_2(\mathscr{H})$. Because $K_{3,2} \nsubseteq^* \operatorname{Incidence}(\mathscr{H})$ the blowup H can not contain a $K_{3,3}$. Furthermore, all vertices in the class of H corresponding to hyperedges of $\mathscr{H}$ have degree 4.
Now the Deletion method, Lemma 6, yields the Claim since

$$\gamma(H) = \frac{|H|-2}{\|H\|-1} = \frac{8+2\cdot 14-2}{2\cdot 14\cdot 4-1} = \frac{34}{111} < \frac{1}{3}.$$

case $d = 2$**.** Similarly to the previous case, we want to make use of a Steiner triple system, meaning a $2-(n,3,1)$ design. Such systems are known to exist for $n \in \mathbb{N}$ with $(n \bmod 6) \in \{1,3\}$, see [43]. Fix $k \in \mathbb{N}$ and let $\mathscr{H}$ be a $2-(6k+1,3,1)$ design. Since every hyperedge contains 3 pairs of vertices and each vertex pair is contained in exactly one hyperedge we obtain $\|\mathscr{H}\| = \frac{\binom{6k+1}{2}}{3} = (6k+1)k$. Let us define $H := \operatorname{Incidence}(\mathscr{H})$, which does not contain a $K_{2,2}$ and every vertex in one part has degree 3.
Again the Deletion method, Lemma 6, yields the Claim since for $k \longrightarrow \infty$ we have

$$\gamma(H) = \frac{|H|-2}{\|H\|-1} = \frac{(6k+1)(k+1)-2}{(6k+1)k\cdot 3-1} \longrightarrow \frac{1}{3}.$$ □

Simultaneously as we published Theorem 17 in [4] on arXiv Hunter, Milojevic, Sudakov and Tomon published the following result.

Theorem 18 (Hunter, Milojevic, Sudakov and Tomon [29])**.** Let $d, t \in \mathbb{N}$ and $H = (A \cupdot B,\ E)$ be a bipartite graph with $\max_{b \in B} \deg_A(b) \leq d$. Then

$$\mathrm{ex}\left(n, \{K_{t,t}, H\text{-ind}\}\right) \leq \left(4|A||B|\right)^{4|H|+10} n^{2-\frac{1}{d}}.$$

The Theorem is stronger than Theorem 16 in the sense that the host graph is complete instead of complete bipartite. Furthermore, H may contain $K_{d,d}$ as a subgraph. However, in their setting the little-o-bound of Theorem 16 can not hold.
Our second main result is a counting result that implies the same asymptotics as Theorem 18 in case that $s = d + 1$.

Theorem 19. Let $d, r \in \mathbb{N}$ with $r \leq d$ and let $H = (A \cupdot B,\ F)$ be a bipartite graph with r many complete vertices $\tilde{A} := \{\, a \in A \mid N_B(a) = B \,\}$ and $\max_{a \in A \setminus \tilde{A}} \deg_B(a) \leq d$. Then there are constants $C, c > 0$ such that for any graph $G \in \mathrm{Free}\left(K_{d+1,d+1}\right)$ with $p := \frac{2\|G\|}{|G|^2} \geq C|G|^{-\frac{1}{d}}$

$$|\mathrm{Isom}_{\mathrm{ind}}(H,\ G)| \geq c \cdot p^{\|H\|} |G|^{|H|} \left(p^{d+1}|G|\right)^{\frac{\tau(H)}{d+1-r}},$$

where $\tau(H) := \sum\limits_{a \in A \setminus \tilde{A}} d - \deg_B(a)$.

Since the lower bound for the number of induced graph isomorphisms is positive the Theorem always guarantees the existence of at least one induced copy of H in G whenever $\|G\| \geq \frac{C}{2}|G|^{2-\frac{1}{d}}$. In case that $t = d+1$ it is stronger than Theorem 18 in the sense that H may have complete vertices, and we are counting induced isomorphisms. However, the constant term C in the result is worse than the constant $\left(4|A||B|\right)^{4|H|+10}$ in Theorem 18.
Furthermore, the lower bound on the number of induced isomorphisms reminds of Sidorenkos Conjecture. This is no coincidence since the proof is inspired by a result in [12] which deals about Sidorenkos Conjecture for bipartite graphs with complete vertices. We also had access to the simplified version of the proof in [13]. Sidorenkos Conjecture is a statement about so-called graphons. For simplicity, we want to only state the version for graphs.

Conjecture 2 (Sidorenkos Conjecture [42])**.** For any bipartite graph H and any graph G with $p := \frac{2\|G\|}{|G|^2}$ one finds that

$$|\mathrm{Hom}(H,\ G)| \geq p^{\|H\|}|G|^{|H|}.$$

However, currently it is unclear, if the bound of Theorem 19 is sharp.
A simplification of the proof of Theorem 19 yields the following Theorem. Albeit we are going to use it in the proof of our main result, we omit a self-reliant proof.

Theorem 20. Let $d, r, t \in \mathbb{N}$ with $t > d \geq r$ and let $H = (A \cupdot B,\ F)$ be a bipartite graph with r many complete vertices $\tilde{A} := \{\, a \in A \mid N_B(a) = B \,\}$ and $\max_{a \in A \setminus \tilde{A}} \deg_B(a) \leq d$. Then there are constants $C, c > 0$ such that for any graph $G \in \mathrm{Free}\left(K_{t,t}\right)$ with $p := \frac{2\|G\|}{|G|^2} \geq C|G|^{-\frac{1}{d}}$

$$|\mathrm{Isom}_{\mathrm{biind}}(H,\ G)| \geq c \cdot p^{\|H\|}|G|^{|H|}.$$

4.2 Proof of the main result

To draw the connection from Theorem 16 to Theorem 17 at first we present the following two Observations. Before this let us introduce the notion of strong neighborhood that we find illustrating throughout the proofs.

Definition 70 (Strong neighborhood)**.** Let G be a graph and $A \subseteq V(G)$ as well as $U \subseteq A$. Then we define

$$N_{\mathrm{strong}}(U,\ A) := \{\, v \in V(G) \mid N_A(v) = U \,\} = N_G(U) \setminus \left(\bigcup_{a \in A \setminus U} N_G(a) \right).$$

Observation 31. Let $d, s \in \mathbb{N}$ and $H = (A \cupdot B,\ E)$ be a bipartite graph with $\max_{b\in B}\deg_A(b) \leq d$ as well as $K_{d,s+1} \not\subseteq^* H$. Then for sufficiently large $k \in \mathbb{N}:\ H \underset{ind}{\subseteq^*} H(k,d,s)$.

Proof of Observation 31. Consider $\mathscr{H} := (A,\ \{\ N_A(b) \mid b \in B\ \})$. For $i \in [d]$ let us define $\mathscr{E}^{(i)} := E(\mathscr{H}) \cap \binom{A}{i}$ as well as $m_i := \max_{e\in\mathscr{E}^{(i)}} |N_{\text{strong}}(e,\ A)|$. Since $K_{d,s+1} \not\subseteq^* H$ we observe that $m_d \leq s$.
Now we define q to be the minimal integer such that

$$(*)\ \forall i \in [d]:\ s\binom{q}{d-i} \geq m_i.$$

Let A' be a set of q many vertices disjoint from A. Let us define an extended hypergraph.

$$\mathscr{H}' := \left(A \cupdot A',\ \mathscr{E}^{(d)} \cup \bigcup_{i\in[d-1]} \left\{ e \cup e' \,\middle|\, e \in \mathscr{E}^{(i)},\ e' \in \binom{A'}{d-i} \right\}\right).$$

We notice that since $\mathscr{H}'$ is d-uniform we have that $I := \text{Incidence}_s(\mathscr{H}') \underset{ind}{\subseteq^*} H(|A| + q, d, s)$.
Thus, it suffices to check that $H \underset{ind}{\subseteq^*} I$. For this we can simply embed the vertices of A in H as the vertices A in I. For $b \in B$ let us denote the vertices in B that have the same neighborhood in A as b by $B(b) := N_{\text{strong}}(N_A(b),\ A)$. We embed $B(b)$ in the blowup vertices corresponding to $\left\{ N_A(b) \cup e' \,\middle|\, e' \in \binom{A'}{d-\deg_A(b)} \right\}$.
This is possible since

$$s \cdot \left| \left\{ N_A(b) \cup e' \,\middle|\, e' \in \binom{A'}{d-\deg_A(b)} \right\} \right| = s\binom{q}{d-\deg_A(b)} \geq m_{\deg_A(b)} \geq |B(b)|.$$

We observe that $q := \left\lceil \frac{|B|}{s} \right\rceil$ suffices the condition $(*)$ since $\forall i \in [d-1]:\ s\binom{q}{d-i} \geq s\left\lceil \frac{|B|}{s} \right\rceil \geq |B| \geq m_i$. In most cases one can choose much smaller q, however in case that $H = K_{d-1,r}$ for some $r \in \mathbb{N}$ the choice $q = \left\lceil \frac{r}{s} \right\rceil$ is optimal. This completes the proof of Observation 31. □

Observation 32. Let $r, d \in \mathbb{N}_0$ with $d \geq r+2$ and $H = (A \cupdot B,\ E)$ be a bipartite graph such that there is $X \in \binom{B}{r}$ with $N_A(X) = A$ as well as $\max_{b\in B\setminus X}\deg_A(b) \leq d$ and $K_{d,d} \not\subseteq H$. Then for large enough $k \in \mathbb{N}$ one finds $H \underset{ind}{\subseteq^*} W(k,d,r)$.

Proof of Observation 32. First observe that $H - B$ fulfills all conditions stated in Observation 31 with $s = d-r-1$ since $K_{d,d-r} \not\subseteq^* H - X$. Thus, we find $k \in \mathbb{N}$ such that $H - X \underset{ind}{\subseteq^*} H(k,d,d-r-1)$ and we conclude that $H \underset{ind}{\subseteq^*} W(k,d,r)$.
This completes the proof of Observation 32. □

In the proof of our main Theorem we are going to find induced hedgehogs using the following Observation.

Observation 33. $\forall s, d \in \mathbb{N}\ \exists k \in \mathbb{N}$ such that for any bipartite graph $H = (A \cupdot B,\ F)$ with $|A| = k$ one has

$$\left(\forall S \in \binom{A}{d},\ a \in A \setminus S:\ \deg_B(S \cup \{a\}) < \frac{\deg_B(S) - s}{k-d}\right) \implies H(k,d,s) \underset{ind}{\subseteq^*} G.$$

Proof of Observation 33. We observe that for any $S \in \binom{A}{d}$ we have that

$$|N_{\text{strong}}(S,\ A)| \geq \deg_B(S) - \sum_{a\in A\setminus S} \deg_B(S \cup \{a\}) \geq \deg_B(S) - (\deg_B(S) - s) = s.$$

This already completes the proof of Observation 33. □

The main tool for proving Theorem 17 is that the existence of a dense bipartite subgraph on specifically imbalanced partition classes implies the existence of the forbidden structures. In the proof we merge the ideas of [44] and [32].

Theorem 21. $\forall k, d, t, s \in \mathbb{N}$ with $k \geq d$ $\forall \eta > 0$ $\exists \xi, \kappa > 0,$ $N \in \mathbb{N}$ such that $\forall n \geq N$ and for all bipartite graphs $G = (A \cup B,\ E)$ fulfilling the following conditions

(a) $|A| \leq \xi n^{\frac{s}{d}}$ and $|B| = n$.

(b) $\|G\| \geq \xi \eta n^{\frac{d+s-1}{d}}$.

(c) $\max_{b \in B} \deg_A(b) \leq \kappa |A|$.

one finds $K_{t,t}$ as a subgraph or a (k, d, s)-hedgehog as an induced subgraph in G.

We remark that the case $t = 1$ is trivial.

Proof of Theorem 21. As a help structure let us define a d-uniform hypergraph on the vertices A.

$$\mathscr{E} := \left\{ e \in \binom{A}{d} \,\middle|\, \deg_B(e) \geq s \right\}, \quad \mathscr{H} := (A,\ \mathscr{E}).$$

Fix $q := k(4kt)^t$. We want to distinguish edges where that have an upper bound on the size of their common neighborhood and edges where we do not have such. For this reason let us define the following coloring:

$$\phi : \mathscr{E} \to \{\text{red}, \text{blue}\}$$
$$e \mapsto \begin{cases} \text{blue} & \deg_B(e) < q \\ \text{red} & q \leq \deg_B(e) \end{cases}$$

Furthermore, we introduce notation for (mono chromatically colored) d-uniform cliques on $\mathscr{H}$ of size q. We adapt the notation of Definition 35.

$$\mathscr{K} := \mathscr{K}_q^{(d)}(\mathscr{H}) = \left\{ K \in \binom{A}{q} \,\middle|\, \binom{K}{d} \subseteq \mathscr{E} \right\}.$$
$$\mathscr{K}(\text{blue}) := \left\{ K \in \mathscr{K} \,\middle|\, \forall e \in \binom{K}{d} :\ \deg_B(e) < q \right\}.$$
$$\mathscr{K}(\text{red}) := \left\{ K \in \mathscr{K} \,\middle|\, \forall e \in \binom{K}{d} :\ \deg_B(e) \geq q \right\}.$$

The idea to prove the Theorem is to arrive at a contradiction when counting the number of q-cliques $|\mathscr{K}|$. To obtain a lower bound on $|\mathscr{K}|$ we will use the high edge density in G and the Hypergraph Removal lemma. To obtain an upper bound on $|\mathscr{K}|$ we will use that in a sufficiently large hyperclique the edges can not have a large common neighborhood in B - there are no red cliques after all.

Lemma 21. $\forall d, s, k, t \in \mathbb{N}$ with $k \geq d \geq 2$ and any bipartite graph $F = (Q \cup R,\ E)$ fulfilling $K_{t,t} \not\subseteq F,\ |Q| = k(4kt)^t$ one has

$$\min_{e \in \binom{Q}{d}} \deg_R(e) \geq q \implies H(k, d, s) \underset{ind}{\subseteq} F.$$

Proof of Lemma 21. Let us call a k-set in Q a *good set* in case that it fulfills a condition similar to the set A in Observation 33.

$$\mathscr{A}_{\text{good}} := \left\{ A \in \binom{Q}{k} \,\middle|\, \forall S \in \binom{A}{d},\ a \in A \setminus S :\ \deg_R(S \cup \{a\}) < \frac{\deg_R(S) - s}{k - d} \right\}.$$

Note that Observation 33 states, that in case $\mathscr{A}_{\text{good}} \neq \emptyset$ one finds a (k, r, s)-hedgehog as an induced subgraph in G. For $S \in \binom{Q}{d}$ define a vertex in $Q \setminus S$ to be S-*bad* in case that it has a large common neighborhood with S.

$$V_{\text{bad}}(S) := \left\{ v \in Q \setminus S \,\middle|\, \deg_R(S \cup \{v\}) \geq \frac{\deg_R(S) - s}{k - d} \right\}.$$

Let us bound the number of bad vertices by double counting the edges between the bad vertices and the neighborhood of S. On the one hand we have that

$$\|V_{\text{bad}}(S),\ N_R(S)\| \geq |V_{\text{bad}}| \frac{\deg_R(S) - s}{k-d} \geq |V_{\text{bad}}| \frac{\deg_R(S) - d}{k} \geq |V_{\text{bad}}| \frac{\deg_R(S)}{2k}$$

where in the last inequality we used that $\deg_R(S) \geq q \geq 2d$. On the other hand the bound on the Zarankievicz function by the Kővári Sós Túran Theorem, Lemma 2, yields that

$$\|V_{\text{bad}}(S),\ N_R(S)\| \leq t\left(\deg_R(S)\,|V_{\text{bad}}(S)|^{1-\frac{1}{t}} + |V_{\text{bad}}(S)|\right).$$

Putting both bounds together we obtain

$$|V_{\text{bad}}| \frac{\deg_R(S)}{4k} \leq |V_{\text{bad}}| \left(\frac{\deg_R(S)}{2k} - t\right) \leq t\deg_R(S)\,|V_{\text{bad}}|^{1-\frac{1}{t}}$$

where in the first inequality we used that $\frac{\deg_R(S)}{2k} \geq \frac{q}{2k} \geq 2t$. We conclude

$$|V_{\text{bad}}| \leq (4kt)^t.$$

Define $\mathscr{A}_{\text{bad}} := \binom{Q}{k} \setminus \mathscr{A}_{\text{good}}$. To proof Claim 21 it suffices to show that $|\mathscr{A}_{\text{bad}}| < \binom{q}{k}$.
Using Union Bound we obtain

$$\begin{aligned}
|\mathscr{A}_{\text{bad}}| &= \sum_{A \in \binom{Q}{k}} \mathbb{1}\left\{\exists S \in \binom{A}{d},\ a \in A \setminus S:\ \deg_R(S \cup \{a\}) \geq \frac{\deg_R(S) - s}{k-d}\right\} \\
&\leq \sum_{A \in \binom{Q}{k}} \left(\sum_{S \in \binom{A}{d}} \left(\sum_{a \in A \setminus S} \mathbb{1}\left\{\deg_R(S \cup \{a\}) \geq \frac{\deg_R(S) - s}{k-d}\right\}\right)\right) \\
&= \sum_{S \in \binom{Q}{d}} \left(\sum_{a \in Q \setminus S} \left(\sum_{\substack{A \in \binom{Q}{k} \\ S \cup \{a\} \subseteq A}} \mathbb{1}\{a \in V_{\text{bad}}(S)\}\right)\right) \\
&= \binom{q-d-1}{k-d-1} \sum_{S \in \binom{Q}{d}} \left(\sum_{a \in Q \setminus S} \mathbb{1}\{a \in V_{\text{bad}}(S)\}\right) \\
&\leq \binom{q-d-1}{k-d-1} \cdot \binom{q}{d} \cdot (4kt)^t = \frac{k-d}{q-d}\binom{q-d}{k-d} \cdot \binom{q}{d} \cdot (4kt)^t = \frac{k-d}{q-d}(4kt)^t \binom{q}{k} < \binom{q}{k}
\end{aligned}$$

where in the last inequality we used that our assumption $q \geq k(4kt)^t$ implies $q - d > (k-d)(4kt)^t$ which implies that $\frac{k-d}{q-d}(4kt)^t < 1$. This completes the proof of Lemma 21. □

By Lemma 21 and our choice of q we deduce that $\mathscr{K} = \mathscr{K}(\text{blue})$. Using this and the Hypergraph Removal lemma we will show

Claim 9. There is a constant $C > 0$ that is independent of $|A|$ and κ such that $|\mathscr{K}| \geq C\binom{|A|}{q}$.

Proof of Claim 9. Ramseys Theorem 1 yields $\Gamma := R^{(d)}(q,q) \in \mathbb{N}$ such that every $\{\text{red}, \text{blue}\}$-coloring of $\binom{[\Gamma]}{d}$ contains a monochromatic hyperclique of size q. Let $C \in \binom{B}{s}$. Observe that the hypergraph $\mathscr{H}$ restricted to the common neighborhood $N_A(C)$ is a clique. By considering disjoint blocks of size Γ in the common neighborhood we find $m(C) := \left\lfloor \frac{\deg_A(C)}{\Gamma} \right\rfloor$ many monochromatic disjoint copies $(U_j(C))_{j \in [m(C)]} \subseteq N_A(C)$ of q-cliques which must be blue by Lemma 21. For our convenience we define

$$\mathscr{C}(C) := \{\, U_j(C) \mid j \in [m(C)] \,\}\,.$$
$$\mathscr{C} := \bigcup_{C \in \binom{B}{s}} \mathscr{C}(C).$$

Since any element in $\mathscr{C}$ is a blue clique it has less than q common neighbors in B. Thus, any element in $\mathscr{C}$ is contained in the neighborhood of at most $\binom{q-1}{s}$ many s-sets of B. We conclude that

$$\sum_{C \in \binom{B}{s}} |\mathscr{C}(C)| \le \binom{q-1}{s} |\mathscr{C}|.$$

By the Definition of Γ we know

$$\sum_{C \in \binom{B}{s}} |\mathscr{C}(C)| \ge \sum_{C \in \binom{B}{s}} \left(\frac{\deg_A(C)}{\Gamma} - 1 \right).$$

Using a double counting argument we observe that

$$\sum_{C \in \binom{B}{s}} \deg_A(C) = \sum_{a \in A} \binom{\deg_B(a)}{s}.$$

Observe that $x \mapsto \binom{x}{s} = \mathbb{1}\{x \ge s\} \prod_{0 \le j < s} \frac{x-j}{s-j}$ is a convex function on $\mathbb{R}$. Using this and Jensens inequality, Lemma 1, let us calculate

$$\sum_{a \in A} \binom{\deg_B(a)}{s} = |A| \sum_{a \in A} \frac{\binom{\deg_B(a)}{s}}{|A|} \ge |A| \binom{\frac{2\|G\|}{|A|}}{s} \ge |A| \left(\frac{\|G\|}{s|A|} \right)^s = \frac{1}{s^s} |A|^{1-s} \|G\|^s.$$

We conclude

$$\begin{aligned}
\binom{q-1}{s} |\mathscr{C}| &\ge \sum_{C \in \binom{B}{s}} \left(\frac{\deg_A(C)}{\Gamma} - 1 \right) \\
&\ge \frac{1}{\Gamma} \sum_{C \in \binom{B}{s}} \deg_A(C) - \binom{|B|}{s} \\
&\ge \frac{1}{s^s \Gamma} |A|^{1-s} \|G\|^s - n^s \\
&\ge \frac{1}{s^s \Gamma} \xi^{1-s} n^{(1-s)\frac{s}{d}} (\xi\eta)^s n^{s\frac{d+s-1}{d}} - n^s \\
&= \frac{\eta^{s^s}}{s\Gamma} \xi n^s - n^s.
\end{aligned}$$

Thus, in case that we choose $\xi > \frac{s^s \Gamma}{\eta^s}$ we obtain $|\mathscr{C}| = \Omega(n^s) = \Omega\left(\left(n^{\frac{s}{d}} \right)^d \right) = \Omega\left(|A|^d \right)$.
Observe that any (blue) hyperedge e from any q-clique in $\mathscr{C}$ can lie in at most $\binom{q-1}{s}$ many q-cliques of $\mathscr{C}$. This is true since for any s-set C in the neighborhood of e in B there is at most one clique in $\mathscr{K}(C)$ that contains e. Thus, to delete all q-cliques from $\mathscr{H}$ one needs to delete at least $\frac{1}{\binom{q-1}{s}} |\mathscr{C}| = \Omega\left(|A|^d \right)$ many hyperedges. Thus, the Hypergraph Removal lemma, Theorem 5, yields the proof of Claim 9. □

Claim 10. For any $K \in \mathscr{K}$ we can fix distinct $T_1(K), T_2(K) \in \binom{K}{d}$ such that $N_B(T_1(K) \cup T_2(K)) \neq \emptyset$.

Proof of Claim 10. Assume for a contradiction that $(N_B(e))_{e \in \binom{K}{d}}$ are pairwise disjoint. Then we would find a (q, d, s)-hedgehog in G with body in A. Furthermore, this hedgehog is induced. Assume for a contradiction that the hedgehog would not be induced, meaning that there is $b \in B$ such that $\deg_K(b) \ge d + 1$. However, in this case b would lie in the common neighborhood of more than one hyperedge $e \in \binom{K}{d}$, a contradiction. □

As already laid out we want to arrive at a contradiction by showing $|\mathscr{K}| = o\left(\binom{|A|}{q}\right)$. Using Claim 10 for any $K \in \mathscr{K}$ we may fix $\tilde{a}(K) \in T_1(K) \setminus T_2(K)$. We can count

$$|\mathscr{K}| = \sum_{K' \in \binom{A}{q-1}} \left(\sum_{a \in A \setminus K'} \mathbb{1}\left\{a = \tilde{a}(K' \cup \{a\})\right\} \right).$$

However, for a fixed $K' \in \binom{A}{q-1}$ we can upper bound the count of $a \in A \setminus K'$ such that a is exactly the fixed vertex in the q-clique $K' \cup \{a\}$ by the fact that in this case a has to lie in the neighborhood of any vertex in the common neighborhood of $T_2(K' \cup \{a\})$. However, the number of possible choices for $T_2(K)$ inside K' is at most $\binom{q-1}{k-1}$. Furthermore, for a fixed $K_2(K' \cup \{a\})$ there are at most $q-1$ common neighbors in B since all cliques are blue. Using requirement (c) for G we bound

$$\begin{aligned}|\mathscr{K}| &= \sum_{K' \in \binom{A}{q-1}} \left(\sum_{a \in A \setminus K'} \mathbb{1}\left\{a = a(K' \cup \{a\})\right\} \right) \\ &= \sum_{K' \in \binom{A}{q-1}} \left(\sum_{T \in \binom{K'}{d}} \left(\sum_{b \in N_B(T)} \deg_A(b) \right) \right) \\ &\leq \binom{|A|}{q-1} \binom{q-1}{k-1} (q-1)\left(\kappa|A|\right).\end{aligned}$$

However, this is a contradiction to Claim 9 if we choose κ small enough. We close the proof of Theorem 21 by remarking that the constant C in Claim 9 comes from the Hypergraph Removal lemma, so we have little control on the choice of κ.
This completes the proof of Theorem 31. □

Now we are able to establish a proof of Theorem 17.

Proof of Theorem 17. Let us assume for a contradiction that there are $r, d, k, t \in \mathbb{N}_0$ with $k \geq d \geq r+2$ as well as some constant $C > 0$ such that for any $\tilde{n}_0 \in \mathbb{N}$ there is $\tilde{n} \in \mathbb{N}$ with $\tilde{n} \geq \tilde{n}_0$ and there is a bipartite graph G on partite sets of size $\tilde{n}$ each with $\|G\| \geq C\tilde{n}^{2-\frac{1}{d}}$ that neither contains $K_{t,t}$ as a subgraph nor contains an induced copy of $W(k, d, r)$.
By an application of the Reduction lemma 11 we may assume that there are constants $C', K > 0$ such that for any $n_0 \in \mathbb{N}$ there is some $n \geq n_0$ and some K-almost regular bipartite graph $G' = (A \cup B,\ E)$ on n vertices and $C'n^{2-\frac{1}{d}}$ edges that neither contains $K_{t,t}$ as a subgraph nor contains an induced copy of $W(k, d, r)$.
Observe that $|A| \cdot \delta(G) \leq \|G\| \leq |B| \cdot \Delta(G)$. Thus, $|A| \leq \frac{\Delta(G)}{\delta(G)}|B| \leq K|B|$. By an analogous argument we also know that $|B| \leq K|A|$. We deduce

$$|A|\left(1 + \frac{1}{K}\right) \leq |A| + |B| = n \leq |A|(1+K).$$

Fix $\eta := \frac{C'}{2K}$ and $s := d - r - 1$ and let $\xi, \kappa > 0$ as well as $N \in \mathbb{N}$ be given by Theorem 21. In the sequel we are going to assume n_0 to be large enough.
We want to find $X \in \binom{B}{r}$ and $B' \subseteq B \setminus X$ as well as $A' \subseteq N_A(X)$ (where we define $N_A(\emptyset) := A$) such that

(a) $|A'| \leq \xi|B'|^{\frac{s}{d}},\ |B'| \geq N$.

(b) $\|A', B'\| \geq \xi\eta|B'|^{\frac{d+s-1}{d}}$.

(c) $\max_{b \in B'} \deg_{A'}(b) \leq \kappa|A'|$.

In case we have found such sets A', B' and X let us define $G' := G[A',\ B']$. Notice $K_{t,t} \nsubseteq G'$ and G' fulfills the requirements of Theorem 21. Thus, we find an induced copy of a (k, r, s)-hedgehog in G' and together with X we find an induced copy of $W(k, d, r)$ in G, a contradiction.

We can find $X \in \binom{B}{r}$ with a large common neighborhood by an averaging argument. Notice again that the case $r = 0$ is trivial since $N_A(\emptyset) = A$. Using the convexity of $x \mapsto \binom{x}{r}$ and Jensens inequality, Lemma 1, we calculate

$$\sum_{X\in\binom{B}{r}} \deg_A(X) = |A| \sum_{a\in A} \frac{\binom{\deg_B(a)}{r}}{|A|} \geq |A| \binom{\frac{\|G\|}{|A|}}{r} \geq |A| \left(\frac{\|G\|}{r|A|}\right)^r \geq \frac{1}{r^r}\left(\frac{K+1}{Kn}\right)^{r-1}\left(C' n^{2-\frac{1}{d}}\right)^r \geq \left(\frac{C'}{Kr}\right)^r n^{r+\frac{d-r}{d}},$$

where we used the inequality $|A| \leq \frac{K}{K+1} n$.
Now the pigeon whole principle yields $X \in \binom{B}{r}$ with $\deg_A(X) = \Omega\left(n^{\frac{d-r}{d}}\right) = \omega\left(n^{\frac{s}{d}}\right)$. Hence, if n is large enough we can choose $A' \in N_A(X)$ such that $|A'| = \left\lfloor \frac{\xi}{2} \cdot |B|^{\frac{s}{d}} \right\rfloor$.
Now we want to delete vertices with in B with a too high degree towards A'. For this purpose fix $\epsilon \in \left(0, \frac{1}{t}\right)$. Let us define the vertices which we want to delete by

$$V_{\text{delete}} := \left\{ b \in B \;\middle|\; \deg_{A'}(b) \geq |A'|^{1-\epsilon} \right\}.$$

Let us bound the number of V_{delete} by double counting the edges between V_{delete} and A'. On the one hand we have

$$\|A', V_{\text{delete}}\| \geq |V_{\text{delete}}||A'|^{1-\epsilon}.$$

On the other hand the bound on the Zarankievicz function by the Kővári, Sós, Turán theorem, Lemma 2, yields

$$\|A', V_{\text{delete}}\| \leq t\left(|A'|^{\frac{t-1}{t}}|V_{\text{delete}}| + |A'|\right).$$

Putting both bounds together we arrive at

$$|V_{\text{delete}}|\left(1 - t|A'|^{\epsilon-\frac{1}{t}}\right) \leq t|A'|^{\epsilon}.$$

Since $\epsilon < \frac{1}{t}$ for large n we have that $t|A'|^{\epsilon-\frac{1}{t}} \leq \frac{1}{2}$. In this case

$$|V_{\text{delete}}| \leq 2t|A'|^{\epsilon}.$$

Let us set $B' := B \setminus V_{\text{delete}}$. Observe that $X \subseteq V_{\text{delete}}$ so $B' \cap X = \emptyset$. It is left to check for the properties (a), (b) and (c).
Ad (a). In case that n is large enough we have that $|B'| \geq |B| - 2t|A'|^{\epsilon} \geq \frac{|B|}{2} \geq N$. We conclude

$$|A'| = \left\lfloor \frac{\xi}{2}|B|^{\frac{s}{d}} \right\rfloor \leq \xi\left(\frac{|B|}{2}\right)^{\frac{s}{d}} \leq \xi|B'|^{\frac{s}{d}}.$$

Ad (b). Observe that $\delta(G) \geq \frac{\text{avdeg}(G)}{K} \geq \frac{2C'}{K} n^{1-\frac{1}{d}}$. Thus, for large n we have that $\delta(G) - |V_{\text{delete}}| \geq \frac{C'}{K} n^{1-\frac{1}{d}}$.

$$\|A',\ B'\| \geq |A'| \cdot \delta(G') \geq |A'|\left(\delta(G) - |V_{\text{delete}}|\right) \geq \frac{\xi}{2} n^{\frac{s}{d}} \frac{C'}{K} n^{1-\frac{1}{d}} = \eta\xi n^{\frac{d+s-1}{d}}.$$

Ad (c). We have $\max_{b\in B'} \leq |A'|^{\epsilon} \leq \kappa|A'|$ in case that n is large enough.
This completes the proof of Theorem 17. □

Now we are well-prepared to prove our main result.

Proof of Theorem 16. Let $d, t \in \mathbb{N},\ d \geq 2$ and let $H = (A \cup B,\ E)$ be a $K_{d,d}$-free bipartite graph that fulfills the degree conditions of Theorem 16 with parameter d. Let $\tilde{A} := \{ a \in A \mid \deg_B(a) > d \}$ as well as $r := |\tilde{A}|$.
case $r \leq d-2$**.** In this case Observation 32 yields $k \in \mathbb{N}$ with $k \geq d$ such that $H \underset{ind}{\subseteq^*} W(k, d, r)$. Now it is easy

to see that

$$\mathrm{ex}^*\left(K_{n,n},\{K_{t,t},H\text{-ind}\}\right)\le \mathrm{ex}^*\left(K_{n,n},\{K_{t,t},W(k,d,r)\text{-ind}\}\right).$$

Thus, the first Claim of Theorem 16 follows directly from Theorem 17.
case $r=d-1$**.** In this case all vertices in $A\setminus\tilde{A}$ have degree at most $d-1$. Notice that H fulfills the requirements for Theorem 20 with parameter $\tilde{d}=d-1,\ \tilde{r}=r$. Thus, Theorem 20 and a contradiction argument yields a constant $C>0$ such that for any $n\in\mathbb{N}$ and any graph $G\in\mathrm{Free}\left(n,\{K_{t,t},H\text{-biind}\}\right)$ one has $\|G\|<\frac{C}{2}n^{2-\frac{1}{d}}$. Since we may have chosen $G\subseteq K_{\lceil\frac{n}{2}\rceil,\lfloor\frac{n}{2}\rfloor}$ the first Claim of Theorem 16 follows by

$$\|G\|=O\left(n^{2-\frac{1}{d-1}}\right)=o\left(n^{2-\frac{1}{d}}\right)\ (n\longrightarrow\infty).$$

Regarding the second Claim of Theorem 16 let us assume for a contradiction that there is a constant $c>0$ such that for any $n_0\in\mathbb{N}$ there is $n\in\mathbb{N}$ with $n\ge n_0$ and a graph $G\in\mathrm{Free}\left(n,\{K_{t,t},H\text{-biind}\}\right)$ such that $\|G\|\ge cn^{2-\frac{1}{d}}$. Now a standard result yields a partition (A,B) of $[n]$ such that $\|G[A,B]\|\ge\frac{\|G\|}{2}\ge\frac{c}{2}n^{2-\frac{1}{d}}$. It follows that $\min\{|A|,|B|\}\ge\frac{c}{2}n^{1-\frac{1}{d}}$. We may assume that $|A|\le|B|$.
Let $\tilde{n}:=|A|$. Let us pick a subset $\tilde{B}\in\binom{B}{\tilde{n}}$ of the vertices in B of maximal degree, formally $\min_{\tilde{b}\in\tilde{B}}\deg_A\left(\tilde{b}\right)\ge\max_{b\in B\setminus\tilde{B}}\deg_A(b)$. We observe

$$\|A,\tilde{B}\|\cdot\frac{|B|}{|\tilde{B}|}\ge\|A,B\|.$$

By our choice of (A,B) and our assumption on $\|G\|$ as well as the fact that $|B|,\tilde{n}\le n$ we deduce

$$\|A,\tilde{B}\|\cdot\frac{|B|}{\tilde{n}}\ge\frac{c}{2}|B|\tilde{n}^{1-\frac{1}{d}}.$$

We conclude

$$\|G\left[A,\tilde{B}\right]\|\ge\frac{c}{2}\tilde{n}^{2-\frac{1}{d}}.$$

Since $H\underset{ind}{\not\subseteq}G\left[A,\tilde{B}\right]$ it follows that

$$\mathrm{ex}\left(K_{\tilde{n},\tilde{n}},\{K_{t,t},H\text{-ind}\}\right)\ge\|G\left[A,\tilde{B}\right]\|\ge\frac{c}{2}\tilde{n}^{2-\frac{1}{d}}.$$

However, since $\tilde{n}$ can become arbitrary large, this is a contradiction to the first Claim. This completes the proof of Theorem 16. □

4.3 Counting induced hedgehogs

In the proof of Theorem 19 we are going to represent vertex mappings from a graph H to a graph G as sequences $S:=(v_x)_{x\in V(H)}\in V(G)^{V(H)}$. We call S an *embedding* in case that it is injective. Furthermore, for $a\in V(H)$ we introduce the short notation

$$S\big|_a:=(v_b)_{b\in N_H(a)}.$$

For the related notion of restricted sequences we remind of Definition 17. For the notion of induced graph isomorphism take a look at Definition 6. For vertex sets U,I with $U\subseteq I$ let us say that two embeddings $S,\tilde{S}\in V(G)^I$ *agree on* U in case that $S\big|_U=\tilde{S}\big|_U$.

Furthermore, for $k \in \mathbb{N}$ and a sequence $W := (v_j)_{j\in[k]} \subseteq V(G)$ we are going to use the abbreviated notation

$$N_G(W) := N_G(\{ v_j \mid j \in [k] \}).$$
$$\deg_G(W) := \deg_G(\{ v_j \mid j \in [k] \}).$$
$$|W| := |\{ v_j \mid j \in [k] \}|.$$

We call S *independent* in case that $\{ v_j \mid j \in [k] \}$ is independent in G.

Proof of Theorem 19. Let us start by fixing the constants that we use in the proof.

$$s := d+1, \quad n_A := |A|, \quad n_B := |B|, \quad n_{B'} := n_B + \tau(G).$$

Furthermore, let us set

$$\alpha := 2^{-r}\left(\sqrt{s}\right)^{-r^2}, \quad \beta := 2^{-n_{B'}}\left(\sqrt{s}\right)^{-n_{B'}^2}, \quad \xi_2 := \frac{1}{4n_{B'}}, \quad \xi_3 := \frac{\beta}{4n_A}, \quad \xi_1 := \left(\frac{\alpha\xi_3}{2^{d+2r+2}}\right)^{\frac{1}{r}}, \quad q := \frac{1}{2}p^r|G|.$$

Let $\tilde{C} > 0$ be the constant obtained from the Packing lemma, Theorem 10, with parameter $\tilde{k} = (d+1)^{(d+1)}$ and $\tilde{d} = d+1-r$ as well as $\tilde{c} = e + (d+1)^{(d+1)}$. We may assume that $\tilde{C} \geq 2$. With this let us define

$$c := \frac{\alpha\beta^2}{2^{n_{B'}+r+6}}\left(\frac{\xi_1}{16\tilde{C}n_A s^{n_B-1}}\right)^{n_A-r}.$$

Furthermore, we fix $C > 0$. We decided not to make C explicit, but to simply assume that C is large enough, mentioning every instance of applying this assumption directly in the proof.

Let us start by sketching the plan for finding many induced isomorphisms from H to G. At first, we introduce a help graph H' where we fill the neighborhoods of all non-complete vertices in A up to size d.

$$B' := B \mathbin{\dot\cup} \bigcup_{a\in A\setminus\tilde{A}} \{ b_{a,j} \mid j \in [d - \deg_B(a)] \}, \text{ where } (b_{a,j})_{a\in A\setminus\tilde{A}, j\in[d-\deg_B(a)]} \text{ are pairwise distinct.}$$
$$H' := \left(A \mathbin{\dot\cup} B',\ E(H) \mathbin{\dot\cup} \left\{ \{a, b_{a,j}\} \;\middle|\; a \in A\setminus\tilde{A},\ j \in [d-\deg_B(a)] \right\}\right).$$

We remark that

$$|B'| = n_{B'}.$$

For any independent embedding $Z \in \mathscr{I}_r(G)$, see Definition 37, that meets some later specified conditions we partition $N_G(Z)$ into $d(Z)+1$ many sets $(X_j(Z))_{0\leq j\leq d(Z)}$, where $d(Z) := \left\lfloor \frac{\deg_G(Z)}{q} \right\rfloor$, such that

- $|X_0(Z)| < q$ and
- $\forall j \in [d(Z)] : \ |X_j(Z)| = q$.

We embed $\tilde{A}$ by Z. For $j \in [d(Z)]$ we are going to find many appropriate embeddings $S := (v_{b'})_{b'\in B'} \subseteq X_j(Z)$. Namely, we require that for any $a \in A \setminus \tilde{A}$ we can find a subset $U_a(S)$ of the common G-neighborhood of the embeddings of the H-neighborhood of a, formally $U_a(S) \subseteq N_G\left(S\big|_a\right)$, such that

(i) S is independent in G.

(ii) $\forall a \in A \setminus \tilde{A} : \ |U_a(S)| \geq \xi_1 p^d |G|$.

(iii) $\forall a \in A \setminus \tilde{A},\ b' \in B' \setminus N_{B'}(a) : \ \deg_{U_a(S)}(v_{b'}) \leq \xi_2 |U_a(S)|$.

Given such embeddings $Z = (v_a)_{a\in\tilde{A}}$ and $(v_b)_{b\in B'}$ we are going to find many embeddings of $A \setminus \tilde{A}$ into G that extend the embedding $(v_x)_{x\in B'\cup\tilde{A}}$ to an induced isomorphism $S' = (v_x)_{x\in V(H')}$ from H' to G.

Given such an induced isomorphism $S' \in V(G)^{V(H')}$ we are going to find out, using the Packing lemma, that there are many other isomorphisms $S'' \in V(G)^{V(H')}$ that agree wit S' on $V(H)$, meaning $S'|_{V(H)} = S''|_{V(H)}$. For clarity, let us again formulate the conditions on an embedding $(v_x)_{x \in V(H)} \subseteq V(G)$ to be an induced isomorphism from H to G.

(a) $G[\{ v_a \mid a \in A \}, \{ v_b \mid b \in B \}] = H$,

(b) $(v_{\tilde{a}})_{\tilde{a} \in \tilde{A}}$ is independent,

(c) $(v_b)_{b \in B}$ is independent,

(d) $(v_a)_{a \in A \setminus \tilde{A}}$ is independent,

(e) $\| \left\{ v_a \,\middle|\, a \in A \setminus \tilde{A} \right\}, \left\{ v_{\tilde{a}} \,\middle|\, \tilde{a} \in \tilde{A} \right\} \| = 0$.

We want to remark that point (*iii*) will be used to guarantee for (a). In order to guarantee for (b) and (c) we are going to apply Lemma 7. In order to guarantee for (d) we are going to apply Lemma 8. The hardest part of the whole proof is to guarantee for (e). The fact that all vertices $a \in A \setminus \tilde{A}$ have degree d in H' and the fact that $K_{d+1,d+1} \nsubseteq G$ is going to be our essential tool. This remark closes the introductory part of the proof.

Since for any graph G we have that $p = \frac{2\|G\|}{|G|^2} \leq 1$, and we require that $p \geq C|G|^{-\frac{1}{d}}$ we have that the statement of the Theorem is trivial in case that $|G| < C^d$. Thus, we may assume that

$$|G| \geq C^d.$$

Using our assumption that C is large enough we may assume that $|G| \geq 4s^{r-1}$ and Lemma 7 assures that we find many independent embeddings $Z = (v_a)_{a \in \tilde{A}}$, namely

$$|\mathscr{I}_r(G)| \geq 2^{-r} \left(\sqrt{s}\right)^{-r^2} |G|^r = \alpha |G|^r. \tag{3}$$

Let us define the set of all blocks

$$\mathscr{X} := \bigcup_{Z \in \mathscr{I}_r(G)} \{ X_j(Z) \mid j \in [d(Z)] \},$$

where for technical reasons we interpret $\mathscr{X}$ as a multiset. Caring about requirement (i), for $X \in \mathscr{X}$ let us define

$$\mathscr{I}(X) := \mathscr{I}_{n_{B'}}(G[X]).$$

Using that

$$q = \frac{1}{2} p^r |G| \geq \frac{C^r}{2} |G|^{\frac{d-r}{d}} \geq \frac{C^r}{2}$$

and our assumption that C is large enough we may assume that $q \geq 4s^{n_{B'}-1}$. Thus, again Lemma 7 assures that

$$\forall X \in \mathscr{X} : \; |\mathscr{I}(X)| \geq 2^{-n_{B'}} \left(\sqrt{s}\right)^{-n_{B'}^2} q^{n_{B'}} = \beta q^{n_{B'}}. \tag{4}$$

Regarding point (ii) for $X \in \mathscr{X}$ define *bad* and *good sequences of* X.

$$S_{\text{bad}}(X) := \left\{ S \in X^d \,\middle|\, \deg_G(S) < 2\xi_1 p^d |G| \right\}.$$
$$S_{\text{good}}(X) := X^d \setminus S_{\text{bad}}(X).$$

Furthermore, let us define X to be a *bad block* if more than a ξ_3-fraction of all contained d-sequences are bad.

$$\mathscr{X}_{\text{bad}} := \left\{ X \in \mathscr{X} \;\middle|\; |S_{\text{bad}}(X)| \geq \xi_3 q^d \right\}.$$
$$\mathscr{X}_{\text{good}} := \mathscr{X} \setminus \mathscr{X}_{\text{bad}}.$$

Regarding point (iii) for $X \in \mathscr{X}$ as well as $S \in S_{\text{good}}(X)$ let us define *correlated* vertices as the vertices thats neighborhood covers a ξ_2-fraction of the common neighborhood of S.

$$V_{\text{correlated}}(S) := \{ u \in V(G) \mid |X_j(u) \cap N_G(S)| > \xi_2 \deg_G(S) \}.$$

Claim 11. Let $X \in \mathscr{X}_{\text{good}}$. Then for any $S := (v_j)_{j\in[d]} \in S_{\text{good}}(X)$ we can choose $U(S) \subseteq N_G(S)$ such that $|U(S)| \geq \xi_1 |G| p^d$ and the following set is small.

$$V_{\text{correlated}}(U(S)) := \left\{ v \in V(G) \setminus U(S) \;\middle|\; \deg_{U(S)}(v) > 2\xi_2 |U(S)| \right\}.$$

Namely, we have $|V_{\text{correlated}}(U(S))| \leq \dfrac{s}{\xi_2}$.

Proof of Claim 11. For convenience let us define $t := |V_{\text{correlated}}(S)|$ and $d := \deg_G(S)$. Since $S \in S_{\text{good}}(X)$ we have that $d \geq 2\xi_1 |G| p^d \geq 2\xi_1 C^d$. By the assumption that C is large enough we may assume that

$$d \geq 2\left(\frac{40s}{\xi_2}\right)^s.$$

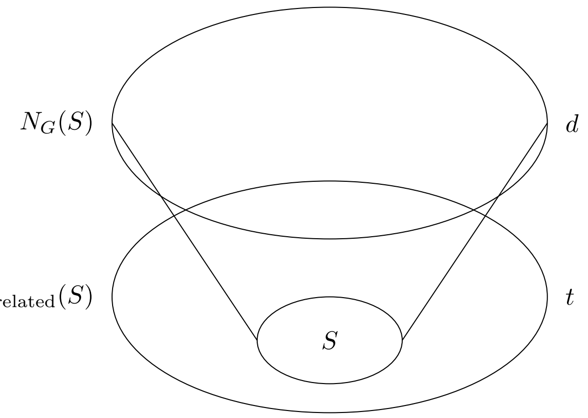

Figure 13: The situation in the proof of Claim 11.

The general idea to prove Claim 11 is to use the Kővári, Sós, Turán theorem and our assumption that G is $K_{s,s}$-free. The problem is that $N_G(S)$ and $V_{\text{correlated}}(S)$ do not have to be disjoint. The bounds obtained by the Kővári, Sós, Turán theorem could still hold in case that $d \gg t$ or $t \gg d$. First, under assumption of some balance restriction for t and d we arrive at a contradiction by a double application of the Kővári, Sós, Turán theorem. After this we will know that $t \leq \frac{d}{2}$ and we will define $U(S) := N_G(S) \setminus V_{\text{correlated}}(S)$. Finally, a third application of the Kővári, Sós, Turán theorem is going to yield the required bound on $|V_{\text{correlated}}(U(S))|$.

Let us assume for a contradiction that

$$t \geq \frac{20s}{\xi_2} d^{1-\frac{1}{s}}. \tag{5}$$

On the one hand by the Definition of correlated vertices we observe that

$$\|G[V_{\text{correlated}}(S) \cup N_G(S)]\| \geq \frac{t \cdot \xi_2 d}{2}$$

where we had to divide by two since the correlated vertices could intersect $N_G(S)$. On the other hand with Observation 8 and with the Kővári, Sós, Turán theorem, especially Corollary 2 since $d \geq 10$, it follows that

$$\|G[V_{\text{correlated}}(S) \cup N_G(S)]\| \leq \text{ex}(K_{t+d}, K_{s,s}) \leq s^{\frac{1}{s}}(t+d)^{2-\frac{1}{s}} \leq s(t+d)^{2-\frac{1}{s}}.$$

Putting both bounds together we obtain

$$td \leq \frac{2s}{\xi_2}(t+d)^{2-\frac{1}{s}} \leq \frac{2s}{\xi_2} \cdot \frac{t^2+2td+d^2}{d^{\frac{1}{s}}}.$$

We can reshape this to

$$\left(1 - \frac{4s}{\xi_2 d^{\frac{1}{s}}}\right) td^{1+\frac{1}{s}} \leq \frac{2s}{\xi_2}\left(t^2+d^2\right).$$

Notice that $d \geq \left(\frac{40s}{\xi_2}\right)^s$ certainly implies that

$$1 - \frac{4s}{\xi_2 d^{\frac{1}{s}}} \geq \frac{1}{2}.$$

Using this and our assumption (5) we deduce

$$\frac{10s}{\xi_2}d^2 = \left(\frac{10s}{\xi_2}d^{1-\frac{1}{s}}\right) \cdot d^{1+\frac{1}{s}} \leq \frac{1}{2}td^{1+\frac{1}{s}} \leq \left(1 - \frac{4s}{\xi_2 d^{\frac{1}{s}}}\right) td^{1+\frac{1}{s}} \leq \frac{2s}{\xi_2}\left(t^2+d^2\right).$$

Calculation yields

$$t^2 \geq 4d^2,$$

meaning that $t \geq 2d$.
To arrive at a contradiction let us consider $X := V_{\text{correlated}}(S) \setminus N_G(S)$. By the previous we have $|X| \geq t-d \geq \frac{t}{2}$. Let us double count the edges between X and $N_G(S)$. On the one hand we have

$$\|X, N_G(S)\| \geq |X| \cdot \xi_2 d.$$

Different than before we did not have to divide by two since X and $N_G(S)$ are disjoint. On the other hand by the Kővári Sós Turán theorem, Lemma 2, we obtain

$$\|X, N_G(S)\| \leq s\left(|X| \cdot d^{1-\frac{1}{s}} + d\right).$$

Notice that $d \geq 2\left(\frac{40s}{\xi_2}\right)^s$ certainly implies that $\xi_2 d \geq 2sd^{1-\frac{1}{s}}$. Using this and putting the two bounds together we arrive at

$$|X| \leq \frac{sd}{\xi_2 d - sd^{1-\frac{1}{s}}} \leq \frac{2s}{\xi_2}.$$

Again using that $d \geq 2\left(\frac{40s}{\xi_2}\right)^s$ we conclude

$$t \leq 2|X| \leq \frac{4s}{\xi_2} < \frac{20s}{\xi_2}d^{1-\frac{1}{s}}.$$

However, this is a contradiction to our assumption (5).
The contradiction argument showed that $t \leq \frac{20s}{\xi_2}d^{1-\frac{1}{s}}$. Using this and $d \geq \left(\frac{40s}{\xi_2}\right)^s$ we calculate

$$t \leq \frac{20sd}{\xi_2 d^{\frac{1}{s}}} \leq \frac{20sd}{\xi_2 \cdot \frac{40s}{\xi_2}} = \frac{d}{2}.$$

Let us define $U(S) \coloneqq N_G(S) \setminus V_{\text{correlated}}(S)$. As required we know that $|U(S)| \geq \frac{d}{2}$. Furthermore, $|U(S)| \geq \dfrac{d}{2} \geq \left(\dfrac{40s}{\xi_2}\right)^s$ and we deduce that

$$\xi_2 |U(S)| \geq 40s|U(S)|^{1-\frac{1}{s}}.$$

Observe that $U(S)$ and $V_{\text{correlated}}(U(S))$ are disjoint since

$$\begin{aligned} V_{\text{correlated}}(U(S)) &= \left\{ u \in V(G) \setminus U(S) \;\middle|\; \deg_{U(S)}(u) > 2\xi_2 |U(S)| \right\} \\ &\subseteq \left\{ u \in V(G) \setminus U(S) \;\middle|\; \deg_{U(S)}(u) > \xi_2 d \right\} = V_{\text{correlated}}(S). \end{aligned}$$

Once more let us double count edges. On the one hand we know that

$$\|V_{\text{correlated}}(U(S)), U(S)\| \geq |V_{\text{correlated}}(U(S))| \cdot 2\xi_2 |U(S)|.$$

On the other hand again the Kővári Sós Turán theorem, Lemma 2, yields

$$\|V_{\text{correlated}}(U(S)), U(S)\| \leq s\left(|V_{\text{correlated}}(U(S))| \cdot |U(S)|^{1-\frac{1}{s}} + |U(S)|\right).$$

Using $\xi_2 |U(S)| \geq s|U(S)|^{1-\frac{1}{s}}$ we conclude that

$$|V_{\text{correlated}}(U(S))| \leq \frac{|U(S)|s}{2\xi_2|U(S)| - s|U(S)|^{1-\frac{1}{s}}} \leq \frac{s}{\xi_2}.$$

This completes the proof of Claim 11. □

For $X \in \mathscr{X}_{\text{good}}$ let us define two types of bad embeddings of B' into X.

$$\begin{aligned} E^{(ii)}_{\text{bad}}(X) &\coloneqq \left\{ S \in X^{B'} \;\middle|\; \exists a \in A \setminus \tilde{A}:\ S\big|_a \in S_{bad}(X, d) \right\}. \\ E^{(iii)}_{\text{bad}}(X) &\coloneqq \left\{ S = (v_b)_{b \in B'} \in X^{B'} \;\middle|\; \exists a \in A \setminus \tilde{A},\ b' \in B' \setminus N_{B'}(a):\ v_{b'} \in V_{\text{correlated}}(U_a(S)) \right\}. \end{aligned}$$

Here the Definition of the latter bad sequences relies on Claim 11, where we define $U_a(S) \coloneqq U\left(S\big|_a\right)$ and the requirement that X is good. Let us define good embeddings

$$E_{\text{good}}(X) \coloneqq \mathscr{I}(X) \setminus \left(E^{(ii)}_{\text{bad}}(X) \cup E^{(iii)}_{\text{bad}}(X)\right).$$

Claim 12. $\forall X \in \mathscr{X}_{\text{good}}:\ |E_{\text{good}}(X)| \geq \frac{\beta}{2} q^{n_{B'}}$.

Proof of Claim 12. First let us upper bound $|E^{(ii)}_{\text{bad}}(X)|$. By Union Bound we obtain

$$\begin{aligned} |E^{(ii)}_{\text{bad}}(X)| &\leq \sum_{S \in X^{B'}} \left(\sum_{a \in A \setminus \tilde{A}} \mathbb{1}\left\{S\big|_a \in S_{\text{bad}}(X)\right\} \right) \\ &= \sum_{a \in A \setminus \tilde{A}} \left(\sum_{S \in X^{B'}} \mathbb{1}\left\{S\big|_a \in S_{\text{bad}}(X)\right\} \right) \leq \sum_{a \in A \setminus \tilde{A}} \left(q^{(n_{B'} - d)} \cdot |S_{\text{bad}}(X)| \right) \leq \xi_3 n_A q^{n_{B'}} = \frac{\beta}{4} q^{n_{B'}}, \end{aligned}$$

where in the last inequality we used that X is a good block.
Now let us upper bound the cardinality of $|E^{(iii)}_{\text{bad}}(X)|$. We already have seen $q \geq \frac{C^r}{2}$ so again using that C is large enough we may assume

$$q \geq 4 n_A n_{B'} \frac{s}{\xi_2 \beta}.$$

Using this, Claim 11 and Union Bound, let us count

$$
\begin{aligned}
|E_{\mathrm{bad}}^{(iii)}(X)| &\leq \sum_{(v_{b'})_{b' \in B'} \in X^{B'}} \left(\sum_{a \in A \setminus \tilde{A}} \left(\sum_{b'' \in B' \setminus N_{B'}(a)} \mathbb{1}\{v_{b''} \in V_{\mathrm{correlated}}(U_a(S))\} \right) \right) \\
&= \sum_{a \in A \setminus \tilde{A}} \left(\sum_{b'' \in B' \setminus N_{B'}(a)} \left(\sum_{(v_{b'})_{b' \in B'} \in X^{B'}} \mathbb{1}\{v_{b''} \in V_{\mathrm{correlated}}(U_a(S))\} \right) \right) \leq n_A n_{B'} \frac{s}{\xi_2} q^{n_{B'}-1} \leq \frac{\beta}{4} q^{n_{B'}}.
\end{aligned}
$$

Thus, we can upper bound

$$
|E_{\mathrm{bad}}^{(ii)}(X) \cup E_{\mathrm{bad}}^{(iii)}(X)| \leq |E_{\mathrm{bad}}^{(ii)}(X)| + |E_{\mathrm{bad}}^{(iii)}(X)| \leq \frac{\beta}{2} q^{n_{B'}}.
$$

Together with $|\mathscr{I}(X)| \geq \beta q^{n_{B'}}$, see (4), this closes the proof of Claim 12. □

Claim 13. $|\mathscr{X}_{\mathrm{good}}| \geq \frac{\alpha}{2^{(r+2)}} |G|^r$.

Proof of Claim 13. First let us show an upper bound on $|\mathscr{X}_{\mathrm{bad}}|$ by a double counting argument. On the one hand

$$
\sum_{X \in \mathscr{X}} |S_{\mathrm{bad}}(X)| = \sum_{S \in V(G)^d} |\{ X \in \mathscr{X} \mid S \in S_{\mathrm{bad}}(X) \}| \leq |G|^d \left(2\xi_1 p^d \cdot |G|\right)^r.
$$

On the other hand

$$
\sum_{X \in \mathscr{X}} |S_{\mathrm{bad}}(X)| \geq \sum_{X \in \mathscr{X}_{\mathrm{bad}}} |S_{\mathrm{bad}}(X)| \geq |\mathscr{X}_{\mathrm{bad}}| \cdot \xi_3 q^d = |\mathscr{X}_{\mathrm{bad}}| \cdot \xi_3 \left(\frac{1}{2}\right)^d (p^r |G|)^d.
$$

Putting together both bounds and plugging in the Definition of ξ_1 we obtain

$$
|\mathscr{X}_{\mathrm{bad}}| \leq \frac{(2\xi_1)^r \, 2^d}{\xi_3} |G|^r \leq \frac{\alpha}{2^{(r+2)}} |G|^r.
$$

Regarding a lower bound for $|\mathscr{X}|$ observe

$$
(|\mathscr{X}| + |\mathscr{I}_r(G)|) \cdot q = \sum_{Z \in \mathscr{I}_r(G)} (d(Z) + 1) \cdot q \geq \sum_{Z \in \mathscr{I}_r(G)} \deg_G(Z) = \sum_{v \in V(G)} |\mathscr{I}_r(G[N_G(v)])|,
$$

where in the first equality we used that we interpret $\mathscr{X}$ as a multiset. Let us again apply Lemma 7 and in a further step Jensens inequality, Lemma 1.

$$
\begin{aligned}
\sum_{v \in V(G)} |\mathscr{I}_r(G[N_G(v)])| &\geq \sum_{v \in V(G)} \mathbb{1}\left\{\deg_G(v) > 4s^{r-1}\right\} |\mathscr{I}_r(G[N_G(v)])| \\
&\geq \sum_{v \in V(G)} \mathbb{1}\left\{\deg_G(v) > 4s^{r-1}\right\} \left(2^{-r} \left(\sqrt{s}\right)^{-r^2} \deg_G(v)^r\right) \\
&= \alpha |G| \sum_{v \in V(G)} \frac{\left(\mathbb{1}\left\{\deg_G(v) > 4s^{r-1}\right\} \deg_G(v)\right)^r}{|G|} \\
&\geq \alpha |G|^{1-r} \left(\sum_{v \in V(G)} \mathbb{1}\left\{\deg_G(v) > 4s^{r-1}\right\} \deg_G(v) \right)^r.
\end{aligned}
$$

Let us show that cutting the small degrees out of the sum is negligible.

$$
\sum_{v \in V(G)} \mathbb{1}\left\{\deg_G(v) > 4s^{r-1}\right\} \deg_G(v) \geq \sum_{v \in V(G)} \left(\deg_G(v) - 4s^{r-1}\right) \geq 2\|G\| - 4s^{r-1}|G| \geq \|G\|,
$$

where we used that $\|G\| \geq C|G|^{2-\frac{1}{d}} \geq C|G| \geq 4s^{r-1}|G|$.
In total, we arrive at the inequality

$$(|\mathscr{X}| + |\mathscr{I}_r(G)|) \cdot \frac{1}{2}p^r|G| \geq \alpha|G|^{1-r}\|G\|^r = \alpha|G|^{1-r}\left(\frac{p|G|^2}{2}\right)^r = \frac{\alpha}{2^r}|G|^{1+r}p^r.$$

Now with

$$|\mathscr{I}_r(G)| \cdot \frac{1}{2}p^r|G| \leq \frac{1}{2}p^r|G|^{1+r} \leq \frac{\alpha}{2^{r+1}}|G|^{1+r}p^r$$

we deduce that

$$|\mathscr{X}| \geq \frac{\alpha}{2^{r+1}}|G|^r.$$

Thus,

$$|\mathscr{X}_{\text{good}}| \geq |G|^r\left(\frac{\alpha}{2^{r+1}} - \frac{\alpha}{2^{r+2}}\right) = |G|^r\frac{\alpha}{2^{r+2}},$$

which closes the proof of Claim 13. □

For $Z = (v_a)_{a\in\tilde{A}} \in \mathscr{I}_r(G)$ and embeddings $S \coloneqq (v_b)_{b\in B} \subseteq N_G(Z)$ let us define the extensions

$$\text{Extensions}(Z,S) \coloneqq \left\{ (v_a)_{a\in A\setminus\tilde{A}} \subseteq V(G) \;\middle|\; (v_x)_{x\in V(H)} \in \text{Isom}_{\text{ind}}(H,\ G) \right\}.$$

Notice that Extensions(Z,S) could be empty in general. For $j \in [d(Z)]$ with $X = X_j(Z) \in \mathscr{X}_{\text{good}}$ we may define $\text{Extensions}_X(Z)$ to be all the extensions of Z to induced isomorphisms from H to G such that all vertices in B are mapped to X.
Now let again $S \coloneqq (v_{b'})_{b'\in B'} \subseteq N_G(Z)$. For each $a \in A$ we call the vertices in G we might choose for embedding a in order to assure the correctness of the edges between the embeddings of a and B' *precandidates.*

$$\text{Precandidates}(a,S) \coloneqq N_{\text{strong}}\left(S,\ S\big|_a\right) \coloneqq \{\, u \in V(G) \mid (N_G(u) \cap \{\, v_{b'} \mid b' \in B' \,\}) = \{\, v_{b'} \mid b' \in N_{B'}(a) \,\} \,\},$$

compare with Definition of strong neighborhood 70. Furthermore, for $a \in A \setminus \tilde{A}$ we refer to the precandidates of a that send no edge towards Z by *candidates.*

$$\text{Candidates}(a,S,Z) \coloneqq \text{Precandidates}(a,S) \setminus \left(\bigcup_{\tilde{a}\in\tilde{A}} N_G(v_{\tilde{a}})\right).$$

The following Observation is our tool to control the edges in G between the embeddings of $\tilde{A}$ and $A \setminus \tilde{A}$.

Claim 14. $\forall a \in A \setminus \tilde{A},\ \tilde{a} \in \tilde{A} :\ |N_G(v_{\tilde{a}}) \cap \text{Precandidates}(a,S)| \leq d.$

Proof of Claim 14. Assume for a contradiction that there exist $a \in A \setminus \tilde{A},\ \tilde{a} \in \tilde{A}$ such that

$$Y \coloneqq N_G(v_{\tilde{a}}) \cap \text{Precandidates}(a,S)$$

fulfills $|Y| \geq d+1$.

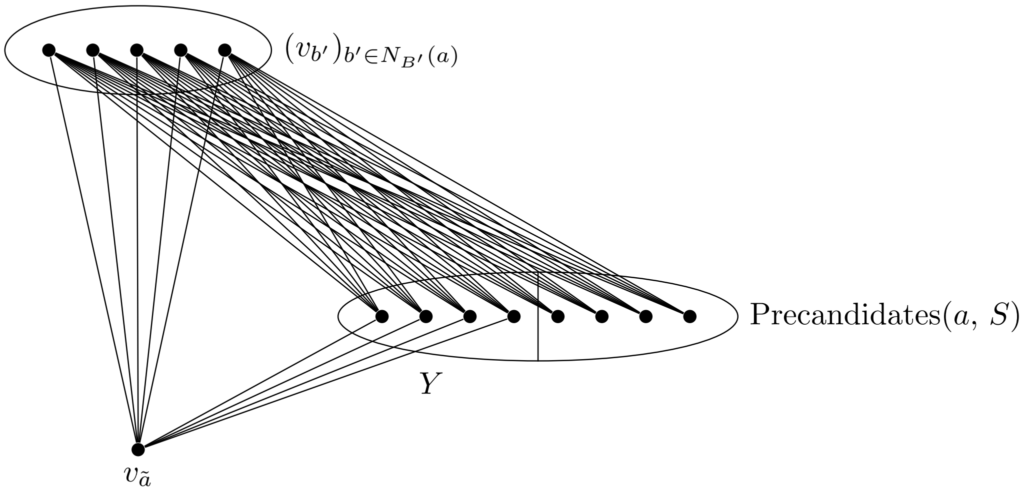

Figure 14: The situation in the proof of Claim 14.

We see that the biinduced subgraph

$$G\left[\left(\{\, v_{b'} \mid b' \in N_{B'}\,(a)\,\} \cup \{v_{\tilde{a}}\}\right),\ Y\right]$$

is a $K_{d+1,|Y|}$, a contradiction. □

Claim 15. Let $Z \in \mathscr{I}_r(G),\ j \in [d(Z)]$ and $X = X_j(Z),\ S \in E_{\text{good}}(X)$ as well as $a \in A \setminus \tilde{A}$. Then

$$|\text{Candidates}(a, S, Z)| \geq \frac{\xi_1}{4} p^d |G|.$$

Proof of Claim 15. By Definition we know that $S\big|_a \in S_{\text{good}}\,(X)$ so by Claim 11 there is $U_a(S) \subseteq N_G\left(S\big|_a\right)$ such that $|U_a(S)| \geq \xi_1 p^d |G|$ and $|V_{\text{correlated}}\,(U_a(S))\,| \leq \frac{s}{\xi_2}$.
First we observe

$$U_a(S) \setminus \left(\bigcup_{b' \in B' \setminus N_{B'}(a)} N_{U_a(S)}\,(v_{b'}) \right) \subseteq \text{Precandidates}(a, S).$$

Since no vertices in $\{\, v_{b'} \mid b' \in B' \setminus N_{B'}\,(a)\,\}$ are *correlated* with $U_a(S)$ we conclude

$$|\text{Precandidates}(a, S)| \geq |U_a(S)| - \sum_{b' \in B' \setminus N_{B'}(a)} \deg_{U_a(S)}\,(v_{b'}) \geq |U_a(S)|\,(1 - 2\xi_2 n_{B'}) \geq \frac{\xi_1}{2} p^d |G|,$$

where we used that by Definition of ξ_2 we have $1 - 2\xi_2 n_{B'} = \frac{1}{2}$. Since $\frac{\xi_1}{2} p^d |G| \geq \frac{\xi_1}{2} C^d$ and we assumed that C is large enough we may assume that $\frac{\xi_1}{2} p^d \geq 2rd$. Using Claim 14 and Union Bound as well as this inequality, we deduce

$$|\text{Candidates}(a, S, Z)| \geq |\text{Precandidates}(a, S)| - rd \geq \frac{\xi_1}{4} p^d |G|.$$

This completes the proof of Claim 15. □

Let $Z \in \mathscr{I}_r(G),\ j \in [d(Z)]$ and $X = X_j(Z),\ S \in E_{\text{good}}(X)$. Let us define the good embeddings of B' into X that agree with S in B by

$$\text{Variants}(S, X) \coloneqq \left\{ \tilde{S} \in E_{\text{good}}(X) \;\middle|\; S\big|_B = \tilde{S}\big|_B \right\}.$$

For $a \in A \setminus \tilde{A}$ we want to choose representative embeddings, one for each class of embeddings that pairwise agree on $B \cup N_{B'}(a)$. Namely, we choose $\text{Variants}_a(S, X) \subseteq \text{Variants}(S, X)$ such that

$$\text{Variants}(S, X) = \sum_{S \in \text{Variants}_a(S,X)} \left\{ \tilde{S} \in \text{Variants}(S, X) \;\middle|\; S\big|_{(N_{(B'\setminus B)}(a))} = \tilde{S}\big|_{(N_{(B'\setminus B)}(a))} \right\}.$$

Furthermore, we want to combine the candidates for a of all $\tilde{S} \in \text{Variants}_a(S, X)$. Let us define

$$\text{Candidates}^*(a, S, Z) := \bigcup_{\tilde{S} \in \text{Variants}_a(S,X)} \text{Candidates}(a, \tilde{S}, Z).$$

Claim 16. Let $Z \in \mathscr{I}_r(G),\ j \in [d(Z)]$ and $X = X_j(Z),\ S \in E_{\text{good}}(X)$ as well as $a \in A \setminus \tilde{A}$. Then

$$|\text{Candidates}^*(a, S, Z)| \geq \frac{\xi_1}{4\tilde{C}} p^d |G| |\text{Variants}_a(S, X)|^{\frac{1}{d+1-r}}.$$

Proof of Claim 16. Let us study the hypergraph

$$\mathscr{F} := \left(\text{Candidates}^*(a, S, Z),\ \left\{ \text{Candidates}(a, \tilde{S}, Z) \;\middle|\; \tilde{S} \in \text{Variants}_a(S, X) \right\} \right).$$

We remark that $\|\mathscr{F}\| = |\text{Variants}_a(S, X)|$ since we allow double hyperedges in $\mathscr{F}$. The key Observation is that a large enough set of hyperedges in $\mathscr{F}$ must have a small intersection. Since the edges are large we can use the Packing lemma, Theorem 10, to show that $|\mathscr{F}|$ is large.
For $\tilde{z} \in \mathbb{N},\ a \in A$ and any $\{ S_j \mid j \in [\tilde{z}] \} \in \binom{\text{Variants}_a(S,X)}{\tilde{z}}$, where for $j \in [\tilde{z}]:\ S_j = (v_{j,b'})_{b' \in B'}$, let us define the set of vertices used for embedding $N_{(B'\setminus B)}(a)$ by some embeddings $(S_j)_{j \in [\tilde{z}]}$.

$$\text{Comb}\left(\{ S_j \mid j \in [\tilde{z}] \}, a\right) := \left\{ v_{j,b'} \;\middle|\; j \in [\tilde{z}],\ b' \in N_{(B\setminus B')}(a) \right\}.$$

We observe

$$|\text{Comb}\left(\{ S_j \mid j \in [\tilde{z}] \}, a\right)|^{\deg_{(B'\setminus B)}(a)} \geq \tilde{z}$$

which yields if we choose $\tilde{z} \geq (d+1)^{\deg_{(B'\setminus B)}(a)}$ that

$$|\text{Comb}\left(\{ S_j \mid j \in [\tilde{z}] \}, a\right)| \geq d + 1.$$

Let us set $z := (d+1)^d$. Observe that for $\{ S_j \mid j \in [z] \} \in \binom{\text{Variants}_a(S,X)}{z}$ the biinduced subgraph

$$G\left[\text{Comb}\left(\{ S_j \mid j \in [z] \}\right), \left(\left\{ v_{\tilde{a}} \;\middle|\; \tilde{a} \in \tilde{A} \right\} \cup \bigcap_{j \in [z]} \text{Candidates}(a, S_j) \right)\right]$$

is complete bipartite which implies that

$$|\bigcap_{j \in [z]} \text{Candidates}(a, S_j)| < d + 1 - r.$$

This shows that for any $W \in \binom{V(\mathscr{F})}{d+1-r}:\ |\{ U \in E(\mathscr{F}) \mid W \subseteq U \}| < z$, which translates into

$$\forall n \in \mathbb{N}:\ \pi_{\mathscr{F}}(n) \leq \sum_{0 \leq i \leq d-r} \binom{n}{i} + z\binom{n}{d+1-r} \leq (e+z)n^{d+1-r},$$

where we used some standard bound on sums of binomial coefficients, that we elaborate in Corollary 3.
Let us set $\delta := \frac{\xi_1}{4} p^d |G|$. Because of Claim 15 and $d + 1 - r < \delta$ we have that $\mathscr{F}$ is (z, δ)-separated. Thus,

Theorem 10 yields that

$$|\mathscr{F}| \geq \delta \left(\frac{\|\mathscr{F}\|}{\tilde{C}} \right)^{\frac{1}{d+1-r}} \geq \frac{\xi_1}{4\tilde{C}} p^d |G| \cdot |\text{Variants}_a(S,X)|^{\frac{1}{d+1-r}},$$

which completes the proof of Claim 16. □

Claim 17. Let $Z = (v_a)_{a \in \tilde{A}} \in \mathscr{I}_r(G)$ such that there is $j \in [d(Z)]$ with $X := X_j(Z) \in \mathscr{X}_{good}$. Then

$$\text{Extensions}_X(Z) \geq \frac{2^{(r+1)}c}{\alpha} p^{\|G\|} |G|^{|H|} \left(p^{d+1}|G|\right)^{\frac{\tau(G)}{d+1-r}} |G|^{-r}.$$

Proof of Claim 17. Let $S \in E_{\text{good}}(X)$. We remark that, since no vertex in $A \setminus \tilde{A}$ is complete

$$\forall \tilde{a} \in \tilde{A},\ a \in A \setminus \tilde{A}:\ v_{\tilde{a}} \notin \text{Candidates}(a).$$

Observe that for distinct $a, a' \in A \setminus \tilde{A}$ the candidate sets could intersect in case that $N_B(a) = N_B(a')$. However, Claim 16 yields that $|\text{Candidates}^*(a,S,Z)| \geq \frac{\xi_1 C^d}{4\tilde{C}}$ and by the assumption that C is large enough we may assume

$$|\text{Candidates}^*(a,S,Z)| \geq n_A \cdot s^{n_B - 1}.$$

Thus, we can equipartition each candidate set into n_A many sets, each of size at least s^{n_B-1}, assign every vertex in A its own part and denote it by $\text{Candidates}^+(a,S,Z)$. Using Claim 16 we deduce

$$\begin{aligned} |\text{Candidates}^+(a,S,Z)| &\geq \left\lfloor \frac{|\text{Candidates}^*(a,S,Z)|}{n_A} \right\rfloor \\ &\geq \frac{|\text{Candidates}^*(a,S,Z)|}{2n_A} \geq \frac{\xi_1}{8\tilde{C}n_A} p^d |G| |\text{Variants}_a(S,X)|^{\frac{1}{d+1-r}}. \end{aligned}$$

Let us choose representative embeddings $\text{Repr}(X) \subseteq E_{\text{good}}(X)$ such that

$$E_{\text{good}}(X) = \sum_{S \in \text{Repr}(X)} \text{Variants}(S,X).$$

We want to show that at least some representatives have a large variant set by double counting. Let us define

$$\begin{aligned} \text{Repr}_{\text{bad}}(X) &:= \left\{ S \in \text{Repr}(X) \;\middle|\; |\text{Variants}(S,X)| \leq \frac{\beta}{4} q^{\tau(G)} \right\}. \\ \text{Repr}_{\text{good}}(X) &:= \text{Repr}(X) \setminus \text{Repr}_{\text{bad}}(X). \end{aligned}$$

On the one hand, using

$$\forall S \in \text{Repr}(X):\ |\text{Variants}(S,X)| \leq q^{(n_{B'} - n_B)} = q^{\tau(G)} \text{ and } |\text{Repr}_{\text{bad}}(X)| \leq q^{n_B}$$

we see

$$\begin{aligned} |E_{\text{good}}(X)| &= \sum_{S \in \text{Repr}(X)} |\text{Variants}(S,X)| \\ &\leq |\text{Repr}_{\text{good}}(X)| \cdot q^{\tau(G)} + |\text{Repr}_{\text{bad}}(X)| \cdot \frac{\beta}{4} q^{\tau(G)} \\ &\leq |\text{Repr}_{\text{good}}(X)| \cdot q^{\tau(G)} + \frac{\beta}{4} q^{n_{B'}}. \end{aligned}$$

On the other hand Claim 12 yields

$$|E_{\text{good}}(X)| \geq \frac{\beta}{2} q^{n_{B'}}.$$

Putting both bound together we obtain

$$|\text{Repr}_{\text{good}}(X)| \geq \frac{\beta}{4} q^{n_B}.$$

In order to assure the independence of the embedding of $A \setminus \tilde{A}$ we use Lemma 8. We already guaranteed that

$$\forall a \in A \setminus \tilde{A}: \ |\text{Candidates}^+(a, S, Z)| \geq s^{n_B - 1}.$$

Thus, if we divide each candidate set into parts of size s^{n_B-1} and possibly one smaller left-over part, we have at least one part of size s^{n_B-1}. For $S \in \text{Repr}_{\text{good}}(X)$ Lemma 8 yields

$$|\text{Extensions}(Z, S\big|_B)| \geq \prod_{a \in A \setminus \tilde{A}} \left\lfloor \frac{|\text{Candidates}^+(a, S, Z)|}{s^{n_B-1}} \right\rfloor \geq \prod_{a \in A \setminus \tilde{A}} \frac{|\text{Candidates}^+(a, S, Z)|}{2s^{n_B-1}}.$$

Using the previous we deduce

$$\begin{aligned} \prod_{a \in A \setminus \tilde{A}} |\text{Candidates}^+(a, S, Z)| &\geq \prod_{a \in A \setminus \tilde{A}} \frac{\xi_1}{8\tilde{C} n_A} p^d |G| |\text{Variants}_a(S, X)|^{\frac{1}{d+1-r}} \\ &= \left(\frac{\xi_1}{8\tilde{C} n_A} p^d |G| \right)^{n_A - r} \left(\prod_{a \in A \setminus \tilde{A}} |\text{Variants}_a(S, X)| \right)^{\frac{1}{d+1-r}}. \end{aligned}$$

Furthemore,

$$\prod_{a \in A \setminus \tilde{A}} |\text{Variants}_a(S, X)| \geq |\text{Variants}(S, X)| \geq \frac{\beta}{4} q^{\tau(G)}$$

where in the second inequality we used the Definition of good representatives. Combining all arguments together we count

$$\begin{aligned} |\text{Extensions}_X(Z)| &\geq \sum_{S \in \text{Repr}(X)} |\text{Extensions}(Z, S\big|_B)| \\ &\geq \sum_{S \in \text{Repr}_{\text{good}}(X)} |\text{Extensions}(Z, S\big|_B)| \\ &\geq |\text{Repr}_{\text{good}}(X)| \cdot \left(\frac{\xi_1}{16\tilde{C} n_A s^{n_B-1}} p^d |G| \right)^{n_A - r} \left(\frac{\beta}{4} q^{\tau(G)} \right)^{\frac{1}{d+1-r}} \\ &\geq \left(\frac{\beta}{4} \right)^{1 + \frac{1}{d+1-r}} \left(\frac{\xi_1}{16\tilde{C} n_A s^{n_B-1}} p^d |G| \right)^{n_A - r} q^{n_B + \frac{\tau(G)}{d+1-r}} \\ &\geq \left(\frac{\beta^2}{2^{n_{B'}+4}} \left(\frac{\xi_1}{16\tilde{C} n_A s^{n_B-1}} \right)^{n_A - r} \right) \cdot p^{d(n_A - r) + r\left(n_B + \frac{\tau(G)}{d+1-r}\right)} |G|^{n_B + n_A - r + \frac{\tau(G)}{d+1-r}}. \end{aligned}$$

Furthermore, we calculate

$$\begin{aligned} p^{d(n_A - r) + r\left(n_B + \frac{\tau(G)}{d+1-r}\right)} |G|^{n_B + n_A - r + \frac{\tau(G)}{d+1-r}} &= p^{\|G\| + \tau(G)\left(1 + \frac{r}{d+1-r}\right)} |G|^{|H| - r + \frac{\tau(G)}{d+1-r}} \\ &= p^{\|G\|} |G|^{|H|} \left(p^{d+1} |G|\right)^{\frac{\tau(G)}{d+1-r}} |G|^{-r}. \end{aligned}$$

The constant term turns out to be

$$\frac{\beta^2}{2^{n_{B'}+4}}\left(\frac{\xi_1}{16\tilde{C}n_A s^{n_B-1}}\right)^{n_A-r} = \frac{2^{(r+2)}c}{\alpha}.$$

This completes the proof of Claim 17. □

Finally, let us put all together. Using Claim 13 and 17 we obtain

$$|\mathrm{Isom}_{\mathrm{ind}}\,(H,\ G)\,| \geq \sum_{Z\in\mathscr{I}_r(G)}\left(\sum_{\substack{j\in[d(Z)]:\\ X_j(Z)\in\mathscr{X}_{\mathrm{good}}}}|\mathrm{Extensions}_{X_j(Z)}(Z)|\right) \geq c\cdot p^{\|G\|}|G|^{|H|}\left(p^{d+1}|G|\right)^{\frac{\tau(G)}{d+1-r}}.$$

This completes the proof of Theorem 19. □

5 Erdős-Hajnal conjecture

The goal of this section is to present the proof of the Erős-Hajnal conjecture for graphs of bounded VC dimension in a comprehensible manner. We give an introduction to the problem and survey related results in section 5.1. Here we also draw the connection to the induced forbidden subgraph problem. In section 5.2 we provide the full proof of the generalization of the Ultra Strong Regularity lemma for graphs of bounded VC dimension to hypergraphs, a result by Fox, Pach and Suk. Here we corrected a minor error in a helping Lemma. The proof of the Erdős-Hajnal conjecture for graph properties of bounded VC dimension can be found in section 5.3.

5.1 Notation and Introduction

Let us start by introducing the key notation of this section. For this purpose let G be a graph.

Definition 71 (Homogeneous set)**.** We call a set $A \subseteq V(G)$ *homogeneous* in case it is empty or a clique. We denote the size of the largest homogeneous set by $\alpha \vee \omega(G) \coloneqq \max\{\omega(G), \alpha(G)\}$.

It turns out to be helpful to introduce the following weaker notion of ϵ-restrictedness.

Definition 72 (ϵ-restricted set)**.** $\Delta_\delta(G) \coloneqq \min\left\{\Delta(G), \Delta\left(\overline{G}\right)\right\}$. For $\epsilon > 0$ let us call a non-empty set $S \subseteq V(G)$ an *ϵ-restricted set* in case $\Delta_\delta(G[S]) \leq \epsilon|S|$. In this case we also call the graph $G[S]$ an *ϵ-restricted graph.*

Definition 73 (Erdős-Hajnal property)**.** Let $\mathscr{C}$ be a hereditary graph property, see Definition 10. We say $\mathscr{C}$ has the Erdős-Hajnal property if

$$\exists C > 0\ \forall G \in \mathscr{C}:\ \alpha \vee \omega(G) \geq |G|^C.$$

Conjecture 3 (Erdős-Hajnal conjecture, [16])**.** Every hereditary graph property has the Erdős-Hajnal property.

While the general Conjecture is open, it is shown in a version for bipartite graphs.

Definition 74 (Homogeneous pair)**.** Let $X, Y \subseteq V(G)$ be two disjoint vertex subsets. We call the pair $\{X, Y\}$ *complete* if G contains all edges $\{\ \{x, y\} \mid x \in X,\ y \in Y\ \}$. We call the pair *anticomplete* if G contains none of the edges in $\{\ \{x, y\} \mid x \in X,\ y \in Y\ \}$. We call the pair *homogeneous* if it is complete or anticomplete.

Analogously as we weakened the notion of homogeneous sets to ϵ-restricted sets, we now introduce the notion of ϵ-restrictedness to pairs of vertex subsets.

Definition 75 (ϵ-restricted towards)**.** Let $X, Y \subseteq V(G)$ be two disjoint vertex subsets and $\epsilon > 0$. We say X is *ϵ-restricted towards* Y if either $\max_{x\in X} \deg_Y(x) \leq \epsilon|Y|$ or $\min_{x\in X} \deg_Y(x) \geq (1-\epsilon)|Y|$. Furthermore, we call $\{X, Y\}$ an ϵ restricted pair in case that X is ϵ-restricted towards Y and Y is ϵ-restricted towards X.

Theorem 22 (Erdős-Hajnal conjecture for bipartite graphs, [17])**.** Let $H = (A \cup B,\ F)$ be a bipartite graph where $a \coloneqq |A|$ and $b \coloneqq |B|$ fulfill $1 \leq a \leq b$. Then for any $n \in \mathbb{N}$ and any bipartite graph $G = (X \cup Y,\ E)$ with $|X| = |Y| = n$ and $H \underset{ind}{\not\subseteq^*} G$ there is $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| = |Y'| = \left\lfloor \left(\frac{n}{b}\right)^{\frac{1}{a}} \right\rfloor$ such that the pair $(X',\ Y')$ is homogeneous.

This is significant as the simple sampling argument in the next Observation demonstrates. We became aware of it in Lemma 3.7 in [9].

Observation 34. For any $n \geq 8$ there is a bipartite graph H on partite sets A, B of size n each, such that for any $s \geq 2\log_2(n)$ there is no $A' \in \binom{A}{s},\ B' \in \binom{B}{s}$ such that the pair $(X',\ Y')$ is homogeneous.

Proof of Observation 34. Fix $n \in \mathbb{N}$ and let $X = \left(X_{(i,\ j)}\right)_{(i,\ j)\in[n]^2} \overset{\text{iid}}{\sim} \text{Be}\left(\frac{1}{2}\right)$. Define a random graph on two disjoint partite sets $A = \{\ a_i \mid i \in [n]\ \}$ and $B = \{\ b_j \mid j \in [n]\ \}$ of size n each by

$$G_X \coloneqq \left(A \cup B,\ \left\{\ \{a_i, b_j\}\ \middle|\ i, j \in [n]\ \text{ with } X_{(i,\ j)} = 1\ \right\}\right).$$

For $s \in [n]$ define the event

$$\mathscr{A}(s) = \left\{ \exists A' \in \binom{A}{s},\ B' \in \binom{B}{s} : \text{ the pair } \{A', B'\} \text{ is complete or anticomplete in } G_X \right\}.$$

By Union Bound and independence

$$\mathbb{P}(\mathscr{A}(s)) \leq 2\binom{n}{s}^2 \left(\frac{1}{2}\right)^{s^2} \leq \left(\frac{ne}{s}\right)^{2s} \left(\frac{1}{2}\right)^{s^2-1}$$

which in turn is less than 1 if

$$2s\left(\log_2(n) + \log_2(e) - \log_2(s)\right) - \left(s^2 - 1\right) < 0.$$

To show this let us first remark that in case $n \geq 8$ and $s \geq 2\log_2(n)$ we have $s \geq 6$. This implies that $\log_2(s) \geq \log_2(6) > \frac{1}{2s} + \log_2(e)$. Completing the proof of Claim 34 we deduce

$$2s\left(\log_2(n) + \log_2(e) - \log_2(s)\right) + 1 < 2s\log_2(n) \leq s^2. \qquad \square$$

In the paper, where Erdős and Hajnal made their Conjecture, namely in [16], they already showed the following.

Theorem 23 (Erdős, Hajnal [16])**.** For every graph H there is $C > 0$ such that for all graphs G

$$H \underset{ind}{\nsubseteq} G \implies \alpha \vee \omega(G) \geq 2^{C\sqrt{\log_2(|G|)}}.$$

In 2017 Fox, Pach, Suk almost showed Conjecture 3 for graphs of bounded VC dimension.

Theorem 24 (Fox, Pach, Suk [20])**.** For every $d \in \mathbb{N}$ there is a function $\phi_H(n) = o(1)$ $(n \longrightarrow \infty)$ such that for any graph G

$$\dim_{\mathrm{VC}}(G) \leq d \implies \alpha \vee \omega(G) \geq 2^{(\log_2(|G|))^{1-\phi_H(|G|)}}.$$

During the work on this thesis however Conjecture 3 has been shown for graphs with bounded VC dimension even in a stronger form by Nguyen, Scott and Seymour. To state the result correctly we need the following Definition.

Definition 76 (Polynomial Rödl property)**.** Let $\mathscr{C}$ be a hereditary graph property. We say $\mathscr{C}$ has the polynomial Rödl property if for every $\epsilon > 0$ every graph of the property contains an ϵ-restricted induced subgraph of size linear in $|G|$ and polynomial in ϵ. Formally

$$\exists C > 0\ \forall \epsilon \in \left(0, \frac{1}{2}\right),\ G \in \mathscr{C}\ \exists H \underset{ind}{\subseteq} G : \ \Delta_\delta(H) \leq \epsilon|H| \text{ and } |H| \geq \epsilon^C |G|.$$

Observation 35. If a hereditary graph property $\mathscr{C}$ has the polynomial Rödl property then it also has the Erdős-Hajnal property.

Proof of Observation 35. Let $C > 0$ be given by the polynomial Rödl property. Choose $C' \in \left(0, \frac{1}{1+C}\right)$ and let $G \in \mathscr{C}$. Let us denote $n := |G|$.
By the polynomial Rödl property for $\epsilon := n^{-C'}$ there is $H \underset{ind}{\subseteq} G$ such that $\Delta_\delta(H) \leq \epsilon$ and $|H| \geq \epsilon^C n = n^{1-C\cdot C'} \geq n^{C'}$ where we used $C' < \frac{1}{1+C}$ in the last inequality.
In case $\delta(H) \geq (1-\epsilon)|H|$ one finds $\|H\| \geq \dfrac{|H| \cdot \delta(H)}{2} > (1-\epsilon)\dfrac{|H|^2}{2}$ and by Observation 1 we find $\omega(H) \geq \frac{1}{\epsilon} = n^{C'}$. In case $\Delta(H) \leq \epsilon|H|$ we find that $\Delta\left(\overline{H}\right) \geq |H| - 1 - \epsilon|H| \geq |H|(1-2\epsilon)$ where we used that $|H| \geq n^{C'} = \frac{1}{\epsilon}$. Analogously as in the former case we find that $\alpha(H) \geq \frac{1}{2\epsilon} = \frac{n^{C'}}{2}$. Since H is an induced subgraph of G a homogeneous set in H is also a homogeneous set in G. Now choose $C'' > 0$ such that $\forall n \in \mathbb{N} : \ n^{C''} \geq 1 \implies n^{C''} \leq \frac{n^{C'}}{2}$. We have shown that $\mathscr{C}$ has the Erdős-Hajnal property with constant C''. $\square$

In this section we will provide the proof to following Theorem.

Theorem 25 (Nguyen, Scott and Seymour [40])**.** Let $\mathscr{C}$ be a hereditary graph property such that $d := \dim_{\mathrm{VC}}(\mathscr{C}) < \infty$, see Definition 52. Then $\mathscr{C}$ fulfills the polynomial Rödl property as well as the Erdős-Hajnal property.

We remark that in our extremal standard setting we have polynomially large homogeneous sets.

Observation 36. For a bipartite graph $H = (A \cup B,\ E)$ with $A \neq \emptyset \neq B$ and $k \in \mathbb{N}$ the hereditary graph property $\mathrm{Free}(\{K_k, H\text{-ind}\})$ fulfills the Erdős-Hajnal property.

Observation 36 could be seen as a Corollary of Theorem 25. Indeed, Observation 22 and Theorem 25 even imply the polynomial Rödl property of $\mathrm{Free}(\{K_k, H\text{-ind}\})$. However, the Erdős-Hajnal property of $\mathrm{Free}(\{K_k, H\text{-ind}\})$ is a simple consequence of the Kővári, Sós, Turán theorem and Turáns Theorem.

Proof of Observation 36. The Kővári, Sós, Turán theorem, especially Corollary 2 yields $\tilde{C} > 0$, we may assume that $\tilde{C} \geq 1$, such that

$$\forall G \in \mathrm{Free}(\{K_k, H\text{-ind}\}) :\ \|G\| \leq \tilde{C}|G|^{2-\frac{1}{k}}.$$

For $G \in \mathrm{Free}(\{K_k, H\text{-ind}\})$ let us calculate

$$\|\overline{G}\| \geq \binom{|G|}{2} - \tilde{C}|G|^{2-\frac{1}{k}} = \left(1 - \frac{1}{|G|} - 2\tilde{C}|G|^{-\frac{1}{k}}\right)\frac{|G|^2}{2} \geq \left(1 - (2\tilde{C}+1)|G|^{-\frac{1}{k}}\right)\frac{|G|^2}{2}.$$

In case $|G| \geq (4\tilde{C}+2)^k$ we have $(2\tilde{C}+1)|G|^{-\frac{1}{k}} \leq \frac{1}{2}$ and Corollary 1 yields

$$\alpha(G) = \omega\left(\overline{G}\right) \geq \frac{1}{2\tilde{C}+1}|G|^{\frac{1}{k}}.$$

Now for some $C > 0$ large enough both in case $|G| \geq (4\tilde{C}+2)^k$ and $|G| \leq (4\tilde{C}+2)^k$

$$\forall G \in \mathrm{Free}(\{K_k, H\text{-ind}\}) :\ \alpha \vee \omega(G) \geq |G|^C. \qquad \square$$

The proof of Theorem 25 uses itself the following Theorem on the Erdős-Hajnal conjecture.

Theorem 26 (Nguyen, Scott and Seymour [40])**.** For a bipartite graph $H = (A \cup B,\ E)$ with $A \neq \emptyset \neq B$ the hereditary graph property $\mathrm{Free}(H\text{-biind})$ fulfills the Erdős-Hajnal property.

5.2 Ultra Strong Regularity lemma for graphs of bounded VC dimension

The main tool for the proof of Theorem 26 is the Ultra Strong Regularity lemma for graphs with bounded VC dimension which we give in a hypergraph version. The presented proof originates from Fox, Pach, Suk in [20]. They generalized earlier versions for graphs, see [1], to uniform hypergraphs.

Definition 77 (ϵ-homogeneous partition)**.** Let $k \in \mathbb{N}_{\geq 2}$ and $\mathscr{H} = ([n],\ \mathscr{E})$ be a k-uniform hypergraph as well as $(V_j)_{j\in[k]} \subseteq [n]$ be a sequence of k non-empty and pairwise disjoint subsets of the vertices of $\mathscr{H}$. Furthermore, let $\epsilon \in \left(0, \frac{1}{2}\right)$. We call $(V_j)_{j\in[k]}$ *ϵ-homogeneous in $\mathscr{H}$* if $\frac{\|V_1,\ldots,V_k\|}{|V_1|\cdot\ldots\cdot|V_k|} \in [0,\epsilon) \cup (1-\epsilon, 1]$. In case $k = 2$ we speak of *ϵ-homogeneous pairs.*

It is worthwhile to compare this Definition with Definitions 74 and 75.

Theorem 27 (Ultra Strong Regularity, Fox, Pach, Suk [20])**.** $\forall d, k \in \mathbb{N},\ k \geq 2\ \exists c > 0\ \forall \epsilon \in \left(0, \frac{1}{4}\right),\ n \in \mathbb{N}$ and any k-uniform hypergraph $\mathscr{H} = ([n],\ \mathscr{E})$ with $\dim^{(1)*}_{\mathrm{VC}}(\mathscr{H}) = d$ there is $K \in \mathbb{N}$ with $\frac{8}{\epsilon} \leq K \leq c\left(\frac{1}{\epsilon}\right)^{3d+1}$ and an equitable partition $(V_j)_{j\in[K]}$ of the vertex set $[n]$ such that $X := \left\{ J \in \binom{[K]}{k} \,\middle|\, (V_j)_{j\in J} \text{ not } \epsilon\text{-homogeneous} \right\}$ is small, namely $|X| \leq \epsilon\binom{K}{k}$.

Note that $\dim_{\mathrm{VC}}^{(1)*}(\mathscr{H})$ describes not the usual dual VC dimension but the generalization of dual VC dimension for graphs to hypergraphs, see Definition 56. However, in case that $k = 2$ by Observation 13 $\dim_{\mathrm{VC}}^{(1)*}(\mathscr{H}) = \dim_{\mathrm{VC}}(\mathscr{H})$, where the latter VC dimension is the open VC dimension of graphs, see Definition 50.

In the proof of a helping Lemma for Theorem 27 presented in [20] there was a small error, which led to the bound $K \le c\left(\frac{1}{\epsilon}\right)^{2d+1}$ instead of $K \le c\left(\frac{1}{\epsilon}\right)^{3d+1}$, see Errata 1 at the end of this section.
In order to prove Theorem 27 we develop an indicator for a sequence of vertex sets in a hypergraph not to be ϵ-homogeneous. Let $k, m \in \mathbb{N}$ and $(W_j)_{j\in[k]}$ be a sequence of pairwise disjoint and non-empty sets, each of the same cardinality m. With a slight abuse of notation we want to identify

$$\bigtimes_{j\in[k]} W_j = \left\{ \{w_1, \dots, w_k\} \;\middle|\; (w_1, \dots, w_k) \in \bigtimes_{j\in[k]} W_j \right\}.$$

Let $\mathscr{H} = \left(\sum_{j\in[k]} W_j,\ \mathscr{E} \right)$ be a hypergraph. The following notion is going to help identify non-ϵ-homogeneous partitions.

Definition 78. $\mathrm{Fringes}_{\mathscr{H}}\left((W_j)_{j\in[k]}\right) := \left\{ (p,\ p') \in \left(\bigtimes_{j\in[k]} W_j\right)^2 \;\middle|\; p \in \mathscr{E},\ p' \notin \mathscr{E} \text{ and } |p \cap p'| = k-1 \right\}.$

Lemma 22 (Fox, Pach, Suk [20])**.** For any $\epsilon \in \left(0, \frac{1}{2}\right)$:

$$(W_j)_{j\in[k]} \text{ not } \epsilon\text{-homogeneous in } \mathscr{H} \implies \left|\mathrm{Fringes}_{\mathscr{H}}\left((W_j)_{j\in[k]}\right)\right| \ge \epsilon^2 m^{k+1}.$$

Proof of Lemma 22. Remember that in case $(W_j)_{j\in[k]}$ not ϵ-homogeneous $\dfrac{\|W_1, \dots, W_k\|}{|W_1| \cdot \ldots \cdot |W_k|} \in [\epsilon, 1-\epsilon]$. This has as consequence that if we independently draw $\{a_1, \dots, a_k\}, \{b_1, \dots, b_k\}$ from $\bigtimes_{j\in[k]} W_j$ uniformly at random then

$$\mathbb{P}\left(\{a_1, \dots, a_k\} \in \mathscr{E}, \{b_1, \dots, b_k\} \notin \mathscr{E}\right) \ge \epsilon(1-(1-\epsilon)) = \epsilon^2.$$

Since every pair of k-sets $(p,\ p') \in F := \mathrm{Fringes}_{\mathscr{H}}\left((W_j)_{j\in[k]}\right)$ intersects in $k-1$ vertices, the two hyperedges can only differ inside one of the sets $(W_j)_{j\in[k]}$. Let us partition F accordingly. For $j \in [k]$ we define

$$\tilde{F}_j := \left\{ (p,\ p') \in \left(\bigtimes_{j\in[k]} W_j\right)^2 \;\middle|\; p\Delta p' \subseteq W_j \right\}.$$
$$F_j := F \cap \tilde{F}_j = \left\{ (p,\ p') \;\middle|\; (p\Delta p') \in \binom{W_j}{2} \right\}.$$

Using the drawn hyperedges $\{a_1, \dots, a_k\}$ and $\{b_1, \dots, b_k\}$ we want to define a sequence of hyperedges entwining around the partition classes. For $0 \le i \le k$ define the following random variables

$$e_i := \{a_1, \dots, a_j, b_{j+1}, \dots, b_k\}.$$

Observe that $e_0 = \{b_1, \dots, b_k\}$ and $e_k = \{a_1, \dots, a_k\}$ and for any $j \in [k]$ the hyperedges e_j and e_{j-1} only differ inside W_j, formally $e_j \Delta e_{j-1} \subseteq W_j$. Furthermore, (e_{j-1}, e_j) is uniformly distributed in $\tilde{F}_j$.
Let us deduce for $j \in [k]$ that

$$\mathbb{P}\left(e_{j-1} \in \mathscr{E}, e_j \notin \mathscr{E}\right) = \mathbb{P}\left((e_{j-1}, e_j) \in F_j\right) = \frac{|F_j|}{|\tilde{F}_j|} = \frac{|F_j|}{m^{k+1}}.$$

Observe that in case of the event $\{e_0 \in \mathscr{E},\ e_k \notin \mathscr{E}\}$ there has to be at least one $j \in [k]$ such that $e_{j-1} \in \mathscr{E}$ and

$e_j \notin \mathscr{E}$. We deduce by Union Bound

$$\mathbb{P}\left(e_0 \in \mathscr{E}, e_k \notin \mathscr{E}\right) \leq \sum_{j\in[k]} \mathbb{P}\left(e_{j-1} \in \mathscr{E}, e_j \notin \mathscr{E}\right).$$

We close the proof of Lemma 22 by putting together the observed equalities and inequalities.

$$|F| = \sum_{j\in[k]} |F_j| = m^{k+1} \sum_{j\in[k]} \mathbb{P}\left(e_{j-1} \in \mathscr{E}, e_j \notin \mathscr{E}\right) \geq m^{k+1}\mathbb{P}\left(e_0 \in \mathscr{E}, e_k \notin \mathscr{E}\right) \geq m^{k+1}\epsilon^2. \qquad \square$$

Proof of Theorem 27. Using Definitions 55, 56 we define the hypergraph

$$\mathscr{F} := \mathscr{N}_{\mathscr{H}}\left[\binom{[n]}{k-1}\right] = \left(\binom{[n]}{k-1}, \left\{ N_{\mathscr{H}}(U) \;\middle|\; U \in \binom{V(\mathscr{H})}{1} \right\}\right).$$

Then by the requirements and Observation 13 we have $\dim_{\mathrm{VC}}(\mathscr{F}) = \dim_{\mathrm{VC}}^{(k-1)}(\mathscr{H}) = \dim_{\mathrm{VC}}^{(1)*}(\mathscr{H}) = d$. The Sauer lemma, Corollary 3, yields that $\forall z \in \mathbb{N}:\ \pi_{\mathscr{F}}(z) \leq ez^d$. We apply the Packing lemma 10 with parameter $(2, d)$ to obtain a constant $C = C(d) > 0$. Furthermore, we fix $\delta' := \frac{\epsilon^3}{ke^k 2^{k+1}}$ and set $\delta := \left\lceil \delta' \binom{n}{k-1} \right\rceil$.
After the initial preparations let us find the required partition of $\mathscr{H}$. In case a pair of vertices a, b inside one of our partition classes has very similar neighborhoods, meaning $|N_{\mathscr{H}}(\{a\}) \Delta N_{\mathscr{H}}(\{b\})|$ is small, we expect them to behave similarly in regard to edges spanned towards sequences of other partition classes. Now if all vertices in all partition classes have pairwise very similar neighborhoods one could come to the idea that there are many ϵ-regular k-sequences of partition classes. This proof is going to show to us that this phenomenon is strong enough to yield the required partition.
Let $l \in \mathbb{N}$ be maximal such that there is $S = \{ s_j \mid j \in [l] \} \in \binom{[n]}{l}$ such that

$$\mathscr{E}(S) := \{ \mathscr{N}(\{s_j\}) \mid j \in [l] \}$$

is δ-separated, see Definition 67. The Packing lemma yields that

$$l \leq C\left(\frac{\binom{n}{k-1}}{\left\lceil \delta' \binom{n}{k-1} \right\rceil}\right)^d \leq C\left(\frac{1}{\delta'}\right)^d = C\left(\frac{ke^k 2^{k+1}}{\epsilon^3}\right)^d.$$

The idea is to assign every vertex v to a partition belonging to a vertex in S that has the most similar neighborhood as v. For this purpose let us define the mapping $\phi : [n] \longrightarrow [l]$

$$v \mapsto \min\{ j \in [l] \mid |\mathscr{N}(\{v\}) \Delta \mathscr{N}(\{s_j\})| < \delta \}.$$

This is well-defined by the maximality of l. Using ϕ we define the partition

$$U_i := \{ v \in [n] \mid \phi(v) = i \} \ \ (i \in [l]).$$

We observe that by the triangle inequality (M3) in Observation 25

$$\forall j \in [l]\ \ v, w \in U_j:\ |N_{\mathscr{H}}(v) \Delta N_{\mathscr{H}}(w)| \leq |N_{\mathscr{H}}(v) \Delta N_{\mathscr{H}}(s_j)| + |N_{\mathscr{H}}(s_j) \Delta N_{\mathscr{H}}(w)| \leq 2\delta.$$

The second requirement for the partition is equitability. We will simply chop up the partition classes $(U_j)_{j\in[l]}$ into appropriate pieces, where we also need to take care of the leftovers.
We propose that a good number of partition classes in our final partition is given by $K := \left\lfloor \frac{8kl}{\epsilon} \right\rfloor$. Let us check that this choice fulfills the required restrictions $\frac{8}{\epsilon} \leq K \leq c\left(\frac{1}{\epsilon}\right)^{3d+1}$ for some constant $c > 0$. The lower bound holds since $k \geq 2$. On the other hand

$$K \leq \frac{8kl}{\epsilon} \leq 8Ck\left(ke^k 2^{k+1}\right)^d \left(\frac{1}{\epsilon}\right)^{3d+1} = c\left(\frac{1}{\epsilon}\right)^{3d+1}$$

if we choose $c := c(d, k) := 8Ck\left(ke^k 2^{k+1}\right)^d$.

For partitioning let us define the sizes of the final partition classes by

$$q_i := \left\lfloor \frac{n}{K} \right\rfloor + \mathbb{1}\{i \in [(n \bmod K)]\} \ \ (i \in [K]).$$

Now we are going to iteratively fill the partition classes $(V_i)_{i\in[K]}$ in the following manner. We go through the sets $(U_j)_{j\in[l]}$ and fill one set V_i after another. In case that we have distributed all vertices from a set U_j but the set V_i we are currently filling is not full yet, we are continuing filling V_i with vertices from U_{j+1}. Let us fix the indices $\mathscr{I}_{\text{mixed}} \subseteq [K]$ of partition classes V_i with vertices mixed from different sets U_j, U_{j+1}.
We remark that this procedure is well-defined and $|\mathscr{I}_{\text{mixed}}| \leq l-1$ since every mixed set corresponds to at least one set U_j and its successor U_{j+1}, where $j \in [l-1]$.
It remains to check that $X := \left\{ J \in \binom{[K]}{k} \,\middle|\, (V_j)_{j\in J} \text{ not } \epsilon\text{-homogeneous} \right\}$ is indeed small. Let us define $\mathscr{I}_{\text{pure}} := [K] \setminus \mathscr{I}_{\text{mixed}}$ and

$$X_1 := \left\{ J \in \binom{[K]}{k} \,\middle|\, J \cap \mathscr{I}_{\text{mixed}} \neq \emptyset \right\}.$$
$$X_2 := \left\{ J \in \binom{\mathscr{I}_{\text{pure}}}{k} \,\middle|\, (V_j)_{j\in J} \text{ not } \epsilon\text{-homogeneous} \right\}.$$

Then $X \subseteq X_1 \cup X_2$ so it suffices to bound X_1 and X_2 in size.
Ad X_1. We observe

$$|X_1| \leq |\mathscr{I}_{\text{mixed}}| \cdot \binom{K-1}{k-1} \leq l \cdot \frac{k}{K}\binom{K}{k} = \frac{lk}{\left\lfloor \frac{8kl}{\epsilon} \right\rfloor}\binom{K}{k} \leq \frac{\epsilon}{4}\binom{K}{k}.$$

Ad X_2. Let us double count the size of

$$F := \bigcup_{J\in\binom{\mathscr{I}_{\text{pure}}}{k}} \text{Fringes}_{\mathscr{H}}((V_j)_{j\in J}) = \left\{ (p,p') \in \left(\bigtimes_{j\in J} V_j \right)^2 \,\middle|\, J \in \binom{\mathscr{I}_{\text{pure}}}{k},\ p \in \mathscr{E},\ p' \notin \mathscr{E},\ |p \cap p'| = k-1 \right\},$$

see Definition 78. Let us sketch the argument for the upper bound. Let us first define all the hyperedges that are spanned between the partition classes $(V_j)_{j\in\mathscr{I}_{\text{pure}}}$ by

$$\mathscr{E}_{\text{pure}} := \left\{ p \in \mathscr{E} \,\middle|\, \exists J \in \binom{\mathscr{I}_{\text{pure}}}{k} :\ p \in \bigtimes_{j\in J} V_j \right\}.$$

For $j \in \mathscr{I}_{\text{pure}}$ and two vertices $b, b' \in V_j$ the number of fringes $(p,p') \in F$ that fulfill $p\Delta p' = \{b, b'\}$ is exactly

$$|N_{\mathscr{E}_{\text{pure}}}(b)\, \Delta N_{\mathscr{E}_{\text{pure}}}(b')\,|,$$

where we interpreted the edges $\mathscr{E}_{\text{pure}}$ as a hypergraph. Thus, we can upper bound this number simply by $|\,(N_{\mathscr{H}}(b)\, \Delta N_{\mathscr{H}}(b'))\,| \leq 2\delta$, take also a look at the visualization in Figure 15. The squares in the neighborhood of b represent $(k-1)$-sets A such that $A \cup \{b\} \in \mathscr{E}$. Square A is filled black in case that $A \cup \{b\} \in \mathscr{E}_{\text{pure}}$.

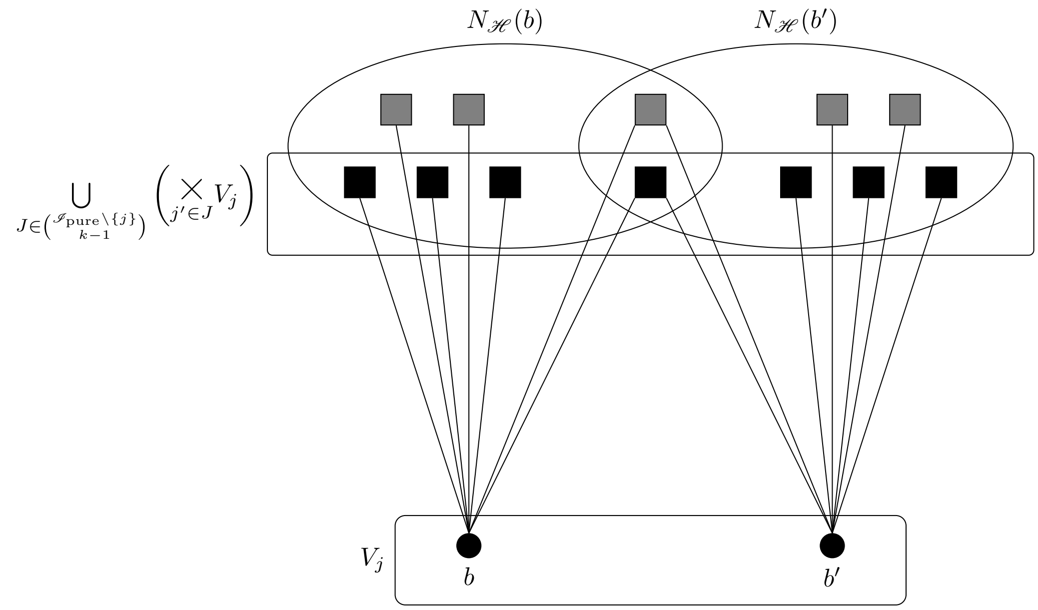

Figure 15: Visualization of the neighborhood of $b, b' \in V_j$ for $j \in \mathscr{I}_{\text{pure}}$.

We can lower bound the sum of the size of all Fringes corresponding to $J \in \binom{\mathscr{I}_{\text{pure}}}{k}$ by the sum of the size of all fringes corresponding to $J \in X_2$. Then we can apply our preparatory work in Lemma 22.

$$|F| = \sum_{J \in \binom{\mathscr{I}_{\text{pure}}}{k}} |\text{Fringes}_{\mathscr{H}}\left((V_j)_{\in J}\right)| \geq \sum_{J \in X_2} |\text{Fringes}_{\mathscr{H}}((V_j)_{j \in J})| \geq |X_2| \epsilon^2 \left\lfloor \frac{n}{K} \right\rfloor^{k+1}.$$

On the other hand, using the pairwise similarity of the neighborhoods of the vertices in each partition class in $(V_i)_{i \in \mathscr{I}_{\text{pure}}}$ as well as $\left\lfloor \frac{n}{K} \right\rfloor + 1 \leq 2 \left\lfloor \frac{n}{K} \right\rfloor$, we count

$$\begin{aligned}
|F| &= \sum_{j \in \mathscr{I}_{\text{pure}}} \left(\sum_{B \in \binom{V_j}{2}} \left(\sum_{J' \in \binom{\mathscr{I}_{\text{pure}}}{k-1}} \left| \left\{ (p, p') \in \text{Fringes}_{\mathscr{H}} \left((V_i)_{i \in J' \cup \{j\}} \right) \,\middle|\, p \Delta p' = B \right\} \right| \right) \right) \\
&\leq \sum_{j \in \mathscr{I}_{\text{pure}}} \left(\sum_{\{b, b'\} \in \binom{V_j}{2}} |N_{\mathscr{H}}(b) \Delta N_{\mathscr{H}}(b')| \right) \\
&< |\mathscr{I}_{\text{pure}}| \binom{\left\lfloor \frac{n}{K} \right\rfloor + 1}{2} 2\delta \leq 2K\delta \left\lfloor \frac{n}{K} \right\rfloor^2.
\end{aligned}$$

Now combining the lower and upper bound we obtain, using $\left\lfloor \frac{n}{K} \right\rfloor \geq \frac{n}{2K}$ and at last plugging in the Definition of δ'

$$\begin{aligned}
|X_2| \leq \frac{2}{\epsilon^2} K\delta \left\lfloor \frac{n}{K} \right\rfloor^{-(k-1)} &\leq \frac{2}{\epsilon^2} K\delta \left(\frac{n}{2K} \right)^{-(k-1)} \\
&= \frac{(2K)^k}{\epsilon^2} \left\lceil \delta' \binom{n}{k-1} \right\rceil n^{-(k-1)} \\
&\leq \frac{(2K)^k}{\epsilon^2} \delta' \left(\frac{ne}{k-1} \right)^{k-1} n^{-(k-1)} \\
&\leq \delta' \frac{2^k e^{k-1}}{\epsilon^2} K \binom{K}{k-1} \\
&= \delta' \frac{2^k e^{k-1}}{\epsilon^2} \frac{k \cdot K}{K-k+1} \binom{K}{k} \leq \delta' \frac{2^{k+1} e^{k-1} k}{\epsilon^2} \binom{K}{k} = \frac{\epsilon}{e} \binom{K}{k}.
\end{aligned}$$

Using this we conclude $|X| \leq \frac{\epsilon}{4} + \frac{\epsilon}{e} < \epsilon$. This closes the proof of Theorem 27. □

Errata 1. In the proof of an analogue of Lemma 22 presented in [20] there has been a small error. There the authors claimed that

Lemma 22*(flawed)**.** $\forall \epsilon \in \left(0, \frac{1}{2}\right) : (W_j)_{j\in[k]}$ not ϵ-homogeneous in $\mathscr{H} \implies |\text{Fringes}_{\mathscr{H}}\left((W_j)_{j\in[k]}\right)| \geq \epsilon m^{k+1}$.
However, this is false as the following small counterexample with $m = 3,\ k = 2$ shows. Let $A = \{a_1, a_2, a_3\},\ B = \{b_1, b_2, b_3\}$ be two disjoint sets of cardinality 3 each. Define the bipartite graph

$$H := (A \cup B,\ \{\{a_1, b_1\}, \{a_1, b_2\}, \{a_1, b_3\}, \{a_2, b_1\}, \{a_3, b_1\}\}).$$

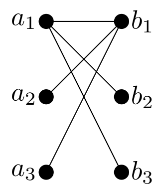

Figure 16: A rendering of the graph H.

Then, since $\frac{\|H\|}{|W_1||W_2|} = \frac{5}{9}$ we have that the partition of H is not $\frac{4}{9}$-homogeneous. Let us list

$$\begin{aligned}\text{Fringes}_H\left((W_j)_{j\in[2]}\right) = \Big\{ & (\{a_2, b_1\},\ \{a_2, b_2\}), (\{a_2, b_1\},\ \{a_2, b_3\}), \\ & (\{a_3, b_1\},\ \{a_3, b_2\}), (\{a_3, b_1\},\ \{a_3, b_3\}), \\ & (\{a_1, b_2\},\ \{a_2, b_2\}), (\{a_1, b_2\},\ \{a_3, b_2\}), \\ & (\{a_1, b_3\},\ \{a_2, b_3\}), (\{a_1, b_3\},\ \{a_3, b_3\})\Big\}.\end{aligned}$$

Thus, the number of Fringes is 8. However, the Claim in the paper would give the lower bound $\frac{4}{9}m^{k+1} = 12$. Our Lemma Claims the lower bound $\left(\frac{4}{9}\right)^2 m^{k+1} = 5\frac{1}{3}$, which however already is not sharp.

5.3 Proof of the Erdős-Hajnal conjecture for graphs of bounded VC dimension

In a first step for a bipartite graph H, $\epsilon > 0$ and any graph G that does not contain a biinduced copy of H let us use the ϵ-regular partition guaranteed by the Ultra Strong Regularity lemma for graphs of bounded VC dimension to obtain a sequence of pairwise ϵ-restricted vertex subsets.

Lemma 23 (Nguyen, Scott and Seymour [40]; Fox, Pach and Suk [20])**.** For any bipartite graph H there is $b = b(H) \in \mathbb{N}$ such that for any $\epsilon \in \left(0, \frac{1}{2}\right)$, $n \in \mathbb{N}$ and any graph $G \in \text{Free}\,(n, H\text{-biind})$ there are $l, m \in \mathbb{N}$ with $l \geq \frac{1}{\epsilon}$ and $m \geq n\epsilon^b$ such that there exists a sequence of pairwise disjoint vertex subsets $(B_j)_{j\in[l]} \subseteq \binom{[n]}{m}$ fulfilling that for any distinct $i, j \in [l]$ the set B_i is ϵ-restricted towards B_j.

Proof of Lemma 23. First we remark that in case $n \leq \left(\frac{1}{\epsilon}\right)^b$ the statement is trivial since we can choose $m = 1$. Thus, we may assume that $n \geq \left(\frac{1}{n}\right)^b$. By Observation 21 there is $d \in \mathbb{N}$ such that any graph that does not contain a biinduced copy of H has VC dimension at most d. The Ultra Strong Regularity lemma, Theorem 27, with parameter $\tilde{k} = 2$ and $\tilde{d} = d$ yields some constant $c = c(d) > 0$ such that for $\tilde{\epsilon} := \frac{\epsilon^2}{6}$ there is $K \in \mathbb{N}$ with $\frac{8}{\tilde{\epsilon}} \leq K \leq c\left(\frac{1}{\tilde{\epsilon}}\right)^{3d+1}$ and an equitable partition $(V_j)_{j\in[K]}$ of $[n]$ such that at most $\tilde{\epsilon}\binom{K}{2}$ of the pairs of partition classes are not $\tilde{\epsilon}$-homogeneous.
The idea of the proof is to use Turáns Theorem to find a large set of partition classes, such that any pair among them is $\tilde{\epsilon}$-homogeneous. In a second step we are going to trim off vertices from these classes such that the leftover classes are pairwise ϵ-restricted.

Since $\epsilon < \frac{1}{2}$ we observe that for some integer $b = b(c,d)$ that only depends on d

$$K \leq c\left(\frac{6}{\epsilon^2}\right)^{3d+1} \leq \left(\frac{1}{\epsilon}\right)^{b-1}.$$

Consider the help graph

$$J := \left([K],\ \left\{ \{i,j\} \in \binom{[K]}{2} \ \middle|\ (V_i,\ V_j)\ \tilde{\epsilon}\text{-homogeneous} \right\}\right).$$

We remark that $\frac{1}{K} \leq \frac{\tilde{\epsilon}}{8} \leq \tilde{\epsilon}$ and calculate

$$\|J\| \geq (1-\tilde{\epsilon})\binom{K}{2} = \left(1-\tilde{\epsilon}-\frac{1}{K}\right)\frac{K^2}{2} \geq (1-2\tilde{\epsilon})\,\frac{K^2}{2}.$$

Hence, Corollary 1 yields $\omega(J) \geq \frac{1}{2\tilde{\epsilon}} = \frac{3}{\epsilon^2}$. Thus, for $l := \left\lceil \frac{1}{\epsilon} \right\rceil$ we may assume that $[l]$ is a clique in J.
Fix $\alpha := \frac{3}{\epsilon}$. For distinct $i,j \in [l]$ let us define a vertex in V_j to be (j,i)-*bad* in case that its behavior towards V_i does not reflect the behavior of V_j towards V_i.

$$V_{\text{bad}}^{(i)}(j) := \left\{ v \in V_j \ \middle|\ \left(\frac{\|V_i,V_j\|}{|V_i||V_j|} \leq \tilde{\epsilon} \text{ and } \deg_{V_i}(v) \geq \alpha\tilde{\epsilon}|V_i|\right) \text{ or } \left(\frac{\|V_i,V_j\|}{|V_i||V_j|} \geq 1-\tilde{\epsilon} \text{ and } \deg_{V_i}(v) \leq (1-\alpha\tilde{\epsilon})|V_i|\right) \right\}.$$

Observe that in case $\frac{\|V_i,V_j\|}{|V_i||V_j|} \leq \tilde{\epsilon}$ we have

$$|V_{\text{bad}}^{(i)}(j)| \cdot \alpha\tilde{\epsilon}|V_i| \leq \|V_i,V_j\| \leq \tilde{\epsilon}|V_i||V_j|.$$

Using this we conclude

$$|V_{\text{bad}}^{(i)}(j)| \leq \frac{|V_j|}{\alpha}.$$

Similarly, by considering the complement graph in case that $\frac{\|V_i,V_j\|}{|V_i||V_j|} \geq 1-\tilde{\epsilon}$ it follows that $|V_{\text{bad}}^{(i)}(j)| \leq \frac{|V_j|}{\alpha}$. Fix $j \in [l]$. We define

$$V_{\text{bad}}(j) := \bigcup_{i \in [l]\setminus\{j\}} V_{\text{bad}}^{(i)}(j).$$

Let us bound the number of bad vertices in V_j.

$$|V_{\text{bad}}(j)| \leq \frac{l-1}{\alpha}|V_j| \leq \frac{1}{\alpha\epsilon}|V_j| = \frac{|V_j|}{3}.$$

We observe $|V_j| \in \left\{\left\lfloor \frac{n}{K}\right\rfloor, \left\lceil \frac{n}{K}\right\rceil\right\}$. Let us choose $X_j \subseteq V_j$ such that $V_{\text{bad}}(j) \subseteq X_j$ and

$$|V_j| - |X_j| = m := \left\lceil \frac{n}{2K} \right\rceil.$$

This is always possible since $|V_{\text{bad}}(j)| \leq \left\lfloor \frac{|V_j|}{3} \right\rfloor$. Finally, we set

$$B_j := V_j \setminus X_j.$$

Using $K \leq \left(\frac{1}{\epsilon}\right)^{b-1}$ we deduce

$$m = \left\lceil \frac{n}{2K} \right\rceil \geq \frac{n\epsilon^{b-1}}{2} \geq n\epsilon^b.$$

Let us close the proof by showing that for all distinct $i,j \in [l]$ the set B_i is ϵ-restricted towards B_j. First, using

$\forall \mu \in \mathbb{R}_0 : \lceil \mu \rceil \leq 2 \left\lceil \frac{\mu}{2} \right\rceil$, we observe that

$$|X_j| = |V_j| - \left\lceil \frac{n}{2K} \right\rceil \leq \left\lceil \frac{n}{K} \right\rceil - \left\lceil \frac{n}{2K} \right\rceil \leq \left\lceil \frac{n}{2K} \right\rceil = m.$$

case $\frac{\|V_i, V_j\|}{|V_i||V_j|} \leq \epsilon$. In this case for $v \in B_i$, using v is not (i,j)-bad, we calculate

$$\deg_{B_j}(v) \leq \deg_{V_j}(v) \leq \alpha\tilde{\epsilon}|V_j| = \alpha\tilde{\epsilon}\,(m + |X_j|) \leq 2\alpha\tilde{\epsilon}m = \epsilon m.$$

case $\frac{\|V_i, V_j\|}{|V_i||V_j|} \geq 1 - \epsilon$. In this case for $v \in B_i$ we calculate

$$\begin{aligned} \deg_{B_j}(v) \geq \deg_{V_j}(v) - |X_j| &\geq (1 - \alpha\tilde{\epsilon})|V_i| - |X_j| \\ &= (1 - \alpha\tilde{\epsilon})(m + |X_j|) - |X_j| \\ &= (1 - \alpha\tilde{\epsilon})m - \alpha\tilde{\epsilon}|X_j| \\ &\geq (1 - 2\alpha\tilde{\epsilon})m = (1 - \epsilon)m. \end{aligned}$$

This completes the proof of Lemma 23. □

The following Theorem is a consequence of the Regularity lemma for graphs. It guarantees copies of graphs in dense graphs that's edges are distributed evenly enough.

Theorem 28 (Rödl [41])**.** $\forall k \in \mathbb{N}$ and $\alpha, \beta \in (0,1)$ with $\alpha < \beta$ there is $\gamma \in (0,1)$ and $N_0 \in \mathbb{N}$ such that every graph G on at least N_0 vertices that fulfills

$$\forall U \subseteq V(G) : \ |U| \geq \gamma|G| \implies \frac{\|G[U]\|}{\binom{|U|}{2}} \in (\alpha, \beta)$$

contains all graphs on k vertices as subgraph.

We are not going to give a proof of Theorem 28. However, we deduce the following Corollary that is going to be applied in the proof of the Erdős-Hajnal conjecture for graphs of bounded VC dimension. It guarantees large ϵ-restricted induced subgraphs in the graphs of any proper hereditary graph property.

Corollary 5. For every proper hereditary graph property $\mathscr{C}$, see Definition 10, and $\epsilon \in \left(0, \frac{1}{2}\right)$ there is $\psi > 0$ such that

$$\forall G \in \mathscr{C}\ \exists H \underset{ind}{\subseteq} G : \ |H| \geq \psi|G| \text{ and } \Delta_\delta(H) \leq \epsilon|H|.$$

Proof of Corollary 5. Since $\mathscr{C}$ is proper there is a graph F such that $F \notin \mathscr{C}$. Now an application of Rödls Theorem 28 with parameter $\tilde{a} = \frac{\epsilon}{2}$, $\tilde{\beta} = 1 - \frac{\epsilon}{2}$ and $\tilde{k} := |F|$ yields $\gamma > 0$ and N_0 with the claimed properties. Let us set $\psi := \min\left\{\frac{1}{N_0}, \frac{\epsilon\gamma}{2}\right\}$. Let $G \in \mathscr{C}$.
case 1 $|G| \leq N_0$**.** In this case $\psi|G| \leq 1$ so the Claim on G is trivial.
case 2 $|G| > N_0$**.** In this case there exists $U \subseteq V(G)$ such that $|U| \geq \gamma|G|$ and $\frac{\|G[U]\|}{\binom{|U|}{2}} \in \left[0, \frac{\epsilon}{2}\right] \cup \left[1 - \frac{\epsilon}{2}, 1\right]$ since otherwise G would contain F as an induced subgraph by Rödls Theorem.
case 2.1 $\|G[U]\| \geq (1 - \frac{\epsilon}{2})\binom{|U|}{2}$**.** In this case $\text{avdeg}(G[U]) \geq (1 - \frac{\epsilon}{2})(|U| - 1)$. A standard result yields an induced subgraph $H \underset{ind}{\subseteq} G[U]$ with $\delta(H) \geq \text{avdeg}(G[U])$. Using $\frac{\epsilon}{2}|U| \geq \frac{\epsilon\gamma}{2}|G| \geq \frac{\epsilon\gamma}{2}N_0 \geq 1$ we deduce

$$\delta(H) \geq \text{avdeg}(G[U]) \geq (1 - \frac{\epsilon}{2})(|U| - 1) \geq (1 - \epsilon)|U| \geq (1 - \epsilon)|H|.$$

Observe further that

$$|H| \geq \delta(H) \geq (1 - \epsilon)|U| \geq (1 - \epsilon)\gamma|G| \geq \frac{\gamma}{2}|G| \geq \psi|G|.$$

This shows that H is the claimed ϵ-restricted induced subgraph in G.

case 2.2 $\|G[U]\| \leq \frac{\epsilon}{2}\binom{|U|}{2}$**.** $\|\overline{G[U]}\| \geq (1-\frac{\epsilon}{2})\binom{|U|}{2}$ so analogously to the previous case we find $H \underset{ind}{\subseteq} G$ with $|H| \geq \psi|G|$ and $\delta(\overline{H}) \geq (1-\epsilon)|H|$ translating into $\Delta(H) = |H| - 1 - \delta(H) \leq |H| - \delta(H) \leq \epsilon|H|$.
We conclude that in both cases we found $H \underset{ind}{\subseteq} G$ with $\Delta_\delta(H) \leq \epsilon|H|$ and $|H| \geq \psi|G|$. This completes the proof of Corollary 5. □

Now we are well-prepared to prove the Erdős-Hajnal conjecture for graphs with bounded VC dimension. The arguments are similar to the ones given in [40], however we changed the structure of the proof into a more linear form.

Proof of Theorem 26. First we are going to prove the following statement by a double induction on m and n.

(∗) For any $m \in \mathbb{N}$ and any bipartite graph H on m vertices there is a constant $C = C(H) > 0$ such that for all $n \in \mathbb{N}$ and any graph $G \in \text{Free}(n, H\text{-biind})$ there is an induced subgraph $\tilde{G} \underset{ind}{\subseteq} G$ with $|\tilde{G}| \geq n^C$ that is a Cograph.

We remark that Theorem 26 is an immediate consequence of (∗) since by Observation 20

$$\forall \tilde{G} \in \mathscr{C}_{\text{Cograph}} :\ \alpha \vee \omega\left(\tilde{G}\right) \geq \sqrt{|\tilde{G}|}.$$

<u>**Induction on m.**</u>
base $m = 2$. We know $H \in \left\{K_2, \overline{K_2}\right\}$ so the graphs in $\text{Free}(H\text{-biind})$ are either complete or empty. This means all graphs in $\text{Free}(H\text{-biind})$ are Cographs themselves.
step $m \geq 3$. Let $H = (A \cup B,\ E)$ be a bipartite graph on m vertices and let $v \in A$. Define $H' := H - v$. Let us prepare some constants.
By induction on m there is $a = a(H') \in \mathbb{N}$, without loss of generality $a > \max\left\{8, \frac{\log_2(|B|)}{128}\right\}$, such that for any graph G

$$H' \underset{biind}{\nsubseteq} G \implies \alpha \vee \omega(G) \geq n^{\frac{1}{a}}.$$
$$\overline{\overline{H'}} \underset{biind}{\nsubseteq} G \implies \alpha \vee \omega(G) \geq n^{\frac{1}{a}}.$$

Let $b = b(H) \in \mathbb{N}$ be given by Lemma 23 and fix $c := 2^{-8}$. Corollary 5 yields $t \in \mathbb{N}$, without loss of generality $t \geq 5ba$, such that for all graphs G

$$H \underset{biind}{\nsubseteq} G \implies \exists G'' \underset{ind}{\subseteq} G \text{ such that } |G''| \geq c^t n \text{ and } \Delta_\delta(G'') \leq c|G'|.$$

With this let us fix $C' := \frac{1}{4at}$.
<u>**Induction on n.**</u>
In the sequel we are going to show the following by induction on n.

$$(\star)\ \forall n \in \mathbb{N},\ G \in \text{Free}(n, H\text{-biind})\ \exists \tilde{G} \underset{ind}{\subseteq} G :\ \tilde{G} \in \mathscr{C}_{\text{Cograph}} \text{ and } |\tilde{G}| \geq n^{C'}.$$

base $n \in \left[2^{4at}\right]$. The Claim $(\star)$ is trivial since $n^{C'} \leq 2$.
step $n > 2^{4at}$. Let $G \in \text{Free}(n, H\text{-biind})$. Then by the Definition of t there is an induced subgraph $G'' \underset{ind}{\subseteq} G$ such that

$$|G''| \geq c^t n \text{ and } \Delta_\delta(G'') \leq c|G''|.$$

Fix $x := n^{-\frac{1}{2at-1}}$ and let $y \in [x^a, c]$ be minimal such that there is an induced subgraph $G' \underset{ind}{\subseteq} G$ with

$$|G'| \geq y^{2t} n \text{ and } \Delta_\delta(G') \leq y|G'|.$$

We remark that this minimum is well-defined since there are only finitely many induced subgraphs of G. Furthermore, we need the lower bound $y \geq x^a$ since otherwise $y = 0$ would be a trivial solution.

The idea of the proof is to find a long sequence of pairwise disjoint and large vertex sets $(B_j)_{j\in[l]}$ in G', such that, in case that $\Delta(G') < y|G'|$ all pairs of the sequence are anticomplete and in case $\Delta\left(\overline{G'}\right) \leq y|G'|$ all pairs of the sequence are complete. Then by induction on n we can find large Cographs on the sets of the sequence and their disjoint sum or disjoint product respectively yields a Cograph large enough for our Claim.
Let us assume for now that $\Delta(G') < y|G'|$. We remark that one can deal with the case $\Delta\left(\overline{G'}\right) < y|G'|$ analogously since $\overline{\overline{H}} \underset{biind}{\nsubseteq} \overline{G}$. We find the sequence $(B_j)_{j\in[l]} \subseteq V(G')$ iteratively, where we start with a vertex set $E_0 := V(G')$. Let us assume that we have found a sequence of length $l' \in \mathbb{N}_0$ of pairwise disjoint vertex subsets $\left((B_j)_{j\in[l']}, E_{l'}\right) \subseteq V(G')$ where

(i) $\forall j \in [l'] : \ |B_j| \geq y^{4t+\frac{1}{2}}|G'|$.

(ii) $|E_{l'}| \geq \left(1 - 2y^{\frac{1}{2}}\right)^{l'} |G'|$.

(iii) $\forall \{i,j\} \in \binom{[l']}{2}$ the pairs $\{B_i, B_j\}$ and $\{B_i, E_{l'}\}$ are anticomplete.

In case that l' is small we know that $|E_{l'}|$ is large and we apply Lemma 23 to find a long sequence of pairwise disjoint and pairwise ϵ-restricted vertex subsets $(C_j)_{j\in[p]} \subseteq E_{l'}$ of size at least $y^{4t+\frac{1}{2}}|G'|$ each.
Our goal is to augment the sequence $(B_j)_{j\in[l']}$ by one set C_i, $i \in [p]$, meaning that we require there are many vertices in $E_{l'}$ that send no edges towards C_i.
For $j \in [p]$ we can bound the number of vertices that are complete towards C_j by the maximal degree condition on G'. It is left to show that there exists $i \in [p]$ with few vertices that send both edges and non-edges towards C_i, we are going to call this kind of vertices *mixed vertices towards C_i*. Using a double counting argument we show the existence of such a set C_i by showing that any vertex in $E_{l'}$ is mixed towards only a few sets. The latter turns out to be the heart of the proof of Theorem 26.
Let $w \in E_{l'}$ and assume for a contradiction that the number r of sets that w is mixed towards is very large, take a look at Figure 17. There the dashed blue lines indicate that a pair of vertex sets is anticomplete. Furthermore, the red lines indicate that in between a pair of vertex sets there are many edges. By reordering the sets we may assume that w is mixed towards the sets $(C_j)_{j\in[r]}$. We are going to consider a help graph J where the vertices correspond to the sets $(C_j)_{j\in[r]}$ and where two partition classes are adjacent in case that there are many edges in between them. Remember that $(C_j)_{j\in[r]}$ is a sequence of pairwise ϵ-restricted sets.
In case that $H' \underset{biind}{\subseteq} J$ we are going to argue that we can find a biinduced copy of H in G, where v is embedded as w. A contradiction. In case that $H' \underset{biind}{\nsubseteq} J$ we are going to apply induction on n and find a large homogeneous vertex set I inside J. This either means that the vertices in the sets corresponding to the elements in I send many many or very few edges towards the other sets corresponding to the elements of I. In Figure 17 the whole graph J is a complete graph. In either case the induced graph of G on the union of the sets corresponding to the elements in I is going to be an y'-restricted graph of size at least $(y')^{2t}n$ for some $y' \in [y^a, y)$. Now in case that $y' \geq x^a$ this would be a contradiction to the minimality of y. However, $y' \geq x^a$ follows from $y \geq y^a$ and the following Claim.

Claim 18. We may assume that $y > x$.

Proof of Claim 18. Let us consider the case $y \in [x^a, x]$. We want to show that $\alpha \vee \omega(G) \geq n^{C'}$, which yields the Claim of $(*)$ since empty and complete graphs are Cographs.
By the Definition of y we know that there exists an induced subgraph $G' \underset{ind}{\subseteq} G$ with $\Delta_\delta(G') \leq y|G'|$ and $|G'| \geq y^{2t}n$. Observe that

$$|G'| \geq y^{2t}n \geq x^{2at}n = n^{1-\frac{2at}{2at-1}} = \frac{1}{x}.$$

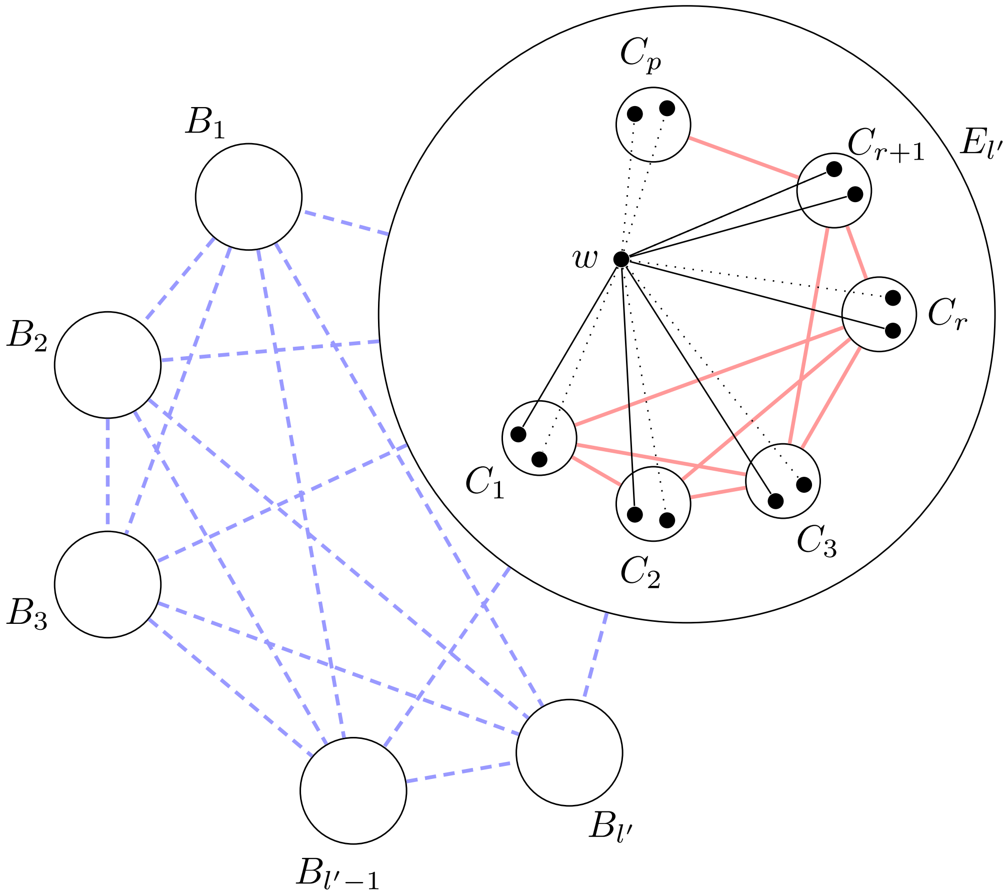

Figure 17: Situation of vertex $w \in E_{l'}$.

case $\delta(G') \geq (1-y)|G'|$. We have

$$\|G'\| \geq (1-y)\frac{|G'|^2}{2} \geq (1-x)\frac{|G'|^2}{2}$$

so Corollary 1 yields that $\omega(G') \geq \frac{1}{x}$.
case $\Delta(G') \leq y|G'|$. We know that $x|G'| \geq xyn \geq x^{a+1}n = n^{1-\frac{a+1}{2at-1}} \geq 1$ which implies $\frac{1}{|G'|} \leq x$. Using this we calculate

$$\delta\left(\overline{G'}\right) \geq |G'|(1-y) - 1 = |G'|\left(1 - y - \frac{1}{|G'|}\right) \geq |G'|(1-2x)$$

and again Corollary 1 yields that $\alpha(G') = \omega\left(\overline{G}\right) \geq \frac{1}{2x}$.
In both cases, using the fact that $2 \leq n^{\frac{1}{4at}}$, we conclude

$$\alpha \vee \omega(G') \geq \frac{1}{2x} \geq n^{\frac{1}{2at-1} - \frac{1}{4at}} \geq n^{\frac{1}{2at} - \frac{1}{4at}} = n^{\frac{1}{4at}} = n^{C'}.$$

This completes the proof of Claim 18. □

Claim 19. There is $l \geq y^{-\frac{1}{4}}$ and a sequence of pairwise disjoint vertex subsets $(B_j)_{j\in[l]} \subseteq V(G')$, each of size greater than $y^{4t+\frac{1}{2}}|G'|$, such that $\forall \{i,j\} \in \binom{[l]}{2}$ the pair $\{X_i, X_j\}$ is homogeneous, see Definition 74.

Let us assume that we had shown Claim 19. Then for each $i \in [l]$ the induction assumption regarding $(\star)$

applied to $G'[B_i]$ yields a Cograph $\tilde{G}_i \underset{ind}{\subseteq} G'[B_i]$ with $|\tilde{G}_i| \geq |B_i|^{C'}$. Observe that

$$\tilde{G} \coloneqq G\left[\bigcup_{i\in[l]} V\left(\tilde{G}_i\right)\right]$$

is a Cograph. Using $|B_i| \geq y^{4t+\frac{1}{2}}|G'| \geq y^{4t+\frac{1}{2}+2t}n$ and $l \geq y^{-\frac{1}{4}}$ let us calculate

$$|\tilde{G}| \geq l\left(y^{6t+\frac{1}{2}}n\right)^{C'} \geq y^{-\frac{1}{4}+C'\left(6t+\frac{1}{2}\right)}n^{C'} \geq n^{C'},$$

where we used $C'\left(6t+\frac{1}{2}\right) = \frac{6t+\frac{1}{2}}{4at} < \frac{2}{a} \leq \frac{1}{4}$ in the last inequality. This completes the proof of $(\star)$ and therefore the proof of Theorem 26. □

It is left to show Claim 19.

Proof of Claim 19. We prove in case $\Delta(G') \leq y|G'|$ that there is a long sequence of pairwise disjoint large subsets of $V(G')$ such that all pairs of sets in the sequence are anticomplete. Notice that in case $\Delta\left(\overline{G'}\right) \leq y|G'|$ since $\overline{\overline{H}} \underset{biind}{\nsubseteq} \overline{G'}$ one can analogously find a similar sequence such that the all pairs of sets in the sequence are complete.

Let $l' \in \mathbb{N}$ be maximal such that there is a sequence of pairwise disjoint vertex sets $(B_j)_{j\in[l']} \subseteq V(G')$ such that

(i) $\forall j \in [l'] :\ |B_j| \geq y^{4t+\frac{1}{2}}|G'|$.

(ii) $|B_{l'}| \geq \left(1-2y^{\frac{1}{2}}\right)^{l'}|G'|$.

(iii) $\forall \{i,j\} \in \binom{[l']}{2}$ the pair $\{B_i, B_j\}$ is anticomplete.

Notice that this is well-defined since for $l' = 1$ the requirements are trivial.

Assume $l' < y^{-\frac{1}{4}}$. In this case with $2y^{\frac{1}{4}} + y^{\frac{1}{2}} \leq 3y^{\frac{1}{4}} \leq 3c^{\frac{1}{4}} = \frac{3}{4} \leq 1$ it follows that

$$|B_{l'}| \geq \left(1-2y^{\frac{1}{2}}\right)^{l'}|G'| \geq \left(1-l'\cdot 2y^{\frac{1}{2}}\right)|G'| > \left(1-2y^{\frac{1}{4}}\right)|G'| \geq y^{\frac{1}{2}}|G'| \tag{6}$$

where in the second inequality we used Bernoulli's inequality, see Observation 3. We deduce that $G[B_{l'}]$ is $y^{\frac{1}{2}}$-restricted.

$$\Delta\left(G'[B_{l'}]\right) \leq \Delta(G') \leq y|G'| \leq y^{\frac{1}{2}}|B_{l'}|. \tag{7}$$

Claim 20. There are disjoint subsets $X, Y \subseteq B_{l'}$ such that $|X| \geq y^{4t}|B_{l'}|$ and $|Y| \geq \left(1-2y^{\frac{1}{2}}\right)|B_{l'}|$ and the pair $\{X, Y\}$ is anticomplete.

Let us assume for now that Claim 20 holds. Then it yields $X, Y \subseteq B_{l'}$ with the given properties. Using (6) we calculate

$$|X| \geq y^{4t}|B_{j'}| \geq y^{4t+\frac{1}{2}}|G'|.$$

Furthermore,

$$|Y| \geq \left(1-2y^{\frac{1}{2}}\right)|B_{l'}| \geq \left(1-2y^{\frac{1}{2}}\right)^{l'+1}|G'|.$$

This however is a contradiction to the maximality of l'. Thus, the contradiction argument showed $l' \geq y^{\frac{1}{4}}$ which completes the proof of Claim 19. □

It is left to prove Claim 20.

Proof of Claim 20. Define $\epsilon := y^{\frac{4t}{b}}$. An application of Lemma 23 yields $p, s \in \mathbb{N}$ with $p \geq \epsilon^{-1} = y^{-\frac{4t}{b}}$ and $s \geq \epsilon^b |B_{l'}| = y^{4t}|B_{l'}|$ and a sequence of pairwise disjoint subsets $(C_j)_{j\in[p]} \in \binom{B_{l'}}{s}$ such that for all distinct $i, j \in [p]$ the set C_i is ϵ-restricted towards C_j.
Denote

$$D := B_{l'} \setminus \left(\bigcup_{j\in[p]} C_j \right).$$

For $i \in [p]$ let us partition D according to how the vertices in D interact with C_i.

$$V_{\text{complete}}(i) := \{ w \in D \mid C_i \subseteq N_{G'}(w) \}$$
$$V_{\text{anticomplete}}(i) := \{ w \in D \mid C_i \cap N_{G'}(w) = \emptyset \}$$
$$V_{\text{crossing}}(i) := \{ w \in D \mid \emptyset \neq C_i \cap N_{G'}(w) \neq C_i \}$$

Our goal in order to prove Claim 20 is to find $i \in [p]$ such that $V_{\text{anticomplete}}(i)$ is huge. For $w \in D$ let us define

$$\mathscr{I}_{\text{crossing}}(w) := \{ i \in [p] \mid \emptyset \neq N_{B_i}(w) \neq B_i \}.$$

Claim 21. $\forall w \in D : \ |\mathscr{I}_{\text{crossing}}(w)| < py.$

Proof of Claim 21. Assume for a contradiction that there is some vertex $w \in D$ fulfilling $|\mathscr{I}_{\text{crossing}}(w)| \geq py$. Without loss of generality we may assume that there exists $r \in \mathbb{N}$ with $r \geq py$ such that $\mathscr{I}_{\text{crossing}}(w) = [r]$. Construct the help graph

$$J := \left([r], \left\{ \{i,j\} \in \binom{[r]}{2} \,\middle|\, \delta\left(G\left[B_i, B_j\right]\right) > (1-\epsilon)s \right\} \right).$$

First we want to remark that for $\{i,j\} \in \binom{[r]}{2} \setminus E(J)$ we know that $\Delta\left(G\left[B_i, B_j\right]\right) < \epsilon s$.
case $H' \underset{biind}{\subseteq} J$**.** In this case we may show that $H \underset{biind}{\subseteq} G$, a contradiction.
Recall that $H = (A \cup B,\ F)$, $v \in A$ and $H' = H - v$. Let $(j_u)_{u\in V(H')} \subseteq [r]$ be the embedding of the vertices of H' into the vertices of J that corresponds to the biinduced copy of H' in J.
Firstly let us embed v as $w_u := w$. Secondly let us care for the vertices in B, where we make use of the assumption that w has neighbors and non-neighbors in all the sets $\{ C_{j_u} \mid u \in B \}$. Simply embed $u \in B$ as w_u in C_{j_u} such that $\{w, w_u\} \in E(G)$ if and only if $\{v, u\} \in E(H)$.
At last, we may embed the vertices $u \in A \setminus \{v\}$. Let us define candidate sets for vertices we can choose to embed u.

$$V_{\text{candidate}}(u) := \{ w_u \in C_{j_u} \mid \forall u' \in B : \{w_u, w_{u'}\} \in E(G) \Longleftrightarrow \{u, u'\} \in E(H) \}.$$

Since all pairs of sets in $(C_{j_u})_{u\in V(H')}$ are ϵ-restricted towards each other by the Definition of J we can bound

$$|V_{\text{candidate}}(u)| \geq |C_{j_u}| - \sum_{u'\in N_B(u)} |\{ w_u \in C_{j_u} \mid \{w_u, w_{u'}\} \notin E(H) \}|$$
$$- \sum_{u'\in B\setminus N_B(u)} |\{ w_u \in C_{j_u} \mid \{w_u, w_{u'}\} \in E(H) \}| \geq s\left(1 - |B|\epsilon\right).$$

Using our assumptions $t \geq 4ab$ and $a > \frac{\log_2(|B|)}{128}$ we calculate

$$\epsilon = y^{\frac{4t}{b}} \leq y^{16a} \leq c^{16a} = 2^{-128a} < 2^{-\log_2(|B|)} = \frac{1}{|B|}.$$

This argument shows that $s\left(1 - |B|\epsilon\right) > 0$ and we deduce that for any $u \in A\setminus\{v\}$ there is at least one candidate.

It is easy to check that indeed

$$G\left[\{\, w_u \mid u \in A\,\}, \{\, w_u \mid u \in B\,\}\right] = H.$$

case $H' \underset{biind}{\not\subseteq} J$. In this case we want to show that y has not been minimal, a contradiction. By the induction hypothesis regarding ($\star$) we find a Cograph $\tilde{J} \underset{ind}{\subseteq} J$ with $|\tilde{J}| \geq |J|^{\frac{1}{a}}$. Observe that $|J| = r \geq py \geq y^{1-\frac{4t}{b}}$. Furthermore, by Observation 20 we find a homogeneous subset $I \subseteq V\left(\tilde{J}\right)$ of size $|I| \geq |\tilde{J}|^{\frac{1}{2}}$. Using $t \geq 4ab$, we conclude

$$|I| \geq |\tilde{J}|^{\frac{1}{2}} \geq |J|^{\frac{1}{2a}} \geq y^{\frac{b-4t}{2ab}} \geq y^{\frac{b-16ab}{2ab}} \geq y^{-6}.$$

Consider the graph

$$\tilde{C} \coloneqq G\left[\bigcup_{j\in I} V\left(C_j\right)\right] = \sum_{j\in I} C_j,$$

see Definition 13. Using $s \geq y^{4t+\frac{1}{2}}|G'|$ and $|G'| \geq y^{2t}n$ we calculate

$$|\tilde{C}| = |I|s \geq y^{-6}s \geq y^{-6+4t+\frac{1}{2}}|G'| \geq y^{6t}n.$$

Using this and defining $y' \coloneqq y^3$ we deduce

$$|\tilde{C}| \geq y^{6t}n = \left(y'\right)^{2t} n.$$

In case that I is independent in J we arrive at

$$\Delta\left(G\left[\tilde{C}\right]\right) \leq s + |I|s\epsilon = |\tilde{C}|\left(\frac{1}{|I|} + \epsilon\right) \leq |\tilde{C}|\left(y^6 + y^{\frac{4t}{b}}\right) \leq |\tilde{C}|2y^6 \leq |\tilde{C}|y^{6-\frac{1}{8}} \leq |\tilde{C}|y^3 \leq y'|\tilde{C}|,$$

where we used that $2 \leq y^{-\frac{1}{8}}$ since $y \leq c = 2^{-8}$.
In case that I is a clique in J it follows analogously that $\delta\left(G\left[\tilde{C}\right]\right) \geq |\tilde{C}|\left(1 - y'\right)$. We conclude that

$$\Delta_\delta\left(G\left[\tilde{C}\right]\right) \leq y'|\tilde{C}|.$$

Using Claim 18 we deduce that $y' = y^3 \geq x^3$ so

$$y' \in [x^3, y).$$

Thus, we have shown that y has not been chosen minimal. This contradiction closes the proof of Claim 21. □

Using Claim 21 let us double count

$$p \cdot \min_{i\in[p]}|V_{\text{crossing}}(i)| \leq \sum_{i\in[p]}|V_{\text{crossing}}(i)| = \sum_{w\in D}|\mathscr{I}_{\text{crossing}}(w)| < |D| \cdot py.$$

Thus, there is $i \in [p]$ such that $|V_{\text{crossing}}(i)| < y|D|$. Using (7) let us roughly estimate $|V_{\text{complete}}(i)| \leq \Delta\left(G'\right) \leq y^{\frac{1}{2}}|B_{l'}|$. Finally, we set $X \coloneqq C_i$ and $Y \coloneqq V_{\text{anticomplete}}(i)$. Then $|X| \geq \epsilon^b|B_{j'}| = y^{4t}|B_{j'}|$ and

$$|Y| = |V_{\text{anticomplete}}(i)| = |B_{j'}| - |V_{\text{complete}}(i)| - |V_{\text{crossing}}(i)| \geq |B_{j'}|\left(1 - 2y^{\frac{1}{2}}\right)$$

which completes the proof of Claim 20. □

Now that we have proven the Erdős-Hajnal conjecture for graphs with bounded VC dimension we can present the proof of Theorem 25.

Proof of Theorem 25. First we remark that by Observation 35 it suffices to show that $\mathscr{C}$ fulfills the polynomial Rödl property. Let $d := \dim_{\mathrm{VC}}(\mathscr{C})$. Theorem 26 yields $\tilde{C} \in (0,1)$ such that

$$(*)\ \forall n \in \mathbb{N},\ F \in \mathrm{Free}\left(n, \mathrm{Incidence}\left(2^{[d+1]}\right)\text{-biind}\right) :\ \alpha \vee \omega(F) \geq n^{\tilde{C}}.$$

Let $\epsilon \in \left(0, \frac{1}{2}\right)$ and $G \in \mathscr{C}$. Let $b \in \mathbb{N}$ be given by Lemma 23 for $H = \mathrm{Incidence}\left(2^{[d+1]}\right)$. Let $q \in \mathbb{N}$ be minimal such that $\epsilon^q(d+1)2^{(d+1)} < 1$ and $2\epsilon^{q\tilde{C}} \leq \epsilon$.
Since $G \in \mathrm{Free}\left(|G|, \mathrm{Incidence}\left(2^{[d+1]}\right)\text{-biind}\right)$ Lemma 23 yields $b = b(H) \in \mathbb{N}$ such that for $\tilde{\epsilon} := \epsilon^q$ we find $l, m \in \mathbb{N}$ with $l \geq \frac{1}{\tilde{\epsilon}}$ and $m \geq \tilde{\epsilon}^b|G|$ and a sequence of pairwise disjoint subsets $(B_j)_{j\in[l]} \subseteq \binom{V(G)}{m}$ such that for any distinct $i, j \in [l]$ the set B_i is $\tilde{\epsilon}$-restricted towards B_j.
The idea of the proof is to apply $(*)$ on the following help graph.

$$J := \left([l],\ \left\{ \{i,j\} \in \binom{[l]}{2} \ \middle|\ \|B_i, B_j\| > (1-\tilde{\epsilon})m^2 \right\}\right).$$

We remark that for $\{i,j\} \in E(J)$ and any $v \in B_i$ we have $\deg_{B_j}(v) > (1-\tilde{\epsilon})m$ and for $\{i,j\} \in \binom{[l]}{2} \setminus E(J)$ and any $v \in B_i$ we have $\deg_{B_j}(v) < \tilde{\epsilon}m$.

Claim 22. $J \in \mathrm{Free}\left(l, \mathrm{Incidence}\left(2^{[d+1]}\right)\text{-biind}\right)$.

Proof of Claim 22. Assume for a contradiction that there are two disjoint vertex subsets $X \in \binom{[l]}{d+1}$ and $Y \in \binom{[l]}{2^{d+1}}$ such that $\{ N_J(y) \mid y \in Y \}$ shatters X. We are going to find a biinduced copy of Incidence $\left(2^{[d+1]}\right)$ in G, a contradiction.
Indeed, if we independently sample v_i uniformly from B_i for any $i \in X \cup Y$ we find that

$$\begin{aligned}\mathbb{P}\left(G\left[\{ v_i \mid i \in X \}, \{ v_j \mid j \in Y \}\right] \neq \mathrm{Incidence}\left(2^{[d+1]}\right)\right) &\leq \sum_{\{i,j\}\in\binom{X\cup Y}{2}} \mathbb{P}\left(\mathbb{1}\{\{v_i,v_j\} \in E(G)\} \neq \mathbb{1}\{\{i,j\} \in E(J)\}\right) \\ &\leq (d+1)\cdot 2^{(d+1)}\tilde{\epsilon} = (d+1)\cdot 2^{(d+1)}\epsilon^q < 1,\end{aligned}$$

where in the last inequality we used the Definition of q. Thus, the probabilistic method yields Claim 22. □

Now, $(*)$ applied to J yields a homogeneous set $I \subset [l]$ in J of size at least $l^{\tilde{C}} \geq \epsilon^{-q\tilde{C}}$. Let us define

$$H := G\left[\bigcup_{i\in I} B_i\right].$$

Observe that $|H| = |I|m \geq \epsilon^{-q\tilde{C}}\epsilon^{bq}|G| = \epsilon^{q(b-\tilde{C})}|G|$, where we remark that $q(b-\tilde{C}) > 0$. It is left to show that H is ϵ-restricted.
case I is an independent set. Then, using $|H| = |I|m$, for any $v \in V(H)$ we calculate

$$\deg_H(v) \leq m + (|I|-1)\,\epsilon^q m \leq |H|\left(\frac{1}{|I|} + \epsilon^q\right).$$

case I is a clique. Then for any $v \in V(H)$

$$\deg_H(v) \geq (|I|-1)(1-\epsilon^q)\,m = |H|\left(1-\frac{1}{I}\right)(1-\epsilon^q) \geq |H|\left(1-\epsilon^q-\frac{1}{I}\right).$$

Furthermore, observe that by $|I| \geq \epsilon^{-q\tilde{C}}$ and the Definition of q

$$\epsilon^q + \frac{1}{|I|} \leq \epsilon^q + \epsilon^{q\tilde{C}} \leq 2\epsilon^{q\tilde{C}} \leq \epsilon.$$

We conclude that H is an ϵ-restricted induced subgraph of G on at least $\epsilon^{q(b-\tilde{C})}|G|$ vertices, which shows the polynomial Rödl property with constant $q(b-\tilde{C})$. This completes the proof of Theorem 25. □

6 Concluding remarks

In this thesis we gave an extensive introduction into the induced Turán problem and its connections to the concept of VC dimension. In general, it would be interesting to better understand the connection between the usual, the biinduced and the induced extremal functions.
Lemma 9 shows that in most cases $\mathrm{ex}(n, \{F, H\text{-ind}\})$ is equal to either $\mathrm{ex}(n, F)(1+o(1))$ or $\mathrm{ex}(n, H)(1+o(1))$. In the manuscript by Hunter, Milojević, Sudakov, and Tomon [29] the authors independently state a Conjecture similar to the question raised at the end of our paper [4].

Conjecture 4 (Hunter, Milojević, Sudakov, and Tomon [29]). For any bipartite graph H $\exists T \in \mathbb{N}$ $\forall t \geq T$: $\mathrm{ex}(n, \{K_{t,t}, H\text{-ind}\}) = O(\mathrm{ex}(n, H))$ $(n \longrightarrow \infty)$.

A first step towards Conjecture 4 might be to prove that for any bipartite graphs H and F we have

$$\forall n \in \mathbb{N} : \ \mathrm{ex}(n, \{H, F\text{-ind}\}) = O(\mathrm{ex}(n, \{H, F\text{-biind}\})) \ \ (n \longrightarrow \infty).$$

Together with Theorem 17 this would translate into

$$\forall k, t \in \mathbb{N}, \ r \in \mathbb{N}_0 \text{ with } k \geq d \geq r+2 : \ \mathrm{ex}(n, \{K_{t,t}, W(k, d, r)\text{-ind}\}) = o\left(n^{2-\frac{1}{d}}\right) \ \ (n \longrightarrow \infty).$$

Furthermore, it would be very nice to resolve Conjecture 1 of Conlon and Lee which they give in [10]. It states that for any $d \in \mathbb{N}$ and any $K_{d,d}$-free bipartite graph H with maximum degree at most d in one part, there is a positive constant $\delta > 0$ such that $\mathrm{ex}(n, H) = O\left(n^{2-\frac{1}{d}-\delta}\right)$. Since this Conjecture is true in case that $d = 2$, see Theorems 12 and 13, one could ask if in this case for any $t \in \mathbb{N}$ there is also some positive constant $\delta' = \delta'(H, t) > 0$ such that $\mathrm{ex}(n, \{\, K_{t,t}, H\text{-ind}\,\}) = O\left(n^{\frac{3}{2}-\delta'}\right)$. The author is sure that at least the proof of Theorem 12 in [11] can be modified to show this strengthening, where one might want to use Lemma 21.
Example 1 raised the interesting problem of determining $\mathrm{ex}(G, H\text{-ind})$ in case that G and H are graphs such that $H \underset{ind}{\subseteq} G$. It appears that currently there are very few results known in this direction.
In Theorem 19 for $d, r \in \mathbb{N}$ with $r \leq d$ and any bipartite graph H that has one partite set A with r complete vertices $\tilde{A}$ such that the non-complete vertices in A have degree at most d, we give a lower bound on the number of induced isomorphisms from H to some other $K_{s,s}$-free graph G of a certain edge density. The main difficulty of the counting is to guarantee that in any copy of H in G the vertices in $\tilde{A}$ do not send any edges towards $A \setminus \tilde{A}$. If one would be interested in the count of induced, labeled subgraphs $\mathscr{I}$ of G that are isomorphic to some graph in $\mathscr{C}_H$, where we define

$$\mathscr{C}_H := \left\{ H + E' \;\middle|\; E' \subseteq \left\{ \{a, \tilde{a}\} \;\middle|\; a \in A \setminus \tilde{A}, \ \tilde{a} \in \tilde{A} \right\} \right\}$$

one would obtain the lower bound

$$|\mathscr{I}| = \Omega\left(|G|^{|H|}\left(\frac{\|G\|}{|G|^2}\right)^{\|H\|}\right) \ \ (|G| \longrightarrow \infty),$$

The author hopes that one could achieve this bound also for the number of induced isomorphisms if one would have a stronger tool at hand than Claim 14. Also compare this to the statement of Theorem 20. Currently, however it is not clear at all if the bound given in Theorem 19 is sharp.
At last, for a bipartite graph H it would be very interesting to show the Erdős-Hajnal conjecture for the graph property $\mathrm{Free}(H\text{-ind})$ instead of the graph property $\mathrm{Free}(H\text{-biind})$. Here again it would help to better understand the relation between the $\{\, H\text{-biind}\,\}$-free graphs and the $\{H\text{-ind}\}$-free graphs.

References

[1] Noga Alon, Eldar Fischer and Ilan Newman. *Efficient Testing of Bipartite Graphs for Forbidden Induced Subgraphs.* SIAM Journal on Computing 37(3): 959–976, (2007).

[2] Noga Alon, Michael Krivelevich, and Benny Sudakov. *Turán numbers of bipartite graphs and related Ramsey-type questions.* Combin. Probab. Comput. 12: 477–494, (2003).

[3] Noga Alon, Lajos Rónyai and Tibor Szabó. *Norm-Graphs: Variations and Applications.* J. Combin. Theory Ser. B 76(2): 280–290, (1999).

[4] Maria Axenovich and Jakob Zimmermann. *Induced Turán problem in bipartite graphs.* arXiv:2401.11296 (2024).

[5] Clark Tabor Benson. *Minimal Regular Graphs of Girths Eight and Twelve.* Canadian Journal of Mathematics 18: 1091—1094, (1966).

[6] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler and Manfred Warmuth. *Learnability and the Vapnik-Chervonenkis dimension.* Journal of the ACM 36(4): 929–965, (1989).

[7] Marthe Bonamy, Nicolas Bousquet, Michał Pilipczuk, Paweł Rzażewski, Stéphan Thomassé, and Bartosz Walczak. *Degeneracy of P_t-free and $C_{\geq t}$-free graphs with no large complete bipartite subgraphs.* J. Combin. Theory Ser. B. 152: 353–378, (2022).

[8] John Adrian Bondy and Miklós Simonovits. *Cycles of even length in graphs.* J. Combin. Theory Ser. B 16(2): 97–105, (1974).

[9] Nicolas Bousquet, Aurélie Lagoutte, Zhentao Li, Aline Parreau and Stéphan Thomassé. *Identifying Codes in Hereditary Classes of Graphs and VC dimension.* SIAM Journal on Discrete Mathematics 29(4): 2047–2064, (2015).

[10] David Conlon, Oliver Janzer, and Joonkyung Lee. *More on the extremal number of subdivisions.* Combinatorica 41: 465–494, (2021).

[11] David Conlon and Joonkyung Lee. *On the extremal number of subdivisions.* Int. Math. Res. Not. IMRN 12: 9122–9145, (2021).

[12] David Conlon, Jacob Fox and Benny Sudakov. *An Approximate Version of Sidorenko's Conjecture.* Geometric and Functional Analysis 20: 1354–1366, (2010).

[13] David Conlon, Jacob Fox and Benny Sudakov. *Sidorenko's Conjecture for a class of graphs: an exposition.* arXiv:1209.0er184 (2012).

[14] Mónika Csikós, Nabil H. Mustafa and Andrey Kupavskii. *Tight Lower Bounds on the VC dimension of Geometric Set Systems.* Journal of Machine Learning Research 20(81): 1–8, (2019).

[15] Reinhard Diestel. *Graph Theory.* 5th edition, Graduate Texts in Mathematics Book 173, (Springer, 2017).

[16] Paul Erdős and András Hajnal. *Ramsey-type Theorems.* Discrete Applied Mathematics 25(1): 37–52, (1989).

[17] Paul Erdős, András Hajnal and János Pach. *A Ramsey-type Theorem for bipartite graphs.* Geombinatorics 10: 64–68, (2000).

[18] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. *Semi-algebraic version of Zarankiewicz's problem.* J. Eur. Math. Soc. 19(6): 1785–1810, (2017).

[19] Jacob Fox, János Pach and Andrew Suk. *Bounded VC dimension implies the Schur-Erdős Conjecture.* 36rd International Symposium on Computational Geometry, Art. No. 46, 8 pp., LIPIcs. Leibniz Int. Proc. Inform., 164, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, (2020).

[20] Jacob Fox, János Pach, and Andrew Suk. *Erdős-Hajnal conjecture for graphs with bounded VC dimension.* 33rd International Symposium on Computational Geometry, Art. No. 43, 15 pp., LIPIcs. Leibniz Int. Proc. Inform., 77, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, (2017).

[21] Jacob Fox and Benny Sudakov. *Dependent Random Choice.* Random Structures & Algorithms 38(1-2): 68–99, (2011).

[22] Zoltán Füredi. *New asymptotics for bipartite Turán numbers.* J. Combin. Theory Ser. A 75(1): 141-–144, (1996).

[23] Zoltán Füredi. *On a Turán type problem of Erdős.* Combinatorica 11: 75–79, (1991).

[24] Zoltán Füredi and Miklos Simonovits. *The history of degenerate (bipartite) extremal graph problems.* in *Erdős Centennial.* Bolyai Soc. Math. Stud. Volume 25: 169–264, (Springer Berlin, 2013).

[25] William Timothy Gowers. *Hypergraph Regularity and the Multidimensional Szemerédi Theorem.* Annals of Mathematics 166(3): 897-946, (2007).

[26] David Haussler. *Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension.* J. Combin. Theory Ser. A 69(2): 217–232, (1995).

[27] David Haussler and Emo Welzl. *-nets and simplex range queries.* Discrete Computational Geometry 2: 127–151, (1987).

[28] Daniel Hug, Wolfgang Weil. *Lectures on Convex Geometry.* 1st Edition, Graduate Texts in Mathematics Book 286: 13 pp., (Springer, 2020).

[29] Zach Hunter, Aleksa Milojević, Benny Sudakov, and István Tomon. *Kővári-Sós-Turán theorem for hereditary families.* arXiv:2401.10853v1 (2024).

[30] Freddie Illingworth. *A note on induced Turán numbers.* arXiv:2105.12503, (2021).

[31] Oliver Janzer. *Improved bounds for the extremal number of subdivisions.* Electron. J. Combin. 26: 1 – 6, (2019).

[32] Oliver Janzer and Cosmin Pohoata. *On the Zarankiewicz problem for graphs with bounded VC dimension.* arXiv:2009.00130 (2021)

[33] Achim Klenke. *Wahrscheinlichkeitstheorie.* 1st. Edition: 144 p. (Springer Berlin Heidelberg, 2006).

[34] Tao Jiang and Robert Seiver. *Turán numbers of subdivided graphs.* SIAM Journal on Discrete Mathematics 26(3): 1238–1255, (2012).

[35] Andrey Kupavskii, Nabil Mustafa and János Pach. *New Lower Bounds for epsilon-Nets.* 32nd International Symposium on Computational Geometry, Art. No. 43, 1 – 16, LIPIcs. Leibniz Int. Proc. Inform., 51, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, (2016).

[36] Tamás Kővári, Vera T. Sós and Pál Turán. *On a problem of K. Zarankiewicz.* Colloquium Mathematicum 3: 50–57, (1954).

[37] Po-Shen Loh, Michael Tait, Craig Timmons, and Rodrigo M. Zhou. *Induced Turán numbers.* Combin. Probab. Comput. 27 no. 2: 274–288, (2018).

[38] Jiří Matoušek. *Lectures on Discrete Geometry.* 1st. Edition, Graduate Texts in Mathematics Book 212: 245 pp., (Springer New York, 2002).

[39] Brendan Nagle, Vojtěch Rödl, and Mathias Schacht. *The counting lemma for regular k-uniform hypergraphs.* Random Structures & Algorithms 28(2): 113–179, (2006).

[40] Tung Nguyen, Alex Scott and Paul Seymour. *Induced subgraph density. VI. Bounded VC dimension.* arXiv:2312.15572 (2024).

[41] Vojtěch Rödl. *On universality of graphs with uniformly distributed edges.* Discrete Mathematics 59(1): 125–134, (1986).

[42] Alexander Sidorenko. *A correlation inequality for bipartite graphs.* Graphs Combin 9: 201–204, (1993).

[43] Thoralf Skolem. *Some Remarks on the Triple Systems of Steiner.* Mathematica Scandinavica 6: 273-–280, (1958).

[44] Benny Sudakov and István Tomon. *Turán number of bipartite graphs with no $K_{t,t}$.* Proc. Amer. Math. Soc. 148: 2811-2818, (2020).

[45] Vladimir Vapnik and Alexey Chervonenkis. *On the uniform convergence of relative frequencies of events to their probabilities.* Theory Probab. Appl. 16: 264–280, (1971).

[46] Vladimir Vapnik and Alexey Chervonenkis. *On the Uniform Convergence of the Frequencies of Occurrence of Events to Their Probabilities.* Edited by Bernhard Schölkopf, Zhiyuan Luo, and Vladimir Vovk in *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik.* 7–12, (Springer Berlin Heidelberg, 2013).