Vertex Folkman numbers and the minimum degree of minimal Ramsey graphs

Hiêp Hàn∗1, Vojtěch Rödl†2, and Tibor Szabó‡3

1Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Chile
2Department of Mathematics and Computer Science, Emory University, USA
3Institute of Mathematics, Freie Universität Berlin, Germany

February 1, 2019

Abstract

We investigate the smallest possible minimum degree of \(r\)-color minimal Ramsey-graphs for the \(k\)-clique. In particular, we obtain a bound of the form \(O(k^2 \log^2 k)\), which is tight up to a \((\log^2 k)\)-factor whenever the number \(r \geq 2\) of colors is fixed. This extends the work of Burr, Erdős, and Lovász who determined this extremal value for two colors and any clique size, and complements that of Fox, Grinshpun, Liebenau, Person, and Szabó who gave essentially tight bounds when the order \(k\) of the clique is fixed.

As a side product our result also yields an improved upper bound on the vertex Folkman number \(F(r, k, k + 1)\) of the \(k\)-clique.

The proof relies on a reformulation of the corresponding extremal function by Fox et al, and combines and refines methods used by Dudek, Eaton, and Rödl.

∗Supported by the Iniciación grant 11150913 and Nucleo Milenio Información y Coordinación en Redes ICM/FIC RC130003, Chile.
†Supported by the NSF grant DMS 1301698.
‡Supported by the GIF grant G-1347-304.6/2016.
1 Introduction

A graph $G$ is $r$-Ramsey for a graph $H$, denoted by $G \to (H)_r$, if every $r$-coloring of the edges of $G$ yields a monochromatic copy of $H$. A graph $G$ is $r$-Ramsey-minimal for $H$ (or $r$-minimal for $H$) if $G \to (H)_r$, but none of the proper subgraphs $G' \subsetneq G$ satisfies $G' \to (H)_r$. Let $\mathcal{M}_r(H)$ denote the family of all graphs $G$ that are $r$-Ramsey-minimal with respect to $H$. Ramsey’s theorem implies that $\mathcal{M}_r(H)$ is non-empty for all integers $r$ and all finite graphs $H$.

We consider the smallest minimum degree $s_r(H)$ an $r$-Ramsey-minimal graph can have, i.e.

$$s_r(H) := \min_{G \in \mathcal{M}_r(H)} \delta(G),$$

where $\delta(G)$ denotes the minimum degree of graph $G$.

In [2] Burr, Erdős, and Lovász addressed the two color case and showed that $s_2(K_{k+1}) = k^2$. Fox, Grinshpun, Liebenau, Person, and Szabó [7] studied $s_r(K_{k+1})$ as a function of $r$. In particular, for every fixed $k$, they determined $s_r(K_{k+1})$ up to a polylogarithmic factor.

Theorem A ([7]).

(i) There exist constants $c, C > 0$ such that for all $r \geq 2$, we have

$$cr^2 \log r \leq s_r(K_3) \leq Cr^2 \log^2 r.$$

(ii) For all $k \geq 3$ there exist constants $c_k, C'_k > 0$ such that, for all $r \geq 3$, we have

$$c_k r^2 \frac{\log r}{\log \log r} \leq s_r(K_{k+1}) \leq C'_k r^2 (\log r)^{8k^2}.$$

The upper bound in (ii) is not polynomial in $k$, but in the same paper the following bound, polynomial in both parameters, was also given.

Theorem B ([7]). For all $k \geq 2$ and $r \geq 3$, $s_r(K_{k+1}) \leq 8k^6 r^3$.

Our main motivation is to investigate $s_r(K_{k+1})$ in the case when the number $r$ of colors is fixed and the order $k$ of the clique tends to infinity. Our main result is the following.

Theorem 1. For every integer $r \geq 2$ there exists a constant $C_r$ such that for every integer $k \geq 3$

$$s_r(K_{k+1}) \leq C_r (k \log k)^2.$$
Remarks.

• For every fixed $r \geq 3$ and $k$ tending to infinity the upper bound of Theorem [1] is tight up to the $(\log^2 k)$-factor. Indeed, $s_r(K_{k+1})$ is monotone in $r$ (see [7]), hence $s_r(K_{k+1}) \geq s_2(K_{k+1}) = k^2$.

• It turns out that the value of the constant $C_r$ is of the order $r^3 \log^3 r$ provided $k \geq \sqrt{r}$ (see Theorem 2' and (1)). This represents an improvement over Theorem B in this range.

Our proof relies on the following translation by Fox et al. [7] of the function $s_r(K_{k+1})$ to another extremal function $P_r(k)$. A sequence of pairwise edge-disjoint graphs $G_1, \ldots, G_r$ on the same vertex set $V$ is called a color pattern on $V$ (with the edges of graph $G_i$ set to have color $i$). For a graph $H$, a color pattern $G_1, \ldots, G_r$ is called $H$-free if none of the $G_i$ contains $H$ as a subgraph.

Consider an $r$-coloring of $V$ with colors $1, \ldots, r$. A graph with colored vertices and edges is called strongly monochromatic if all its vertices and edges have the same color. We define $P_r(k)$ to be the smallest integer $N$ such that there exists a $K_{k+1}$-free color pattern $G_1, \ldots, G_r$ on an $N$-element vertex set $V$ with the property that any coloring of $V$ with the colors $1, \ldots, r$ yields a strongly monochromatic copy of $K_k$.

**Theorem C ([7]).** For all integers $r \geq 2$, $k \geq 2$, we have $s_r(K_{k+1}) = P_r(k)$.

To find the desired color pattern and hence bound $P_r(k)$ from above we introduce a function called color pattern density Folkman number (or just color pattern number). For a real $\alpha > 0$ and positive integers $r$ and $k$, let $CP_r(\alpha, k)$ be the smallest integer $N$ for which there exists a $K_{k+1}$-free color pattern $G_1, \ldots, G_r$ on an $N$-element vertex set $V$ such that for every $i \in [r]$ the graph $G_i$ induces a copy of $K_k$ in any subset $U \subseteq V$ of size $\alpha|V|$.

It follows from Theorem 2 below that $CP_r(\alpha, k)$ exists. Further, the two functions $s_r$ and $CP_r$ connect naturally, as for every $k$

$$s_r(K_{k+1}) = P_r(k) \leq CP_r(1/r, k)$$

holds. To see this, let $G_1, \ldots, G_r$ be a $K_{k+1}$-free color pattern on $|V| = CP_r(1/r, k)$ vertices such that each $G_i$ induces a copy of $K_k$ in any subset of $V$ of size $|V|/r$. We claim that for any coloring of the vertex set $V$ with the colors $1, \ldots, r$, there is a strongly monochromatic $K_k$ in any vertex color.
class of size at least $|V|/r$ (in particular in the largest one). Indeed, if the color class $U_i \subseteq V$, belonging to color $i$, has size at least $|V|/r$, then $G_j[U_i]$ contains a $K_k$ for every $j \in [r]$. In particular for the color $j = i$ this copy is strongly monochromatic. Thus inequality (1) follows.

Now Theorem 1 immediately follows from (1) and the following Theorem 2.

**Theorem 2.** For every integer $r \geq 2$ there exists a constant $C_r$ such that

$$CP_r(1/r,k) \leq C_r(k \log k)^2.$$  \hspace{1cm} (2)

Our proof of Theorem 2 combines the modification of a construction given by Dudek and Rödl [3] with that of Eaton and Rödl [4] and extends the result of Dudek and Rödl who investigated the closely related notion of vertex Folkman numbers (see e.g. [9, 3]). Let $F(r, k, m)$ denote the minimum number $N$ such that there is a graph $G$ on $N$ vertices with the property that $G$ does not contain a $K_m$, yet any $r$-coloring of the vertices of $G$ yields a monochromatic copy of $K_k$. Clearly,

$$F(r, k, k + 1) \leq CP_1(1/r,k) \leq CP_r(1/r,k).$$  \hspace{1cm} (3)

To show their upper bound on $F(r, k, k + 1)$ Dudek and Rödl [3] in fact showed that $CP_1(1/r, k) = O(k^2 \log k)$ for every $r \geq 2$ and used the first inequality of (3). Our Theorem 2 implies a $\log^2 k$-factor improvement via the second inequality of (3).

**Corollary 1.** For every integer $r \geq 2$ there exists a constant $C_r$ such that for every integer $k \geq 3$

$$F(r, k, k + 1) \leq C_r k^2 \log k.$$  \hspace{1cm} (3)

Logarithm in our paper is always of natural base. Further, we regularly omit floor and ceiling signs whenever they are immaterial due to the asymptotic nature of our statements.

## 2 The construction

The main purpose of this section is to describe the construction behind Theorem 2. Using this construction we will establish Theorem 2 for large $k$. 

4
Theorem 2. There exists an absolute constant $k_0$ such that for all $k > k_0$ and all $r < k^2$ the following holds

$$CP_r(1/r, k) < 80^3(r \log r)^3(k \log k)^2.$$ (4)

We postpone the proof and first show that Theorem 2 follows from Theorem 2' and a result from [7].

Proof of Theorem 2. We will apply Theorem 2' for large $k$ and Theorem A for small $k$.

The following upper bound on $CP_r(1/r, k)$ was derived in [7] within the proof of Theorem A. There it was established that

$$CP_r(1/r, k) < C'_k r^2 (\log 2r)^8k^2,$$ (5)

for all $k, r \geq 2$ and for some constant $C'_k$ (depending on $k$).

Let $k_0$ be the constant from Theorem 2 and let $\tilde{r} := \max\{k_0, \sqrt{r}\}$. Then Theorem 2 holds with

$$C_r = \max \left\{ 80^3(r \log r)^3, \tilde{C}_r r^2 (\log 2r)^8\tilde{r}^2 \right\},$$

where $\tilde{C}_r = \max_{k \leq \tilde{r}} C'_k$.

Indeed, for $k > \max\{k_0, \sqrt{r}\}$ Theorem 2 immediately follows from (4) of Theorem 2 and for $k \leq \tilde{r}$, it follows from (5). □

Therefore it only remains to prove Theorem 2'.

Proof of Theorem 2'. For every $k > \sqrt{r}$ sufficiently large we are going to construct $r$ edge-disjoint graphs $G_1, \ldots, G_r$ on a vertex set $V$ of size at most $80^3(r \log r)^3(k \log k)^2$, such that the following two properties hold.

(P1) For every $i \in [r]$ and every subset $U \in \binom{V}{|V|/r}$ the graph $G_i[U]$ contains a copy of $K_k$.

(P2) For every $i \in [r]$ the graph $G_i$ contains no copy of $K_{k+1}$.

1While the notion of color pattern density Folkman number $CP_r$ was not explicitly defined in [7], graphs on $n$ vertices with the property in the definition of $CP_r(1/r, k)$ were introduced and called $(n, r, k)$-critical. This concept was then used in the proof of the upper bound of Theorems A and B together with inequality (1) (which appears as Lemma 4.1 in [7]). In fact all previous upper bounds for $s_r(K_{k+1})$ (e.g. in [2, 7]) are derived by proving the same upper bound for $CP_r(1/r, k)$ and then applying (1).
The rest of this section is concerned with the description of the construction of $G_1, \ldots, G_r$. The proofs of (P1) and (P2) are presented in Sections 3 and Section 4, respectively. Besides the condition $k > \sqrt{r}$ we also require $k > k_0$, where $k_0$ is an absolute constant, the existence of which will be clear from the proofs.

First we choose the parameters of our construction:

$$z = \log k, \quad x = (15r \log r)z, \quad y = 6\sqrt{x}. \quad (6)$$

By Chebyshev’s Theorem there exists a prime $q$ such that

$$\frac{kxy}{z} \leq q + 1 \leq 2\frac{kxy}{z}. \quad (7)$$

The desired graphs are then constructed as follows.

1. Let $PG(q)$ be the projective plane over the $q$-element field, with point set $\mathcal{P}$ and line set $\mathcal{L}$. We take the “blow-up” of $\mathcal{P}$ replacing the points $v \in \mathcal{P}$ by pairwise disjoint $z$-element sets $Z_v = \{(v, i) : i \in [z]\}$. Our graphs $G_1, \ldots, G_r$ will be constructed on the vertex set $V = \bigcup_{v \in \mathcal{P}} Z_v$, which has size $|V| = (q^2 + q + 1)z < (q + 1)^2z \leq \frac{(2kx(6\sqrt{x}))^2}{z} < 80^3(r \log r)^3(k \log k)^2$.

For every line $L \in \mathcal{L}$ let $M(L) = \bigcup_{v \in L} Z_v$ denote the blow-up of $L$. Note that each such blown-up line has size

$$|M(L)| = (q + 1)z = \Theta((r \log r)^{3/2}k \log^{3/2} k).$$

The set of all blown-up lines is denoted by $\mathcal{M}$. For technical reasons we fix a subset $\hat{\mathcal{L}} \subseteq \mathcal{L}$ of the lines such that $|\hat{\mathcal{L}}|$ is divisible by $r$ and

$$q^2 + q + 1 = |\mathcal{L}| \geq |\hat{\mathcal{L}}| \geq q^2 + q + 1 - r + 1.$$  

Let $\hat{\mathcal{M}} = \{M(L) : L \in \hat{\mathcal{L}}\}$ be the family of corresponding blown-up lines. Then

$$|\hat{\mathcal{M}}| = |\hat{\mathcal{L}}| = \Theta((r \log r)^3k^2 \log k).$$
2. For every $M \in \hat{\mathcal{M}}$ we construct a random graph $H_M$ on vertex set $M$ as follows. First, we choose a uniform random partition of $M$ into $k + 1$ partition classes $M = R_0 \cup R_1 \cup \cdots \cup R_k$ where $|R_i| = x$ for $1 \leq i \leq k$ and $R_0 = M \setminus \bigcup_{i \in [k]} R_i$ with $|R_0| = (q + 1)z - kx$. Note that now there are two partitions on each $M$: the fixed partition given by $M = M(L) = \bigcup_{v \in L} Z_v$ and the random partition given by $M = R_0 \cup R_1 \cup \cdots \cup R_k$.

We call a set $S$ of vertices crossing with respect to a partition $(Z_v)_{v \in P}$ if $S \subseteq \bigcup_{i \in I} P_i$ and every class $P_i$ contains at most one vertex of $S$. As the edges of the graph $H_M$ we take exactly those pairs $a, b \in M$ which are crossing with respect to both partitions $(Z_v)_{v \in L}$ and $(R_i)_{i \in [k]}$. In other words, the edges of $H_M$ are obtained from the edges of the $k$-partite Turán graph on the random partition classes $R_1 \cup \cdots \cup R_k$, with those being deleted which are non-crossing with respect to $(Z_v)_{v \in L}$. In particular $\{a, b\} \cap R_0 = \emptyset$.

3. We partition the set $\hat{\mathcal{M}}$ of blown-up lines uniformly at random into $r$ parts $\mathcal{M}_1, \ldots, \mathcal{M}_r$ of equal size $|\hat{\mathcal{M}}|/r$. As the graphs $G_i$, $i \in [r]$, we take the union of all the graphs $H_M$ with $M \in \mathcal{M}_i$:

$$G_i := \bigcup_{M \in \mathcal{M}_i} H_M.$$ 

Note the following important facts about the two types of pairs of vertices in our construction.

**Fact.**

- **Pairs of vertices which are non-crossing with respect to the fixed partition $(Z_v)_{v \in P}$ are not adjacent in any $H_M$;**

- **Any pair of vertices $x, y \in V$ which is crossing with respect to $(Z_v)_{v \in P}$ has a unique blown-up line $M^*_{x,y} \in \mathcal{M}$ that contains it.**

These two facts imply that the graphs $H_M$ in the definition of $G_i$ are pairwise edge-disjoint, and that $G_i$ and $G_j$ are edge-disjoint for every $i \neq j$. Further, by definition, $xy \in E(G_i)$ if and only if $M^*_{x,y} \in \mathcal{M}_i$ and $xy \in E(H_{M^*_{x,y}})$.

To conclude the proof of Theorem 2', in the next two sections we establish the following two lemmas.
Lemma 1. There exists a $k_0$ such that for every $k > k_0$ and $k > \sqrt{r}$ the following holds. Let $\mathcal{K} \subseteq \hat{\mathcal{M}}$ be an arbitrary family of $\frac{\lvert \hat{\mathcal{M}} \rvert}{r}$ blown-up lines and let $H = \bigcup_{M \in \mathcal{K}} H_M$ be the corresponding random graph, defined above. Then we have
\[
P\left[ \exists U \subseteq \left(\frac{V}{\lvert V \rvert / r}\right) \text{ such that } H[U] \not\supseteq K_k \right] < \frac{1}{2r}.
\]

Lemma 2. Let $k \geq k_0$, where $k_0$ is a large enough absolute constant. With probability larger than $1/2$ the graph $G_i$ is $K_{k+1}$-free for every $i \in [r]$.

Using Lemma 1 with $\mathcal{K} = \mathcal{M}_i$ for all $i \in [r]$ together with the union bound implies that (P1) fails to hold with probability less than $1/2$. Lemma 2 explicitly states that (P2) holds with probability greater than $1/2$. Hence our construction satisfies both properties with non-zero probability.

3 Every coloring yields a monochromatic $K_k$

Lemma 1 itself is a consequence of the following two lemmas. For a set $U \subseteq V$ of size $\frac{|V|}{r}$ the blown-up line $M \in \mathcal{M}$ is called $U$-good if $|U \cap M| \geq \frac{9}{10} \cdot \frac{|M|}{r} = \frac{9}{10r}(q+1)z$, otherwise it is called $U$-bad. The following extension of a lemma from [4] shows that almost all of the $q^2 + q + 1$ blown-up lines are $U$-good.

Lemma 3. Suppose that $k > \sqrt{r}$. Then, for every set $U \subset V$ of size $\frac{|V|}{r}$ there are at most $100qr$ blown-up lines in $\mathcal{M}$ which are $U$-bad. In particular, for every $\mathcal{K} \subseteq \mathcal{M}$ with $|\mathcal{K}| = \frac{|\hat{\mathcal{M}}|}{2r}$ there are at least $\frac{|\hat{\mathcal{M}}|}{2r}$ blown-up lines in $\mathcal{K}$ which are $U$-good.

Our second lemma states that the intersection of a $\frac{|V|}{r}$-subset with a good line is very likely to contain a $k$-clique.

Lemma 4. Let $U \subset V$ be a set of size $|U| = \frac{|V|}{r}$ and let $M \in \hat{\mathcal{M}}$ be a $U$-good line. Then, provided $k$ is large enough, we have
\[
P\left[ H_M[U \cap M] \not\supseteq K_k \right] < k^{-2.1 \log(er)}.
\]

We postpone the proofs of the two lemmas and first show how to derive Lemma 1.
Proof of Lemma 1. Let \( U \subseteq V \) be an arbitrary subset of size \( |U| = \frac{|V|}{r} \). Since \( H = \bigcup_{M \in \mathcal{K}} H_M \) we infer that if \( H[U] \) does not contain a copy of \( K_k \) then none of the graphs \( H_M[U \cap M] = H[U \cap M] \) with \( M \in \mathcal{K} \) do. This implies that

\[
\mathbb{P}[H[U] \not\supseteq K_k] \leq \prod_{M \in \mathcal{K}} \mathbb{P}[H[U \cap M] \not\supseteq K_k] = \prod_{M \in \mathcal{K}} \mathbb{P}[H[U \cap M] \not\supseteq K_k],
\]

where the last equality follows since the random graphs \( H_M \) are independent for different \( M \). We can now estimate further by taking the product only over the \( U \)-good lines \( M \in \mathcal{K} \) and apply Lemma 3 and Lemma 4 to obtain

\[
\mathbb{P}[H[U] \not\supseteq K_k] \leq \prod_{M \in \mathcal{K}, \text{\(U\)-good}} \mathbb{P}[H[U \cap M] \not\supseteq K_k] \leq \prod_{M \in \mathcal{K}, \text{\(U\)-good}} \frac{1}{k^{2.1 \log(\epsilon r)}} \leq k^{-\frac{2.1 \log(\epsilon r)}{2r}|\mathcal{M}|}.
\]

Taking the union bound and assuming that \( k \) is sufficiently large, the probability that there is a \( U \subseteq V \) with \( |U| = \frac{|V|}{r} \) and \( H[U] \not\supseteq K_k \) is at most

\[
\left( \frac{|V|}{|V|/r} \right)^k \cdot \frac{2.1 \log(\epsilon r)}{2r}|\mathcal{M}| \leq (er)^{|V|/r} k^{-\frac{2.1 \log(\epsilon r)}{2r}|\mathcal{M}|}
\]

\[
\leq \exp \left( \frac{|\mathcal{M}| z}{r} \log(\epsilon r) - \log k \frac{2.1 \log(\epsilon r)}{2r} (|\mathcal{M}| - r + 1) \right)
\]

\[
\leq \exp \left( -\frac{\log(\epsilon r)}{21r} |\mathcal{M}| \log k \right) < \frac{1}{2r}.
\]

\[ \square \]

To complete the proof of Lemma 1 we now establish Lemma 3 and Lemma 4. The proof of the former will rely on the expansion property of the Erdős-Rényi \([5, 6]\) polarity graph \( ER_q \). The vertex set of \( ER_q \) is the set of points of the projective plane \( PG(q) \). The vertices \([x_0, x_1, x_2]\) and \([y_0, y_1, y_2]\) form an edge in \( ER_q \) if \( x_0 y_0 + x_1 y_1 + x_2 y_2 = 0 \) (where we allow loops). Notice that both the point set \( \mathcal{P} \) and the line set \( \mathcal{L} \) from our construction can be identified with \( V(ER_q) \), furthermore point \([x_0, x_1, x_2]\) \( \in \mathcal{P} \) is incident to line \([y_0, y_1, y_2]\) \( \in \mathcal{L} \) if and only if the corresponding vertices are adjacent in \( ER_q \).

We will need the Expander Mixing Lemma \([11, 12]\).

**Lemma 5.** Let \( \Gamma = (V, E) \) be a \( d \)-regular graph with at most one loop at each vertex. Let \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \) be the eigenvalues of the adjacency matrix
of $\Gamma$ and let $\lambda = \max\{\lambda_2, \vert\lambda_n\vert\}$ denote the second largest absolute value of the eigenvalues. Then, for all $A, B \subseteq V$, not necessarily disjoint, we have

$$\left\| e(A, B) - \frac{d}{n} \|A\|\|B\| \right\| \leq \lambda \sqrt{\|A\|\|B\|},$$

where $e(A, B) = \{(a, b) \in A \times B: \{a, b\} \in E\}$.

For the application of the Expander Mixing Lemma we recall that the polarity graph $ER_q$ is $(q+1)$-regular, has $q^2 + q + 1$ vertices, and the second largest absolute value of its eigenvalues is equal to $\sqrt{q}$ (see e.g. [8]).

We now turn to the proof of Lemma 3. Recall from the construction that $V = \mathcal{P} \times [z]$ where $\mathcal{P}$ is the point set of the projective plane $PG(q)$. For any subset $U \subseteq V$ of size $|U| = \frac{|\mathcal{L}|}{r}$ let $B \subseteq \mathcal{L}$ be the set of all lines $L$ whose blow-up $M = M(L)$ is $U$-bad. To establish Lemma 3 we need to show that $|B| \leq 100 qr$.

Let $U_i := \{u \in \mathcal{P} : (u, i) \in U\}$ and let $e(U_i, B)$ denote the number of point-line incidences, i.e., pairs $(u, L) \in U_i \times B$ such that $u \in L$. Note that $\sum_{i \in [z]} |U_i \cap L| = |U \cap M(L)|$ and due to $U$-badness of $M(L)$ for each $L \in B$ we infer

$$\sum_{i \in [z]} e(U_i, B) = \sum_{i \in [z]} \sum_{L \in B} |U_i \cap L| = \sum_{L \in B} |U \cap M(L)| < \frac{9}{10r} |B|(q + 1)z. \tag{8}$$

To give a lower bound on the sum, we interpret both $U_i$ and $B$ as vertex subsets of the polarity graph $ER_q$. We bound $e(U_i, B)$ using Lemma 5 with $d = q + 1, n = q^2 + q + 1$, and $\lambda = \sqrt{q}$, which yields $e(U_i, B) \geq \frac{d}{n} |U_i| |B| - \sqrt{q |U_i| |B|}$ for every $i \in [z]$. This implies

$$\frac{9}{10} \cdot \frac{d}{n} |U||B| \geq \sum_{i=1}^{z} e(U_i, B) \geq \sum_{i=1}^{z} \left( \frac{d}{n} |U_i||B| - \sqrt{q |U_i||B|} \right)$$

$$= \frac{d}{n} |U||B| - \sqrt{q|B|} \sum_{i=1}^{z} \sqrt{|U_i|}$$

where the upper bound follows from [8]. This means that

$$\frac{d}{10n} |B||U| \leq \sqrt{q|B|} \sum_{i=1}^{z} \sqrt{|U_i|} \leq \sqrt{q|B|} \sqrt{n},$$

10
where we used the Cauchy-Schwarz inequality and that \( \sum_{i=1}^{z} |U_i| = |U| = nz/r \). The bound resolves to
\[
\sqrt{|B|} \leq \frac{10\sqrt{qrn}}{d} \leq 10\sqrt{qr},
\]
proving the first claim. The second claim follows from (7) and \( k > \sqrt{r} \) as
\[
|\hat{M}| = q^2 + q + 2 - r > q\frac{kxy}{z} - r > q \cdot 6\sqrt{r}(15r \log r)^{3/2} - r > 200qr^2.
\]

It remains to prove Lemma \ref{lem:point-set}.

**Proof of Lemma \ref{lem:point-set}**. Given a set \( U \subset V \) of size \( |V|/r \) and a \( U \)-good line \( M \in \hat{M} \) we want to show that \( \Pr(K_k \not\subseteq H_M[U \cap M]) < k^{-2.1 \log(εr)} \).

Recall that we have a fixed partition of \( M \) given by \( M(L) = \bigcup_{v \in L} Z_v \) for some \( L \in \mathcal{L} \). Let \( U_v = U \cap Z_v \) for each \( v \in L \). This constitutes a partition of \( U \cap M = \bigcup_{v \in L} U_v \). Further, we have a random partition of \( M \) given by \( M = R_0 \cup \cdots \cup R_k \).

Consider the auxiliary bipartite graph \( B \) with the partition classes \( [k] \) and \( L \), in which \( i \in [k] \) is adjacent to \( v \in L \) if and only if \( R_i \cap U_v \neq \emptyset \). Then \( U \cap M \) contains a copy of \( K_k \) if and only if there is a matching saturating all vertices from \( [k] \). Indeed, such a matching is equivalent to the existence of \( k \) points \( u_1, \ldots, u_k \in U \) which are in distinct \( U_v \)'s and in distinct \( R_i \)'s. In other words, we have a point set of size \( k \) in which each pair of vertices are crossing with respect to both partitions \((Z_v)_{v \in L} \) and \((R_i)_{i \in [k]} \). By the definition of \( H_M \), these points are pairwise adjacent, hence form a \( k \)-clique in \( H_M[U \cap M] \).

For an index set \( J \subseteq [k] \) let \( R_J = \bigcup_{i \in J} R_i \) and for a subset \( I \subseteq L \) let \( U_I = \bigcup_{v \in I} U_v \). By Hall’s theorem, a matching saturating the vertex set \( [k] \) in the bipartite graph \( B \) does not exist if and only if there is a set \( J \subseteq [k] \) and a set \( I \subseteq L \) of size \( |I| = |J| - 1 \) such that the neighborhood of \( J \) in \( B \) is completely contained in \( I \). By the definition of \( B \) this means that for every \( j \in J \) and \( v \in L \setminus I \) we have that \( R_j \cap U_v = \emptyset \). In other words we have \( R_J \cap U_{L \setminus I} = \emptyset \) and hence we conclude that
\[
R_J \subseteq M \setminus U_{L \setminus I} = U_I \cup (M \setminus U).
\]

So far we have argued that if \( H_M[U \cap M] \) does not contain a \( K_k \), then there is a set \( J \subseteq [k] \) and a set \( I \subseteq L \) of size \( |I| = |J| - 1 \), such that (9) is satisfied.
Next we will give an upper bound for the probability that the random partition $R_0, \ldots, R_k$ satisfies (9) for a fixed subset $J \subseteq [k]$ of size $j$ and fixed subset $I \subseteq L$ of size $|I| = j - 1$ and then finish the argument by the union bound. Note that $|R_J| = jx$ and let $t = |U_I \cup (M \setminus U)|$. Then we have

$$t \leq (j - 1)z + |M \setminus U| \leq kz + |M| \left(1 - \frac{9}{10r}\right) \leq |M| \left(1 - \frac{4}{5r}\right),$$

since $M$ is a $U$-good line (and hence $|U \cap M| \geq \frac{9}{10r} |M|$) and $kz \leq \frac{(q+1)x}{10r} z \leq \frac{(q+1)}{10r} |M|$ due to (7).

Observe that once $I$ is fixed then the subset $U_I$ and hence also $t$ are fixed, while once $J$ is fixed the set $R_J$ is just a uniformly chosen random subset of size $jx$. Then the probability that the random partition $R_0 \cup \cdots \cup R_k$ satisfies (9) is equal to the probability that a uniformly chosen set of size $jx = |R_J|$ is subset of a fixed set of size $t$. The latter is at most

$$\binom{t}{jx} = \binom{t}{M} \cdots \binom{|M| - jx + 1}{M} \leq \left(\frac{t}{|M|}\right)^{jx} \leq \left(1 - \frac{4}{5r}\right)^{jx}.$$

The union bound over all $J \subseteq [k]$ of size $j$ and $I \subseteq [q+1]$ of size $|I| = j - 1$ then yields

$$\mathbb{P}\left[H_M(U \cap M) \geq K_k\right] \leq \sum_{j \in [k]} \binom{k}{j} \binom{q+1}{j-1} \left(1 - \frac{4}{5r}\right)^{jx} \leq \sum_{j \in [k]} (k(q+1)e^{-\frac{4}{5r}x})^j.$$

Since $k(q+1) \leq \frac{2k^2x6\sqrt{x}}{z} = 12 \cdot 153/2 k^2(\log k)^{1/2}(r \log r)^{3/2} < (kr)^3$ holds for large enough $k$ and every $r \geq 2$, we can further bound again using $k$ sufficiently large:

$$\sum_{j \in [k]} (k(q+1)e^{-\frac{4}{5r}x})^j \leq \sum_{j \in [k]} \exp \left[j \left(3 \log kr - x \frac{4}{5r}\right)\right] \leq \sum_{j \in [k]} \exp \left[-7.65j \log k \log r\right] \leq k \exp \left[-7.65 \log r \log k\right] < k^{-2.1 \log (er)}. $$

This finishes the proof. \qed
4 $K_{k+1}$-freeness

The goal of this section is to prove Lemma 2, which we do at the end. Let $G = \bigcup G_i$ be the union of the $r$ edge disjoint graphs $G_1, \ldots, G_r$ we constructed on the vertex set $V$. That is, $E(G) = \bigcup_{M \in \widetilde{M}} E(H_M)$.

We start with bounding the probability that a given subset $T \subseteq V$, which is crossing with respect to the partition $(Z_v)_{v \in \mathcal{P}}$, induces a clique of some $G_i$. Let us denote by $\mathcal{M}(T)$ the family of those blown-up lines $M$ whose corresponding random graph $H_M$ might contribute an edge to a potential clique on $T$. That is, formally let $\mathcal{M}(T) := \{M \in \widetilde{M} : |M \cap T| \geq 2\}$. (Note that if $T$ is not crossing, then it does not induce a clique in any of the $G_i$, since the sets $Z_v$ contain no edge of $G$.)

Lemma 6. For every subset $T \subset V$ which is crossing with respect to $(Z_v)_{v \in \mathcal{P}}$ we have

$$\mathbb{P}[G_i[T] \text{ is a clique for some } i] \leq \left(\frac{1}{r}\right)^{|\mathcal{M}(T)|} \cdot \left(\frac{1}{r}\right)^{|\mathcal{M}(T)|-1}.$$  

Proof. The induced graph $G_i[T]$ is a clique for some $i$ if and only if $G[T]$ is a clique and there is an index $i$, such that all blown-up lines in $\mathcal{M}(T)$ are selected into $\mathcal{M}_i$. Since the presence of edges in $G$ depends only on the random partitions within the blown-up lines, and this is independent from the random choice of the partition $\widetilde{M} = \bigcup \mathcal{M}_i$, we have that

$$\mathbb{P}[G_i[T] \text{ is a clique for some } i] = \mathbb{P}[G[T] \text{ is a clique}] \cdot \left(\frac{1}{r}\right)^{|\mathcal{M}(T)|-1}. \quad (10)$$

Since every crossing pair of vertices $a, b$ is contained in a unique blown-up line $M_{a,b}^*$, the induced graph $G[T]$ is a clique if and only if for all $x, y \in T$ we have $M_{x,y}^* \in \widetilde{M}$ and for all $M \in \mathcal{M}(T)$ the subgraphs $H_M[M \cap T] = G[M \cap T]$ are cliques. Since the graphs $G[M \cap T]$ are edge-disjoint for different $M$ and are chosen independently, we have that

$$\mathbb{P}[G[T] \text{ is a clique}] \leq \prod_{M \in \mathcal{M}(T)} \mathbb{P}[G[M \cap T] \text{ is a clique}]. \quad (11)$$

The graph $G[M \cap T]$ is a clique if and only if $M \cap T$ is crossing with respect to the random partition $(R_i)_{i \in [k]}$. In particular if $G[M \cap T]$ is a clique then
the uniform random set $\bigcup_{i=1}^{k} R_i$ of size $kx$ should contain the fixed set $M \cap T$. Because of uniformity we can instead calculate the probability that a fixed subset of size $kx$ of $M$ contains a uniformly chosen $|M \cap T|$-element subset. By the choice of our parameters $kx$ represents at most $1/y$-fraction of $M$. Hence the probability that all $|M \cap T|$ elements of the random subset are from the fixed $kx$-element subset is less than $(1/y)^{|M \cap T|}$.

\[ P\left[G[M \cap T] \text{ is a clique}\right] \leq P\left[M \cap T \subseteq \bigcup_{i=1}^{k} R_i\right] \leq \left(\frac{1}{y}\right)^{|M \cap T|} \]

for any $M \in \mathcal{M}(T)$.

Together with (10) and (11), this concludes the proof of the lemma. \qed

Motivated by the estimate of the previous lemma, we classify the $(k+1)$-element subsets $S \subseteq V$ that are crossing with respect to the fixed partition $(Z_v)_{v \in P}$, based on how large $\sum_{M \in \mathcal{M}(S)} |M \cap S|$ is.

Let $\gamma = \log^2 k$. A subset $S \subseteq V$ of size $(k+1)$, which is crossing with respect to the partition $(Z_v)_{v \in P}$, is called of Type 1 if $\sum_{M \in \mathcal{M}(S)} |S \cap M| \geq (k+1)\gamma$, and is called of Type 2 otherwise.

Let $A_\geq (A_\prec)$ denote the event that there is a subset $S \subseteq V$ of Type 1 (of Type 2), such that $G_i[S] \cong K_{k+1}$ for some $i \in [r]$.

Claim 1. We have $P(A_\geq) < 1/4$.

Proof. We prove a stronger statement and bound by $1/4$ the probability that there is a $(k+1)$-subset $S$ of Type 1 that induces a clique in the graph $G = \bigcup_i G_i$. Using Lemma 6 and the union bound we obtain that the probability of this is at most

\[ \binom{|V|}{k+1} \left(\frac{1}{y}\right)^{(k+1)\gamma} < \exp\left(6(k+1) \log kr - (k+1)\gamma \log y\right). \]

Here we used that $|V| \leq k^3r^6$ when $k$ is large enough and $r \geq 2$. By the choice of $\gamma = \log^2 k$ and $y \geq \sqrt{r}$, this probability is less than $1/4$ for large enough $k$ and $r \geq 2$. \qed

Next we deal with subsets of Type 2.

Claim 2. We have $P(A_\prec) < 1/4$. 

14
For the proof of the claim we first establish that any set of Type 2 which is relevant must have at least half of it contained in a single blown-up line.

**Lemma 7.** Let $S$ be a $(k + 1)$-element set of Type 2, which is crossing with respect to $(Z_v)_{v \in \mathcal{P}}$ and for every $x, y \in S$ we have $M_{x,y}^* \in \hat{\mathcal{M}}$. Then there exists a blown-up line $M \in \mathcal{M}(S)$ such that $|M \cap S| \geq \frac{k + 1}{2}$.

**Proof.** The following simple fact will be useful.

**Observation 1.** Let $H$ be a hypergraph on the vertex set $W$ such that for any vertex $u$ the union of the hyperedges containing $u$ is the whole $W$. Then there is a hyperedge containing at least $\frac{|W|^2}{\sum_{K \in E(H)} |K|}$ vertices.

**Proof.** By averaging there must be a vertex $u \in W$ which is contained in at most $\frac{\sum_{K \in E(H)} |K|}{|W|}$ hyperedges. As the edges containing $u$ cover all of $W$ one of them must be of size at least

\[
\frac{|W|}{\sum_{K \in E(H)} |K|} = \frac{|W|^2}{\sum_{K \in E(H)} |K|}.
\]

To prove Lemma 7, we fix a blown-up line $M_1 \in \mathcal{M}(S)$ which maximizes $|M \cap S|$ among all $M \in \mathcal{M}(S)$. We will show that $|M_1 \cap S| \geq \frac{k + 1}{2}$. First, to establish a weaker bound on $|M_1 \cap S|$, we apply Observation 1 for the hypergraph with edge set \{ $M \cap S : M \in \mathcal{M}(S)$ \} on vertex set $S$. Note that for every $u, a \in S$ there exists an $M \in \mathcal{M}(S)$ with $u, a \in M$, hence the hyperedges incident to any vertex $u$ cover all of $W$ one of them must be of size at least

\[
|M_1 \cap S| > \frac{|S|^2}{(k + 1)\gamma} = \frac{k + 1}{\gamma}.
\]

Suppose now for a contradiction that $|M_1 \cap S| < \frac{k + 1}{2}$. Then the set $S' = S \setminus M_1$ has size larger than $\frac{k + 1}{2}$, and we consider the hypergraph with edge set \{ $M \cap S' : M \in \mathcal{M}(S)$ \} with vertex set $S'$. We apply Observation 1 to this hypergraph, noting that the presumption of the claim is satisfied for the same reason as above. We conclude that there is a set $M_2 \in \mathcal{M}(S)$, such that

\[
|M_2 \cap S'| \geq \frac{|S'|^2}{(k + 1)\gamma} > \frac{k + 1}{4\gamma}.
\]
Note that $|M_1 \cap M_2 \cap S| \leq 1$ since $S$ is crossing with respect to $(Z_v)_{v \in \mathcal{P}}$ and $M_1 \cap M_2 = Z_v$ for some $v \in \mathcal{P}$. If existent, let $m$ denote the point in $M_1 \cap M_2 \cap S$.

Consider any pair $(a, b)$ of points such that $a \in (M_1 \cap S) \setminus \{m\}$ and $b \in (M_2 \cap S')$. As $S$ is crossing with respect to $(Z_v)_{v \in \mathcal{P}}$ and $a, b \in S$, there is a unique blown-up line $M_{a,b}$ containing the two points $a$ and $b$, which therefore satisfies $M_{a,b} \in \mathcal{M}(S)$. In the following we show that these lines are all distinct and use it to derive a contradiction to $\sum_{M \in \mathcal{M}(S)} |M \cap S| < (k + 1)\gamma$.

Let $(a, b)$ and $(c, d)$ be two such pairs of vertices, that is $a, c \in (M_1 \cap S) \setminus \{m\}$, and $b, d \in M_2 \cap S'$. We show that if $M_{a,b}^* = M_{c,d}^*$ then $(a, b) = (c, d)$. Note that $M_{a,b}^* = M_{c,d}^*$ implies $a, b, c, d \in M$. If $a \neq c$ then, due to $a, c \in M_1 \cap M$ and the uniqueness of $M_{a,c}^*$, we have $M_1 = M_{a,c}^* = M \ni b, d$. However, $b, d \in S'$ as well, which is a contradiction since $S' = S \setminus M_1$ is disjoint from $M_1$. Thus $a = c$. If $b \neq d$ then, due to $b, d \in M_2 \cap M$ and the uniqueness of $M_{b,d}^*$, we have $M_2 = M_{b,d}^* = M \ni a$. However, $a \in (M_1 \cap S) \setminus \{m\}$ which is disjoint from $M_2$. This again yields a contradiction, implying that $b = d$. Hence $M_{a,b}^* \neq M_{c,d}^*$ for distinct pairs $(a, b) \neq (c, d)$, so the number of lines in $\mathcal{M}(S)$ is at least

$$((|M_1 \cap S| - 1) \cdot |M_2 \cap S'| \geq \left(\frac{k + 1}{\gamma} - 1\right) \frac{k + 1}{4\gamma}.$$  

The contribution of each of these lines to the sum $\sum_{M \in \mathcal{M}(S)} |M \cap S|$ is at least two, so

$$\sum_{M \in \mathcal{M}(S)} |M \cap S| \geq 2 \cdot \left(\frac{k + 1}{\gamma} - 1\right) \frac{k + 1}{4\gamma}.$$  

This is greater than $(k + 1)\gamma$ for large enough $k$, which is a contradiction to $S$ being of Type 2. Therefore $|M_1 \cap S| \geq \frac{k + 1}{2}$ and the lemma follows. \hfill $\square$

With Lemma 7 established, let us return to the proof of Claim 2.

**Proof of Claim 2.** We claim that the event $A_<$ (i.e., that there exists a Type 2 subset $S \subseteq V$ inducing a $K_{k+1}$ in one of the $G_i$) is only possible if the event $A^*$ that

there is a blown-up line $M \in \hat{\mathcal{M}}$, with a subset $T \subseteq M$ of size $t = \left\lceil \frac{k}{2} \right\rceil$, which is crossing with respect to $(Z_v)_{v \in \mathcal{P}}$, and a vertex $x \in V \setminus M$, such that $T \cup \{x\}$ induces a clique in one of the $G_i$, 

16
also holds. Indeed, the special blown-up line $M$ is delivered by Lemma 7 with $T$ being any $t$-subset of $M \cap S$. The appropriate vertex $x \in V \setminus M$ then must exist because $G_i[S] \cong K_{k+1}$, but $M$ does not contain a $(k+1)$-clique by construction.

We will bound $\mathbb{P}(A_<)$ via bounding $\mathbb{P}(A^*)$. Let us fix a crossing subset $T$ of size $t$ and a vertex $x \in V$, such that $T$ is contained in some blown-up line $M \in \hat{M}$ for which $x \notin M$. We estimate, using Lemma 6, the probability that $T' = T \cup \{x\}$ induces a clique in $G_i$ for some $i$. Note that the unique blown-up lines $M_{x,w} \in \mathcal{M}(T')$ containing the pairs $\{x,w\}$ are pairwise distinct for distinct $w \in T$ and they are also different from $M$ (otherwise two distinct elements of $T$ would be covered by two different lines). In particular, we have $\sum_{M \in \mathcal{M}(T')} |M \cap T'| \geq t + 2t = 3t$. Since $|\mathcal{M}(T')| \geq t + 1$, Lemma 6 implies that $\mathbb{P}\left[G_i[T \cup \{x\}] \cong K_{t+1} \text{ for some } i \right] \leq \frac{1}{y^t} \cdot \frac{1}{t^t}$.

We estimate $\mathbb{P}(A^*)$ with the union bound. The number of choices for $T' = T \cup \{x\}$ is at most $|\hat{M}||V| (q+1)^t z^t |V|$, as first one chooses $M = M(L)$, then a $t$-element subset of its $q + 1$ partition classes from $(Z_v)_{v \in L}$, then a vertex from each of the selected $t$ partition classes (each of size $z$), and finally the vertex $x \in V \setminus M$.

Using $(q+1)z \leq 2kxy$, $|\hat{M}||V| \leq |P|^2 z = (q^2 + q + 1)^2 z \leq k^5 r^6 \log^6 r$ for large enough $k$, $y = 6\sqrt{x}$, and $t \geq \frac{k}{2}$ we obtain that

$$\mathbb{P}(A_<) \leq \mathbb{P}(A^*) \leq |\hat{M}||V| \binom{q+1}{t} z^t \cdot \left( \frac{1}{y^3 r} \right)^t \leq k^5 r^6 \log^6 r \left( \frac{e(q+1)z}{ty^3} \right)^t \cdot \frac{1}{t^t} \leq k^5 \left( \frac{2e k y x}{2} \cdot \frac{r^6 \log^6 r}{r^t} \right)^t \leq k^5 \left( \frac{4e}{36} \right)^{k/2} \cdot \frac{r^6 \log^6 r}{r^{k/2}} < \frac{1}{4},$$

for every $k$ large enough and $r \geq 2$.

**Proof of Lemma 2.** Note that by construction the sets $Z_v$ contain no edge of any $G_i$, $i \in [r]$. Hence a $(k + 1)$-clique of a $G_i$ can only be induced by a subset that is crossing with respect to the partition $(Z_v)_{v \in P}$. Such subsets can be either of Type 1 or Type 2, so the probability that some $G_i$ contains a clique of size $k + 1$ is bounded by $\mathbb{P}(A_>) + \mathbb{P}(A_<)$. Hence Claim 1 and Claim 2 imply Lemma 2. \qed
5 Concluding remarks and open questions

In this paper we gave an upper bound on the function \( s_r(K_{k+1}) = P_r(k) \), formulated in terms of the color pattern (density Folkman) number \( CP_r(1/r, k) \). The bound is of the form

\[
s_r(K_{k+1}) = P_r(k) \leq CP_r(1/r, k) \leq C_r(k \log k)^2
\]

and is tight up to the logarithmic factor. Furthermore, it improves the upper bound of [7] in the range when the number \( r \) of colors is not too large compared to the order \( k \) of the clique.

The function \( CP_1(\cdot, k) \) is very closely related to the Folkman number \( F(\cdot, k, k + 1) \), and as there is only one graph involved in the pattern, we prefer to refer to \( CP_1(\cdot, k) \) as the density Folkman number. Via the relation from [3], i.e.,

\[
F(r, k, k + 1) \leq CP_1(1/r, k) \leq CP_r(1/r, k)
\]

our bound from above also yields a \((\log^2 k)\)-factor improvement on the known upper bound for \( F(r, k, k + 1) \).

1. A non-trivial lower bound? It is already mentioned in [3], that the vertex Folkman number \( F(r, k, k + 1) \) lacks any non-trivial (i.e. non-linear) lower bound. In fact there is no better lower bound known for the density Folkman number either, while the upper bound on \( CP_1(\frac{1}{r}, k) \) obtained in this paper is super-quadratic. Although we were unable to prove any nontrivial lower bound, we tend to think that the upper bound is closer to the true order of magnitude and make the following conjecture.

**Conjecture 1.** For any fixed \( r \geq 2 \) and any fixed \( \varepsilon > 0 \) we have \( CP_1(\frac{1}{r}, k) \geq k^{2-\varepsilon} \) provided \( k \) is sufficiently large.

In fact, it would already be interesting to show, e.g., that \( CP_1(\frac{1}{2}, k) \geq 1000k \). In other words, any \( K_{k+1} \)-free graph on 1000k vertices must have a vertex subset of size 500k which \( K_k \)-free. Recall that by [2], we have \( CP_2(\frac{1}{2}, k) \geq P_2(k) = k^2 \). This is in striking contrast with our lack of knowledge for Conjecture 1.
2. Separation of the two and three color case? The monotonicity of $s_r(K_k)$ in the number of colors is immediate once we know $s_r(K_{k+1}) = P_r(k)$, hence $s_3(K_{k+1}) = s_2(K_{k+1}) = k^2$. Is there something stronger possible? We believe the answer is yes.

**Conjecture 2.** $s_3(K_{k+1}) \gg k^2$ as $k \to \infty$.

A somewhat weaker version of this would be the following.

**Conjecture 3.** There exists fixed $\alpha > 0$ and an integer $\ell \in \mathbb{N}$ such that $CP_\ell(\alpha, k) \gg k^2$ as $k \to \infty$.

3. Asymptotic equality of the two types of Folkman numbers? The density Folkman number $CP_1(1/r, k)$ serves as an obvious upper bound to the vertex Folkman number $F(r, k, k+1)$. In fact, all known upper bounds for $F(r, k, k+1)$ are derived by finding a monochromatic $K_k$ in the largest color class, hence they also serve as upper bounds for $CP_1(1/r, k)$. It would be interesting to decide whether these two functions have the same order of magnitude. We put forward the following conjecture.

**Conjecture 4.** $F(r, k, k+1) = \Theta \left( CP_1 \left( \frac{1}{r}, k \right) \right)$.

References


