Recent Progress on the Erdős-Rogers Functions

Haocong (Jim) Li

Advisor: Prof. Jacques Verstraëte

Department of Mathematics

University of California, San Diego

In partial fulfillment of the requirements

for the Mathematics Honors Program

June 2024

Acknowledgements

I hereby express my greatest gratitude to Prof. Jacques Verstraëte for the guidance and direction of this thesis. Prof. Verstraëte inspired me to find and commit my interest towards graph theory and related topics. He also inspired me to set a paradigm for a life as a scholar. I will also thank former UCSD student Xun Gong and current UCSD students Siyuan (Samuel) Zheng and Chen-an (Jack) Chou for latex helps. Note that this is not an exhausting list - I will thank everyone that explained to me mathematical concepts during my mathematical studies and everyone that helped me on my life in general.

Contents

1 Introduction

1.1 Notations

Let $G = (V, E)$ be a graph. We denote $V := V(G)$ to be the set of vertices of G, $v(G) := |V(G)|$; we denote $E := E(G)$ to be the set of edges of G, $e(G) = |E(G)|$. We say $x \in G$ if x is a vertex of G, and we say $X \subseteq G$ if X is a subset of $V(G)$. For $X \subseteq G$, let $G[X]$ be the induced subgraph on G with vertex set X.

1.2 The Erdős-Rogers Function

Definition 1.2.1 (The Erdős-Rogers function)**.** Let F, G be some graphs, and let H be a G-free graph on n vertices. We define $f(H, F, G) := \max\{|K| : K \subseteq H,\ H[K] \text{ is } F\text{-free}\}$. We then define the Erdős-Rogers function $f(n, F, G)$ to be as follows:

$$f(n, F, G) = \min_{H \text{ is } G\text{-free}, |H|=n} f(H, F, G) = \min_{H \text{ is } G\text{-free}, |H|=n} \max\{|K| : K \subseteq H,\ H[K] \text{ is } F\text{-free}\}. \tag{1.2.1}$$

For a brief overview of the progress of the Erdős-Rogers function in the last three decades, one may refer to the introduction part of [10] for more detail. The purpose of this paper is to inspect how tools from entropy, finite geometry and arithmetic progressions paves the ways towards the probabilistic methods that will be used to deduce the results of certain Erdős-Rogers functions.

2 Lower Bound on $f(n, K_s, K_{s+1})$ for $s \geq 3$

Remark: Need to show that $c(r)$ is bounded above!!! There is such a lack of rigor!!! in this section, we will prove the best known lower bound on $f(n, K_s, K_{s+1})$ for $s \geq 3$ through arguments on independent sets of K_{s+1}-free graphs. Through the results from Shearer in 1995, we will deduce the following lower bound:

Theorem 2.0.1. For any $s \geq 3$, $f(n, K_s, K_{s+1}) = \Omega(\sqrt{n \log n / \log\log n})$

To prove this theorem, we need the next theorem by Shearer:

Theorem 2.0.2 (Shearer 1995 [19])**.** For any $r \geq 4$, Let G be a K_r-free graph with the size of vertices $|V(G)| = n$ with $d = \Delta(G)$, the maximum degree over all vertices in G. Let $\alpha(G)$ denote the maximum size over all independent sets of G. Then,

$$\alpha(G) = \Omega(\frac{n \log d}{d \log\log d}). \tag{2.0.1}$$

Proof of Theorem 2.0.1: For any K_{s+1}-free graph G choose $v \in V(G)$ such that $|N(V)| = d$. Notice that $N(V)$ is K_s-free. Since an independent set must be K_s-free for any $s \geq 2$,

$$f(n, K_s, K_{s+1}) = \Omega(\min_d(\max\{d, \frac{n \log d}{d \log\log d}\})). \tag{2.0.2}$$

Note that the minimum occurs precisely (up to constant factors) when $d = \frac{n \log d}{d \log\log d}$, i.e., $d = \sqrt{n \log d / \log\log d}$, as $\frac{\log d}{d \log\log d}$ is a decreasing function with respect to d. So $d \in [\sqrt{n}, \sqrt{n \log n}/\sqrt{\log\log n}]$ as $\frac{\log d}{\log\log d}$ is an increasing function whose image is in $[1, \frac{\log n}{\log\log n}]$. Since $\frac{\log d}{d \log\log d}$ is a decreasing function with respect to d, and when $d = \sqrt{n \log n / \log\log n}$, $\frac{n \log d}{d \log\log d} = \frac{n \log n}{\log\log n} \cdot \sqrt{\frac{\log\log n}{n \log n}} = d$ up to constant factors. Therefore, $d = \sqrt{n \log n / \log\log n}$ gives the desired leading-term equality, which implies the theorem. □

2.1 A Short Introduction to Entropy

This section is a short exposition on the essences of entropy that will be enough to understand Shearer's proof of Theorem 2.0.2. The section itself is condensed, with lots of definitions and short theorems in [14], and the main observation is lemma 2.1.15.

Definition 2.1.1. Let $(\Omega, \mathcal{F})$ be a finite probability space, i.e. $\Omega = \{\omega_1, ..., \omega_n\}$ for some $n \in \mathbb{N}$, $\mathcal{F} = P(\Omega)$ is the set of subsets of Ω. Let $X : \Omega \to \Xi$ be a discrete random variable with distribution p_x and support $S = \{x \in \Omega : p_X(x) > 0\}$. Then $H(X)$, the entropy of X is defined as follows:

$$H(X) := \sum_{x \in S} p_X(x) \log(\frac{1}{p_X(x)}) = \sum_{x \in S} -p_X(x) \log(p_X(x)). \tag{2.1.1}$$

Remark 2.1.2. We extend the definition of $H(X)$ on Ω by putting $p_X(x) \log(1/p_X(x))$ to be 0 for all $x \notin S$.

Remark 2.1.3. We can similarly define the entropy of X (with the properties defined above) with different bases of logarithm. For example, the $\log 2$ or $\log_2$ entropy of X, denoted $H_2(X)$ is defined to be:

$$H_2(X) := \sum_{x \in S} p_X(x) \log_2(\frac{1}{p_X(x)}) = \sum_{x \in S} -p_X(x) \log_2(p_X(x)) = \frac{H(X)}{\log 2}. \tag{2.1.2}$$

Theorem 2.1.4. Let X be a discrete random variable with finite support S with $|S| = k$, then $0 \leq H(X) \leq \log k$, and $H(X) = \log k$ iff $p_X(x) = 1/k \ \forall x \in S$.

Proof. By Jensen's Inequality, $H(X) = \sum_{x \in S} p_X(x) \log p_X(x) \leq \log(\sum_{x \in S} p_X(x)/p_X(x)) = \log k$, and equality holds iff $p_X(x) = 1/k \ \forall x \in S$. Also, $H(X) = \sum x \in S - p_X(x) \log(p_X(x))$ and since $p_X(x) \leq 1$, $-\log(p_X(x)) \geq 0$ so the lower bound follows. □

Definition 2.1.5. For random variables X and Y with supports $S \subseteq \Omega_1$ and $T \subseteq \Omega_2$, the conditional entropy of X given $Y = y$ is

$$H(X|Y = y) = \sum_{x \in S} p_{X|Y=y}(x) \log(\frac{1}{p_{X|Y=y}(x)}) \tag{2.1.3}$$

where $p_{X|Y=y}$ is the distribution X conditioned on $Y = y$. The conditional entropy, denoted $H(X|Y)$, is

$$H(X|Y) = \sum_{y \in T} p_Y(y) H(X|Y = y) = \sum_{x \in S,\, y \in T} p_{X,Y}(x, y) \log(\frac{1}{p_{X|Y=y}(x)}). \tag{2.1.4}$$

Corollary 2.1.6. $H(X|X) = 0$.

Proof of Corollary.

$$\begin{aligned} H(X|X) &= \sum_{x \in S} p_X(x) H(X|X = x) \\ &= \sum_{x \in S} p_X(x) \sum_{y \in S} p_{X|X=x}(y) \log\left(\frac{1}{p_{X|X=x}(y)}\right) \\ &= \sum_{x} 0 = 0. \end{aligned} \tag{2.1.5}$$

□

Definition 2.1.7. the mutual information between X and Y, denoted $I(X;Y)$, is

$$I(X;Y) := H(X) - H(X|Y) = E_{X,Y}[\log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}] \tag{2.1.6}$$

Remark 2.1.8. $I(X;Y)$ is reflexive by the above definition.

Remark 2.1.9. $I(X;X) = H(X)$ by Corollary 2.1.6.

Definition 2.1.10. The random variable $\hat{X} = (X_1, ..., X_n)$ where $X_i : \Omega_i \to \mathbb{R}$ is called a random vector (with dimension n), and the distribution associated to $\hat{X}$ is $p_{\hat{X}}(\hat{x}) = p_{(X_1,...,X_n)}((x_1, ..., x_n))$ where $\hat{x} = (x_1, ..., x_n)$.

Definition 2.1.11. For random vectors $\hat{X} = (X_1, ..., X_n)$ and $\hat{Y} = (Y_1, ..., Y_n)$ of the same dimensions, $(\hat{X}, \hat{Y})$ is memoryless if for $\hat{x} = (x_1, ..., x_n)$, $\hat{y} = (y_1, ..., y_n)$,

$$p_{\hat{Y}|\hat{X}=\hat{x}}(\hat{y}) = \prod_{i=1}^{n} p_{Y_i|X_i=x_i}(y_i). \tag{2.1.7}$$

Remark 2.1.12. $(\hat{X}, \hat{X})$ is memoryless since

$$p_{\hat{X}|\hat{x}}(\hat{y}) = \begin{cases} 1 & \text{if } y_i = x_i \ \forall i; \\ 0 & \text{otherwise.} \end{cases} = \prod_{i=1}^{n} p_{X_i|x_i}(y_i). \tag{2.1.8}$$

Theorem 2.1.13. For n-dimensional memoryless random vectors $\hat{X}$ and $\hat{Y}$,

$$I(\hat{X}, \hat{Y}) \leq \sum_{i=1}^{n} I(X_i; Y_i). \tag{2.1.9}$$

Proof. Note that

$$I(\hat{X}, \hat{Y}) = E_{\hat{X},\hat{Y}}[\log \frac{p_{\hat{X},\hat{Y}}(\hat{x}, \hat{y})}{p_{\hat{X}}(\hat{x}) p_{\hat{Y}}(\hat{y})}] \tag{2.1.10}$$

and

$$\begin{aligned}
\sum_{i=1}^{n} I(X_i;Y_i) &= \sum_{\hat{x},\hat{y}} E_{X_i,Y_i}[\log \frac{p_{Y_i|X_i=x_i}(y_i)}{p_{Y_i}(y_i)}] \\
&= E_{\hat{X},\hat{Y}}[\log \frac{\prod_{i=1}^{n} p_{Y_i|X_i=x_i}(y_i)}{\prod_{i=1}^{n} p_{Y_i}(y_i)}] \\
&= E_{\hat{X},\hat{Y}}[\log \frac{p_{\hat{Y}|\hat{X}=\hat{x}}(\hat{y})}{\prod_{i=1}^{n} p_{Y_i}(y_i)}] \\
&= E_{\hat{X},\hat{Y}}[\log \frac{p_{\hat{X},\hat{Y}}(\hat{x},\hat{y})}{p_{\hat{X}}(\hat{x})\prod_{i=1}^{n} p_{Y_i}(y_i)}]
\end{aligned} \tag{2.1.11}$$

So

$$\begin{aligned}
I(\hat{X},\hat{Y}) - \sum_{i=1}^{n} I(X_i;Y_i) &= E_{X_i,Y_i}[\log \frac{p_{\hat{X},\hat{Y}}(\hat{x},\hat{y})}{p_{\hat{X}}(\hat{x})p_{\hat{Y}}(\hat{y})} \cdot \frac{p_{\hat{X}}(\hat{x})\prod_{i=1}^{n} p_{Y_i}(y_i)}{p_{\hat{X},\hat{Y}}(\hat{x},\hat{y})}] \\
&= E_{X_i,Y_i}[\log \frac{\prod_{i=1}^{n} p_{y_i}(y_i)}{p_{\hat{Y}}(\hat{y})}] \\
&= E_{Y_i}[\log \frac{\prod_{i=1}^{n} p_{y_i}(y_i)}{p_{\hat{Y}}(\hat{y})}] \\
\text{(Jensen's Inequality)} &\leq \log(E_{Y_i}[\frac{\prod_{i=1}^{n} p_{y_i}(y_i)}{p_{\hat{Y}}(\hat{y})}]) \\
&= \log(\sum_{\hat{y}} \prod_{i=1}^{n} p_{Y_i}(y_i)) \\
&= 0 \text{ by the observation that } \sum_{\hat{y}} \prod_{i=1}^{n} p_{Y_i}(y_i)) = 1.
\end{aligned} \tag{2.1.12}$$

Hence, theorem follows. □

Corollary 2.1.14. For $\hat{X} = (X_1, ..., X_n)$, $H(\hat{X}) \leq \sum_{i=1}^{n} H(X_i)$.

Proof. The corollary follows from Remark 2.1.9. □

Lemma 2.1.15 (Kleitman, Shearer, and Sturtevant [12])**.** Let $F \subseteq P([n])$ be a collection of distinct subsets of $[n]$, where $i \in [n]$ occurs in a proportion α_i over all elements in F. Then

$$\log |F| \leq \sum_{i=1}^{n} H(\alpha_i), \tag{2.1.13}$$

where $H(\alpha) = -\alpha \log \alpha - (a - \alpha)\log(1-\alpha)$ for all $\alpha \in (0, 1]$.

Proof. let $F := \{S_1, ..., S_r\}$ be such a collection. let p be a probability measure on F such that $p(S_j) = 1/r$ for all $j \in [r]$. $\forall i \in [n]$, let $X_i : F \to \{0, 1\}$ be a random variable where $\forall j \in [r]$,

$$X_i(S_j) = \begin{cases} 1 & \text{if } i \in S_j \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.14}$$

So $p(X_i^{-1}(1)) = \alpha_i$ for all i, which implies $H(X_i) = H(\alpha_i)$. Now let $S = (X_1, ..., X_n)$ be a random vector. By equation 2.1.14, $S : F \to \{0,1\}^n$ is injective. So S is uniformly distributed with the probability measure p. Therefore,

$$H(S) = \sum_{i=1}^{r} p(S_i) \log \frac{1}{p(S_i)} = r \cdot \frac{1}{r} \cdot \log r = \log r = \log |F|. \tag{2.1.15}$$

By Corollary 2.1.14,

$$\log |F| = H(S) \leq \sum_{i=1}^{n} H(X_i) = \sum_{i=1}^{n} H(\alpha_i). \tag{2.1.16}$$

Lemma follows. □

2.2 Proof of Theorem 2.0.2

Lemma 2.2.1 (Shearer [19])**.** Let G be a K_r-free graph with $r \geq 3$, $I(G)$ be the set of independent sets of G, and $\bar{\alpha}(G)$ be the average size of independent sets in G. Then as $|I(G)| \to \infty$,

$$\bar{\alpha}(G) = \Omega \frac{\log |I(G)|}{\log \log |I(G)|}. \tag{2.2.1}$$

Proof. Let $m := |V(G)|$, $k = \alpha(G)$, $\phi = \frac{\bar{\alpha}(G)}{m}$. Then,

Claim 1. We have the following three properties:

(1) $\bar{\alpha}(G) = m\phi$,

(2) $|I(G)| \geq 2^k$,

(3) $|I(G)| \leq 2^{mH_2(\phi)}$.

Proof of Claim. Note that (1) and (2) follows immediately by definition. For (3), denote $V(G) := \{y_1, ..., y_m\}$ let $I(G) := \{S_1, ..., S_\gamma\}$. let p be a probability measure on F such that $p(S_j) = 1/\gamma$ for all $j \in [\gamma]$. $\forall i \in [m]$, let $X_i : I(G) \to \{0,1\}$ be a random variable where $\forall j \in [\gamma]$,

$$X_i(S_j) = \begin{cases} 1 & \text{if } y_i \in S_j \\ 0 & \text{otherwise.} \end{cases} \tag{2.2.2}$$

Now let $X := (X_1, ..., X_m)$ be the random vector with the uniform distribution. Then by lemma 2.1.15,

$$\log |I(G)| \leq \sum_{i=1}^{m} H(X_i) = \sum_{i=1}^{m} H(\alpha_i) \leq_{(*)} mH(\frac{\sum_{i=1}^{m} \alpha_i}{m}) =_{(**)} mH(\frac{\bar{\alpha}(G)}{m}), \tag{2.2.3}$$

where $(*)$ follows from the observation that $\sum_{i=1}^{m} |I(G)|\alpha_i$ counts the sum of occurrences of each vertex over all independent sets of S, which equals $|I(G)|\bar{\alpha}(G)$, and $(**)$ follows from the fact that $f(x) := -x \log x - (1-x) \log(1-x)$ is concave on $(0,1)$ as by calculus methods, its second derivative is

$$f''(x) = -\frac{1}{1-x} - \frac{1}{x}. \tag{2.2.4}$$

Dividing both sides of the equation 2.2.3 by $\log 2$ and exponentiating over 2, we get $|I(G)| \leq 2^{mH_2(\bar{\alpha}(G)/m)} = 2^{mH_2(\phi)}$. □ (claim)

By the above claim we have

$$\bar{\alpha}(G) = m\phi = mH(\phi) \cdot \frac{\phi}{H(\phi)} \geq \frac{\phi}{H(\phi)} \cdot \frac{\log |I(G)|}{\log 2}. \tag{2.2.5}$$

So it remains to find a lower bound for $\phi/H(\phi)$ in terms of $|I(G)|$. By calculus methods (i.e. on value of the function and its first and second derivative), $\phi/H(\phi)$ is increasing, so it's enough to find a lower bound for ϕ and replace $H(\phi)$ with a larger value when ϕ reaches its lower bound. Since $mH_2(\phi) \geq \log_2 |I(S)| \geq k$, $H(\phi) \geq k/m$.

Claim 2. $\forall r \geq 3$ and K_r-free graph G with vertex size $|V(G)| = m$, $\alpha(G) = k := k_r(m) \geq m^{\frac{1}{r-1}}$.

Proof of Claim. Let $d = \Delta(G)$, the maximum degree over all vertices in G. If $r = 3$, then by Turan's theorem (Caro-Wei bound [1]), which states that for all H with $|V(H)| = n$, $\Delta(H) = d$, then $\alpha(H) \geq \frac{n}{d+1}$ $k \geq \max(d, m/d) \geq \min_d(\max(d, \frac{m}{d+1}))$. Note that $\max(d, \frac{m}{d+1})$ obtains minimum when $d = \frac{m}{d+1}$, which implies $d = m^{1/2} = m^{1/(3-1)}$. So claim true for $r = 3$. Now suppose the claim is true for $r = 1, ..., n-1$. For $r = n$, let d be defined as above, then $k \geq \max(k_{n-1}(d), m/d) \geq \max(d^{1/(n-2)}, m/d)$, which obtains minimum when $d^{1/(n-2)} = m/d \Rightarrow d = m^{\frac{n-2}{n-1}}$, which gives $k \geq m^{1-(n-1)}$. □ (claim)

So $H(\phi) \geq m^{-\frac{r-2}{r-1}}$ by claim, and we thus have

$$\begin{aligned} -\phi \log \phi - (1-\phi)\log(1-\phi) &\geq m^{-\frac{r-2}{r-1}} \\ \phi(\log \phi - \log(1-\phi)) + \log(1-\phi) &\leq -m^{-\frac{r-2}{r-1}} \\ \phi \log \frac{\phi}{1-\phi} + \log(1-\phi) &\leq -m^{-\frac{r-2}{r-1}} \\ \log \frac{\phi^\phi}{(1-\phi)^{\phi-1}} &\leq -m^{-\frac{r-2}{r-1}} \\ \phi^\phi (1-\phi)^{1-\phi} &\leq e^{-m^{-\frac{r-2}{r-1}}}. \end{aligned} \tag{2.2.6}$$

Note that $g(x) = x^x$ is lower bounded by $1/e$ by a calculus argument (i.e. on value of the function and its first and second derivative), so we have

$$\begin{aligned} \phi^\phi / e &\leq e^{-m^{-\frac{r-2}{r-1}}} \\ \phi \log(e^{-1}\phi) &\leq -m^{-\frac{r-2}{r-1}} \\ \phi \log \frac{1}{e^{-1}\phi} &\geq m^{-\frac{r-2}{r-1}}. \end{aligned} \tag{2.2.7}$$

So for large enough m,

$$\phi \geq \frac{c(m)}{m^{\frac{r-2}{r-1}} \log m}. \tag{2.2.8}$$

for some constant $c(m)$. Notice that if $H(\phi) = m^{-\frac{r-2}{r-1}}$, then the inequality in the equation 2.2.8 still holds. By a calculus argument, $H(\phi)$ is increasing on $(0, e/2)$. So for large enough m and for $\phi = \frac{c(m)}{m^{\frac{r-2}{r-1}} \log m}$, $H(\phi) \leq m^{-\frac{r-2}{r-1}}$, which implies that $\phi/H(\phi) \geq c(m)/\log m$. So

$$\bar{\alpha}(G) \geq \frac{\phi}{H(\phi)} \cdot \frac{\log |I(G)|}{\log 2} \geq \frac{c(m)}{\log m} \cdot \frac{\log |I(G)|}{\log 2}. \tag{2.2.9}$$

Note that $|I(G)| \geq 2^k \geq 2^{m^{1/(r-1)}}$, so $\log m = O(\log \log |I(G)|)$. Hence,

$$\bar{\alpha}(G) = \Omega(\frac{\log |I(G)|}{\log \log |I(G)|}). \tag{2.2.10}$$

□

Theorem 2.2.2 (Shearer [19])**.** For any n-vertex d-regular graph G that is K_r-free for all $r \geq 4$. Let $\bar{\alpha}(G)$ be the average size of an independent set in G. Then for $d = d(n)$ where $d(n) \to \infty$ as $n \to \infty$, $\bar{\alpha}(G) \geq c(r) n \frac{\log d}{d \log \log d}$ for some constant $c(r)$, i.e. $\bar{\alpha}(G) = \Omega(\frac{n \log d}{d \log \log d})$.

Proof. Let p be a uniform distribution of $I(G)$, i.e. $\forall S \subseteq I(G)$, $p(S) = \frac{1}{|I(G)|}$. For all $x \in V(G)$, let $T := N(x)$ denote the neighborhood of x, $p_x := p(\{S \in I(G),\ x \in S\})$, $d\bar{p}_x$ be the average number of neighbors of x being in some element in $I(G)$, $H_x := G[V(G) \backslash \{x \cup N(x)\}]$ be the induced subgraph of G with vertex sets $V(G) \backslash \{x \cup N(x)\}$, and for all $S \subseteq N(x)$ (s can be $\varnothing$), let $f(S)$ be the probability that there exists some independent set F_{H_x} s.t. $E(S, F_{H_x}) = \varnothing$, i.e. there is no edge between S and F_{H_x} and $V(T) \backslash V(S) = N_{G,T}(F_{H_x})$ where $N_{G,T}(F_{H_x})$ denote the neighborhood of vertex sets F_{H_x} under the graph G. Now note that each $F \in I(G)$ is a union of some $F_x \in I(H)$ (F_x can be $\varnothing$) and some unique element in $I(N(x)) \cup \{x\}$. Note that each such F_x has a neighborhood $N_{T,G}(F_x) \subseteq T$ which determines some S_{F_x} defined above. Also note that

$$|I(G)| = \sum_{\varnothing \subseteq F_x \subseteq I(H)} (1 + 1 + |I(S_{F_x})|) \tag{2.2.11}$$

as there are three categories of ways that we can build an independent set in G from F_x, the first one of which is to keep F_x, not adding or deleting any vertex when F_x is nonempty, the second one of which is to add x to F_x as $N(x) \cap F_x = \varnothing$, and the third one of which is to add some independent set of S_{F_x}. Now, by double counting,

$$|I(G)| = \sum_{\varnothing \subseteq F_x \subseteq I(H)} (1 + 1 + |I(S_{F_x})|) = \sum_{\varnothing \subseteq S \subseteq T} |I(H_x)| f(S)(|I(S)| + 1 + 1). \tag{2.2.12}$$

So, since $\sum_{\varnothing \subseteq S \subseteq T} f(S) = 1$,

$$\begin{aligned} p_x &= \frac{|I(G)| - \sum_{\varnothing \subseteq S \subseteq T} |I(H_x)| f(S)(|I(S)| + 1)}{|I(G)|} \\ &= 1 - \frac{1 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|}{2 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|} \\ &= \frac{1}{2 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|}. \end{aligned} \tag{2.2.13}$$

Moreover,

$$d\bar{p}_x = \frac{\sum_{\varnothing \subseteq S \subseteq T} \sum_{F_S \in I(S)} f(S)|F_S|}{2 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|} = \frac{\sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|\bar{\alpha}(S)}{2 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|} \tag{2.2.14}$$

and so

$$\bar{p}_x = \frac{\sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|\bar{\alpha}(S)}{d(2 + \sum_{\varnothing \subseteq S \subseteq T} f(S)|I(S)|)}. \tag{2.2.15}$$

Note that by lemma 2.2.1, there exists some constant $c = c(r-1)$ such that $\bar{\alpha}(S) \geq c(r-1)\frac{\log |I(S)|}{\log \log |I(S)|}$ as $S \subseteq T$ is K_{r-1}-free for all S. Now let $\lambda = \lambda(d)$ be some parameter that will be set later, let $w = \sum_{\varnothing \subseteq S \subseteq T,\, |I(S)| \geq \lambda} f(S)|I(S)|$, $y = c(r-1)\log \lambda / \log \log \lambda$. Then we have $\forall S \subseteq T$ s.t. $|I(S)| \geq \lambda$, $\bar{\alpha}(S) \geq y$, and so

$$p_x \geq \frac{1}{2 + \lambda + w} \quad \text{and} \quad \bar{p}_x \geq \frac{1}{\frac{d}{wy}(2 + \lambda + w)}. \tag{2.2.16}$$

Note that if $yw/d \geq 1$ then $\bar{p}_x \geq p_x$ and $d/yw \leq 1 \Rightarrow \bar{p}_x \geq \frac{1}{2+\lambda+d/y}$; if $yw/d \leq 1$ then $\bar{p}_x \leq p_x$ and $d/yw \geq 1 \Rightarrow p_x \geq \frac{1}{2+\lambda+d/y}$. Let $p_{x,F}$ be the probability that $x \in F$ for a fixed $F \in I(G)$. Then $\forall x \in F$, $p_{x,F} = \frac{1}{|I(G)|}$. Therefore,

$$\begin{aligned} \bar{\alpha}(G) &= \sum_{F \in I(G)} \frac{|F|}{|I(G)|} = \sum_{F \in I(G)} \sum_{x \in F} p_{x,F} = \sum_{x \in G} \sum_{F \ni x} p_{x,F} \\ &= \sum_{x \in G} p_x = \frac{\sum_{x \in I(G)} d\bar{p}_x}{d} = \sum_{x \in G} \bar{p}_x. \end{aligned} \tag{2.2.17}$$

and

$$\begin{aligned} 2\bar{\alpha}(G) &= \sum_{x \in G} (p_x + \bar{p}_x) \geq \sum_{x \in G} \max(p_x, \bar{p}_x) \geq \sum_{x \in G} \frac{1}{2 + \lambda + d/y} \\ &= \frac{n}{2 + \lambda + d/y} = \frac{n}{2 + \lambda + c(r-1)\frac{d \log \log \lambda}{\log \lambda}} \\ &= \frac{n \log \lambda}{\lambda \log \lambda + 2 \log \lambda + c(r-1) d \log \log \lambda}. \end{aligned} \tag{2.2.18}$$

Now let $\lambda = d/\log d$, then up to constant multiple of leading terms by some c',

$$\bar{\alpha}(G) \geq c' n \frac{\log d}{d \log \log d}. \tag{2.2.19}$$

□

Proof of Theorem 2.0.2. For K_r-free ($r \geq 4$) n-vertex graph G, let $D := \Delta(G)$, and let "$d = \delta(G)$, i.e. the minimum degree over all vertices of G. We apply the following process to make a D-regular graph Γ: First we label vertices of G to be $\{v_1, ..., v_n\}$, and next we take $2(D-d)$ disjoint copies $G_1, ..., G_{D-d}$ of G with vertices $\{v_1^j, ..., v_n^j\} \in G_j$ for all $j \in [D-d]$ and $\forall i, j \in [D-d]$, the function $f_{ij} : G_i \to G_j$, $f_{ij}(v_k^i) = v_k^j$ for all $k \in [n]$ is an isomorphism, i.e. $\{v_k^i, v_{k'}^i\}$ is an edge if and only if $\{v_k^j, v_{k'}^j\}$ is an edge. We call the set $\Gamma_k := \{v_k^1, ..., v_k^{2(D-d)}\}$ the k^{th} corresponding vertex set of G. Now for any corresponding vertex set of G, we bipartition Γ_k into $A_k := \{v_k^1, ..., v_k^{D-d}\}$ and

$B_k = \Gamma_k \backslash A_k$. Then, corresponding vertex-wise, if $|N(V_k)| = r_k$ for some $r_k \in \{d, d+1, ..., D\}$, then we make a $(D - r_k)$-regular bipartite graph with parts A_k and B_k on Γ_k. It follows that Γ is D-regular and K_r-free $\forall r \geq 4$ because non-corresponding vertices on different copies are disconnected and the planted bipartite graphs on each corresponding vertex set is triangle-free and thus K_r-free for $r \geq 4$. Now by Theorem 2.2.2,

$$\bar{\alpha}(\Gamma) \geq c' \cdot 2(D-d)n\frac{\log D}{D \log\log D}, \tag{2.2.20}$$

which implies that $\exists F \subseteq I(\Gamma)$ s.t. $|F| \geq \bar{\alpha}(\Gamma)$ and so averaging over $2(D-d)$ copies of G, $\exists F_{G_j} \subseteq F \cap G_j$ with $|F_{G_j}| \geq \bar{\alpha}(G) \geq c' n \frac{\log d}{d \log\log d}$. Hence we have

$$\alpha(G) \geq |F_{G_j}| = \Omega(\frac{n \log d}{d \log\log d}), \tag{2.2.21}$$

where $d = d(n)$ such that $d(n) \to \infty$ as $n \to \infty$. $\square$

3 Upper Bound on $f(n, K_s, K_{s+1})$ for $s \geq 3$: Preliminaries

3.1 Finite Geometry Background

This section will mainly cover the finite geometry background behind Mubayi and Verstraëte's results [15] and will cover roughly chapter 1 and 2 of Barwick and Ebert's *Unitals in Projective Planes* [3]. We will briefly introduce the finite geometric properties behind Hermitian unitals, from projective spaces, finite fields, and α-sesquilinear forms to the construction of the Hermitian unitals. Another purpose of this section is to exemplify a connection between finite geometry and its combinatorial applications. More details regarding the connection between finite geometry and extremal combinatorics can be found in [2] and [6].

3.1.1 Projective Plane

Definition 3.1.1. A *projective plane* $\mathcal{P} = (P, L)$ is a pair of two sets P, called *points* and L, a collection of subsets of P, called *lines*, such that the following conditions hold:

(1). $\forall p_1, p_2 \in \mathbb{P}$, $\exists$ unique $l \in L$ s.t. $p_1, p_2 \in l$,

(2). $\forall l_1, l_2 \in L$, $l_1 \cap l_2 = p$ for some $p \in P$,

(3). $|P| \geq 4$ and for all $l \in L$, $|L| \geq 3$,

(4). $\exists p_1, p_2, p_3, p_4$ such that no triple is contained in a common line l.

The following two lemmas are examples of some interesting results from the axioms. It's not hard to prove them, so I will leave the proof as an exercise on people that try to understand projective plane in more detail.

Lemma 3.1.2. The last axiom of the above definition is equivalent to the following:

- $\exists l_1, l_2, l_3, l_4$ such that all triples $l_a, l_b, l_c \in \{l_1, l_2, l_3, l_4\}$ have empty intersection, i.e. $l_a \cap l_b \cap l_c = \varnothing$.

Lemma 3.1.3. A projective plane must have at least 7 points and 7 lines.

Definition 3.1.4. The *dual* of the projective plane $\mathcal{P} = (P, L)$, $\mathcal{P}' = (P', L')$, is the geometric structure produced by reversing the containment of points and lines, i.e., we put $L' = P$ and $P' = L$ and for any $l \in L = P'$, $p \in P = L'$, we say $l \in p$ in $\mathcal{P}'$ iff $p \in l$ in $\mathcal{P}$.

Remark 3.1.5. The dual of any projective plane is a projective plane by Definition 3.1.1 and Lemma 3.1.2.

Definition 3.1.6. For projective planes $\mathcal{P}_1 = (P_1, L_1)$ and $\mathcal{P}_2 = (P_2, L_2)$, a function $\phi : \mathcal{P}_1 \to \mathcal{P}_2$ is an *isomorphism* if $\phi|_{P_1} : P_1 \to P_2$, $\phi|_{L_1} : L_1 \to L_2$ are bijections and for all $p_1, l_1 \in \mathcal{P}_1$, $\phi(p_1) \in \phi(l_1)$ if and only if $p_1 \in l_1$. $\mathcal{P}_1$ and CP_2 are *isomorphic*, denoted $\mathcal{P}_1 \cong \mathcal{P}_2$, if such ϕ exists. We call the isomorphism $\phi : \mathcal{P}_1 \to \mathcal{P}_1$ an *automorphism* or a *collineation* of $\mathcal{P}_1$.

Upon understanding of the following theorem, we use the fact that every finite-dimensional vector space V over any field F has a finite basis (i.e. a linearly independent set that spans V), and all bases of F have the same cardinality. The theorem can be proved using induction, and I will omit this proof as it deviates our focus on this section.

Theorem 3.1.7 (Dimension Theorem)**.** For each n-dimensional vector space V over some field F and for any subspaces $A, B \subseteq V$,

$$\dim(A) + \dim(B) = \dim(A + B) + \dim(A \cap B). \tag{3.1.1}$$

Note that by the Gram-Schmidt Process, any finite-dimensional inner product space has an orthonormal basis. Also for any $n \in \mathbb{N}$, for the n-dimensional vector space V over some field F, we can always assign an inner product to V: Indeed, we first choose a basis $\{e_1, e_2, ..., e_n\}$ of V and define an inner product $\langle \cdot, \cdot \rangle$ as follows: $\forall x = \sum_{i=1}^n a_i e_i$, $y = \sum_{i=1}^n b_i e_i$, we let $\langle x, y \rangle := \sum_{i=1}^n a_i b_i$. It follows that $(V, \langle \cdot, \cdot \rangle)$ is an inner product space with an orthonormal basis $\{e_1, e_2, ..., e_n\}$. So for any n-dimensional vector space V over F, we may recognize $V = F^n := \{\sum_{i=1}^n a_i e_i \mid a_1, ..., a_n \in F,\ \{e_1, ..., e_n\}$ is an orthonormal basis of $(V, \langle \cdot, \cdot \rangle)\}$, and may also write each element $x \in V$ in coordinates, i.e. for $x = \sum_{i=1}^n a_i e_i$, $x = (a_1, a_2, ..., a_n)^T$.

3.1.2 The Projective Plane $PG(2, F)$

With the dimension theorem, we now construct the classical projective plane $PG(2, F)$. On the 3-dimensional vector space F^3, we let $\mathcal{P} = PG(2, F) = (P, L)$ where $P := \{p \subseteq F^3 :$ p is a 1-dimensional subspace and $L := \{l \subseteq F^3 : L$ is a 2-dimensional subspace$\}$. For any $p \in P$, $l \in L$, we say $p \in l$ if and only if $p \subseteq l$ in F^3. Then we have the following proposition:

Proposition 3.1.8. $\mathcal{P}$ defined above is a projective plane.

Proof. By the dimension theorem, since 2 distinct 1-dimensional subspaces are linearly independent, they span a unique 2-dimensional subspace $l \subseteq F^3$; since for any 2 distinct 2-dimensional subspaces A, B, $\dim(A \cap B) \geq 1$, $\dim(A \cap B) = 1$, which implies that $A \cap B = p$ for some 1-dimensional subspace p. Lastly, for 1-dimensional subspaces e_1, e_2, e_3 and $e_1 + e_2 + e_3$, note that any triple will span F^3, which implies that they do not lie in any 2-dim subspace l, which satisfies the last axiom. $\square$

Definition 3.1.9 (Homogeneous Coordinates)**.** Note that any 1-dimensional subspace p of $PG(2, F)$ is generated by some vector $\vec{v} = [x_1, x_2, x_3]^T$: it consists of all F-multiples of $[x_1, x_2, x_3]^T$, i.e. $p = \{\lambda[x_1, x_2, x_3]^T : \lambda \in F\}$. Therefore, we may represent p with any nonzero multiple of $[x_1, x_2, x_3]^T$, and, hence, we call $[x_1, x_2, x_3]^T$ the *homogeneous coordinates* of p.

Hence, the reason underlying the notation of $PG(2, F)$ to be a projective plane constructed from $V = F^3$ is addressed through the following definition:

Definition 3.1.10 (Projective dimension)**.** Through the definition of homogeneous coordinates, we naturally recognize 1-dimensional subspaces to be "points" and 2-dimensional subspaces to be "lines" on the projective space, also since 2 distinct 2-dimensional subspaces span F, we now define the *projective dimension* $\dim_{proj}(x) = \dim(x) - 1$ for any $x \in F^3 \backslash \{0\}$.

Remark 3.1.11. Under the above definition, F^3 under $PG(2, F)$ has projective dimension 2.

3.1.3 Projective Plane over Finite Fields

We first recall the following facts from finite fields, which can be found in any book or course that includes field theoretic topics:

- Any finite field F has characteristic p for some prime number p,
- $F \cong \mathbb{F}_q$ for some prime power $q = p^n$ $(n \in \mathbb{N})$, where $\mathbb{F}_q$ is the Galois field of order p^n,
- The multiplicative group $\mathbb{F}_q^\times = (\mathbb{F}_q \backslash \{0\}, \times)$ is cyclic, and any generator of this group is called a *primitive element* of the field.
- For $q = p^n$, The mapping $\phi : \mathbb{F}_q \to \mathbb{F}_q$, $x \in \mathbb{F}_q \mapsto x^q$ is a field isomorphism, and we call this mapping the *Frobenius automorphism.*

For $F = \mathbb{F}_q$ where q is some prime power, we use $PG(2, q)$ to represent the projective plane $PG(2, \mathbb{F}_q)$. Under the notion of homogeneous coordinates, when $F = \mathbb{F}_q$ where q is a prime power, we have the following properties:

- $[x_1, x_3, x_3]^T$ equals one of $[0, 0, 1]$, $[0, 1, z]$ or $[1, y, z]$ for some $y, z \in \mathbb{F}_q$. Therefore, $PG(2, q)$ contains $q^2 + q + 1$ points.
- For any $[x_1, x_2, x_3], [y_1, y_2, y_3] \in \mathbb{F}_q^3$, there exists a line $l = \text{span}\{[x_1, x_2, x_3]^T, [y_1, y_2, y_3]^T\} = \{\lambda_1[x_1, x_2, x_3]^T + \lambda_2[y_1, y_2, y_3]^T : \lambda_1, \lambda_2 \in \mathbb{F}_q\}$. So each line contains $q + 1$ points.

By a standard counting argument over each fixed point, we observe the third property:

- Each point is contained in $\frac{q^2+q+1-1}{q} = q + 1$ lines.

Now by double counting on point-line incidence pairs (p, l) where $p \in l$, we deduce the fourth property:

- $|L| = \frac{(q^2+q+1)(q+1)}{q+1} = q^2 + q + 1$ lines, where q^2+q+1 on the numerator represents the number of points, $q+1$ on the numerator represents the number of lines containing each $p \in P$, and the $q+1$ on the denominator represents the number of points contained in l for each $l \in L$.

Example 3.1.12. The Fano plane $PG(2,2)$ is the smallest projective plane, i.e. the projective plane with the smallest size of P and L.

3.1.4 Projective Geometry

Definition 3.1.13 (Projective Geometry)**.** Following from [7], We generalize projective planes to projective geometries $\mathcal{P} = (P, L)$ where P is a set called points and L is a set of subsets of P, called lines, under the following three axioms:

- $\forall p_1, p_2 \in P$, there exists a unique $l \in L$ such that $p_1, p_2 \in l$,
- $\forall l \in L$, $|l| = |\{p \in P : p \in L\}| \geq 3$,
- $\forall l_1, l_2 \in L$ such that $l_1 \cap l_2 = \{p\}$ for some $p \in P$, if $q, r \neq p$ are 2 distinct points on l_1 and $s, t \neq p$ are 2 distinct points on l_2, then the lines l_{qr} containing q and r and l_{st} containing s and t intersect at some point $p' \in P$.

Remark 3.1.14. Following from the above three axioms, a projective geometry is a projective plane if and only if for any $l_1, l_2 \in L$, $\varnothing \neq l_1 \cap l_2 \in P$.

Definition 3.1.15. For some field F and an n-dimensional vector space V over F where $n \geq 3$, any hyperplane H is represented by some linear equation

$$a_1X_1 + a_2X_2 + ... + a_nX_n = 0 \tag{3.1.2}$$

for some $a_1, a_2, ..., a_n$ not all zero. Viewing H under the projective geometry $PG(n-1, F)$, note the H consists of points whose homogeneous coordinates are solutions to the above equation. We then call $[a_1, a_2, ..., a_n]^T$ the *homogeneous dual coordinates* of H in $PG(n-1, F)$. Note that the "homogeneity" follows similarly from the definition of homogeneous coordinates of $PG(n-1, F)$ as $[a_1, ..., a_n]$ and $\lambda[a_1, ..., a_n]$ ($\lambda \in F^\times$) will represent the same hyperplane.

Definition 3.1.16 (Semilinear transformation)**.** Let F be a field and V be a vector space over F, and α be an automorphism of F. A function $T : V \to V$ is called a *semilinear transformation* with *companion automorphism* α if the following hold:

- $\forall v, w \in V$, $T(v + w) = T(v) + T(w)$;
- $\forall v \in V, \lambda \in F$, $T(\lambda v) = \lambda^\alpha T(v)$.

If T is a bijection, then we call T a *nonsingular semilinear transformation*; If α is the identity automorphism, then we call T a *linear transformation.*

We state the following theorem as a blackbox, and one can gain a quick insight through [17].

Theorem 3.1.17 (Fundamental Theorem of Projective Geometry)**.** Every automorphism (collineation) of $PG(n-1, F)$ ($n \geq 3$) is induced by a nonsingular semilinear transformation of the underlying vector space.

Remark 3.1.18. By definition of semilinear transformations, note that the set of all automorphisms of $PG(n-1, F)$ ($n \geq 3$) is a group under composition of functions. We call this group $P\Gamma L(n, F)$.

Definition 3.1.19 (Homography Subgroup)**.** Let V be an n-dimensional vector space over some field F where $n \geq 3$. The subgroup of $P\Gamma L(n, F)$ consisting of all automorphisms induced by nonsingular linear transformations of V is called the *homography subgroup* or *subgroup of projectivities*, denoted $PGL(n, F)$.

3.1.5 Linear Algebra

This subsection is mainly section 1.5 of Barwick and Ebert's book [3], with some additional references on Simeon Ball's *Finite Geometry and Combinatorial Applications* [2].

Definition 3.1.20 (α-sesquilinear form)**.** Let F be a field and α be an automorphism of F. Let V be a vector space over F. An *α-sesquilinear form* $s : V \times V \to F$ is defined to be such that:

- $\forall v_1, v_2, w_1, w_2 \in V$, $s(v_1 + v_2, w_1) = s(v_1, w) + s(v_2, w)$ and $s(v_1, w_1 + w_2) = s(v_1, w_1) + s(v_2, w_2)$,
- $\forall \lambda \in F$, $v, w \in V$, $s(\lambda v, w) = \lambda s(v, w)$,
- $\forall \lambda \in F$, $v, w \in V$, $s(v, \lambda w) = \lambda^\alpha s(v, w)$.

Note that by the second axiom, if s is an α-sesquilinear form and $v, w \in V$, then $s(v, 0) = s(0, w) = 0$. We call α the *companion automorphism* of s and we say s is *nondegenerate* if the only $v \in V$ satisfying $s(v, w) = 0$ for all $w \in V$ is $v = 0$.

Definition 3.1.21 (Orthogonal complement of subspaces)**.** For all vector space V over any field F with an α-sesquilinear form s, let W be any subspace of V. Then we define $W^\perp = \{v \in V : s(v, w) = 0 \; \forall w \in W\}$, called the *orthogonal complement* of W.

Definition 3.1.22 (Dual space)**.** Let V be a finite-dimensional vector space over some field F. We let V^* be the vector space consisting all linear maps $f : V \to F$ and call this vector space the *dual space* of V

Remark 3.1.23. V^* is a well-defined vector space by the definition and properties of linear maps.

Proposition 3.1.24. Let $\{e_1, e_2, ..., e_n\}$ be a basis of V. Let $S := \{e^1, e^2, ..., e^n\}$ where e^i is defined to be such that for any $i, j \in [n]$, $e^i(e_j) = \delta^i_j$, which is the Kronecker delta, i.e. $\delta^i_j = 0$ if $j \neq i$ and 1 if $j = i$. Then S is a basis of V^*.

Proof. Since the uniqueness of linear maps are dependent on where they send the basis $e_1, ..., e_n$, we have the following bijection:

$$f \in V^* \longleftrightarrow (f(e_1), ..., f(e_n)) \in F^n \longleftrightarrow f(e_1)e^1 + f(e_2)e^2 + ... + f(e_n)e^n \in V^*. \tag{3.1.3}$$

Claim follows. □

Definition 3.1.25. From the proposition above, we call B the *standard dual basis* of V^*.

Definition 3.1.26 (Annihilator)**.** For $W \subseteq V$ a subspace of a finite-dimensional vector space, we let $W^\circ = \{f \in V^* : f(w) = 0\ \forall w \in W\}$, called the *annihilator* of W. Similarly, if A is a subspace of V^*, we let $A^\circ := \{v \in V : f(w) = 0\ \forall f \in A\}$, called the *annihilator* of A.

Remark 3.1.27. The following construction is based on Folland's construction of the dual of dual spaces in section 5 of [8]. Let V be a finite-dimensional vector space over some field F. We now let $\hat{V} := \{\hat{v} \in V^{**} : v \in V\}$ where $\hat{v}$ is defined to be such that for any $f \in V^*$, $\hat{v}(f) = f(v)$. Note that $\hat{V} \subseteq V^{**}$ is a subspace since for all $a, b \in F$, $f, g \in V^*$, $\hat{v}(af + bg) = (af + bg)(v) = af(v) + bg(v) = a\hat{v}(f) + b\hat{v}(g)$. Also note that for a basis $\{e_1, ..., e_n\}$ of V, $\{\hat{e}_1, ..., \hat{e}_n\}$ is a basis of $\hat{V}$ by the setup of $\hat{V}$. So $\phi : V \to \hat{V}$, $\phi(v) = \hat{v}\ \forall v \in V$ is a linear bijection (also called an isomorphism) of vector spaces. Hence, $\dim(V^{**}) = \dim(V^*) = \dim(V) = \dim(\hat{V})$ so $\hat{V} = V^{**}$. Moreover, $A^\circ = \phi^{-1}(\{\hat{v} \in \hat{V} : \hat{v}(f) = 0\ \forall f \in A\})$.

Theorem 3.1.28. Let V be a finite-dimensional vector space over some field F and W be a subspace, and let A be a subspace of V^*. Then,

$$\dim(V) = \dim(W) + \dim(W^\circ) = \dim(A) + \dim(A^\circ). \tag{3.1.4}$$

Proof. By the above remark, it's enough to show that $\dim(V) = \dim(W) + \dim(W^\circ)$. Now we let $\psi : V^* \to W^*$, $\psi(f) = f|_W$ be the restriction map. Then note that ψ is a homomorphism on the abelian group $(V^*, +)$, with the kernel $\{f \in V^* : f(w) = 0\ \forall w \in W\} = W^\circ$. By a standard linear algebraic argument, we have $\dim(V) = \dim(W) + \dim(W^\circ)$. □

Theorem 3.1.29. Let V be a finite-dimensional vector space equipped with a nondegenerate α-sesquilinear form s. Let W be a subspace of V. Then

$$\dim(V) = \dim(W) + \dim(W^\perp). \tag{3.1.5}$$

Proof. For all $v \in V, w \in W$, let $s_w(v) := s(v, w)$ and $A_w := \{s_w : w \in W\}$, so $A_W \subseteq V^*$ is a subspace. Then we have

$$\begin{aligned} W^\perp &= \{v \in V : s(v, w) = 0\ \forall w \in W\} \\ &= \{v \in V : s_w(v) = 0\ \forall w \in W\} = A_W^\circ. \end{aligned} \tag{3.1.6}$$

So $\dim(V) = \dim(A_W^\circ) + \dim(A_W) = \dim(W^\perp) + \dim(A_W)$. Now it remains to show $\dim(A_W) = \dim(W)$. Note that the mapping $w \in W \mapsto s_w \in A_W$ is surjective (by definition) and injective since if there exists $w, w' \in W$, $w \neq w'$ such that $s_w = s_{w'}$ then $s_w - s_{w'} = 0$, which contradicts s being nondegenerate. Also for a basis $\{w_1, ..., w_n\}$ of W, $\{s_{w_1}, ..., s_{w_n}\}$ spans A_W and is linearly independent since if there exists $\lambda_1^\alpha, ..., \lambda_n^\alpha$ not all zero such that $\sum_{i=1}^n \lambda_i^\alpha s_{w_i} = 0$, then $\sum_{i=1}^m \lambda_i w_i = 0$, which is a contradiction. Thus, $\dim(W) = \dim(A_W)$. □

Remark 3.1.30. Note that for a finite dimensional vector space V over some field F, the mapping of subspaces $w \subseteq V \to W^\perp$ reverses containment, i.e. $\forall U \subseteq W$, $U^\perp \supseteq W^\perp$ by definition. Also, by theorem 3.1.29, it is injective. Therefore, when F is a finite field, the mapping is bijective.

Definition 3.1.31. For a projective geometry G, a *correlation* of G is a bijection of subspaces of G that reverses containment. In particular, a correlation interchanges points and hyperplanes.

We include the following theorem from [5] as a blackbox:

Theorem 3.1.32 (Birkoff-von Neumann)**.** Let $n \geq 3$, F be a field, $PG(n-1, F)$ be the classical projective geometry over F, and ρ be a correlation of $PG(n-1, F)$. Then there exists a nondegenerate α-sesquilinear form s which induced ρ, i.e. $W^\rho = W^\perp$ for any subspace $W \subseteq V$.

Definition 3.1.33. For any α-sesquilinear form s over some vector space V, we say s is *reflexive* if for any $v, w \in V$ satisfying $s(v, w) = 0$, $s(w, v) = 0$.

Theorem 3.1.34. Let $PG(n-1, F)$ be the classical projective geometry on the n-dimensional vector space V over some field F and ρ be a correlation with the associated sesquilinear form s. Then s is reflexive if and only if ρ has order 2, i.e. s is reflexive if and only if $W^{\perp\perp} = W$ for any subspace $W \subseteq V$.

Proof. For the forward direction, for any subspace $W \subseteq V$, note that $W^{\perp\perp} = \{w \in V : s(w, v) = 0 \ \forall v \in W^\perp\} = \{w \in V : s(v, w) = 0 \ \forall v \in W^\perp\} \subseteq W$. Also note that by theorem 3.1.29, $\dim(W) = \dim(W^{\perp\perp})$. So $W = W^{\perp\perp}$.

For the backward direction, if $W^{\perp\perp} = W$, then for any $v, w \in V$, $s(v, w) = 0 \Leftrightarrow v \in \langle w \rangle^\perp \Rightarrow \langle w \rangle^{\perp\perp} \subseteq \langle v \rangle^\perp \Rightarrow s(w, v) = 0$. □

Definition 3.1.35. When a correlation ρ has order 2, i.e. ρ^2 is identity, we call ρ a *polarity.*

The following theorem by Birkoff and von Neumann [5] is a classification of all nondegenerate reflexive sesquilinear forms s that are associated with some polarity ρ. Again, we include this theorem as a blackbox:

Theorem 3.1.36. Let ρ be a polarity of $PG(n-1, F)$ where $n \geq 3$ and F is some field; let s be an associated nondegenerate reflexive sesquilinear form with companion automorphism α. Then (s, α) is precisely one of the following:

- $\alpha = \mathrm{id}^F$ is the identity automorphism and $s(v, w) = s(w, v)$ for all $v, w \in V$. If the characteristic of F is 2, then $s(v, v) \neq 0$ for all v.
- $\alpha = \mathrm{id}_F$ and $s(v, v) = 0$ for all $v \in V$.
- α has order 2 and $s(v, w) = 0$ for all $v \in V$.

Definition 3.1.37. The polarity s satisfying the first, second, and third property of the above theorem are called *orthogonal, symplectic, and unitary* polarities respectively, and the associated α-sesquilinear forms are called *symmetric bilinear, skew-symmetric bilinear, and Hermitian* forms respectively.

We now define the Gram matrix G, which is one of the most essential concepts in section 3.1. We first fix a basis $B = \{e_1, e_2, ..., e_n\}$ of an n-dimensional vector space V over F, and construct an $n \times n$-dimensional matrix G where the (i, j)th entry $G_{ij} = s(e_i, e_j)$. For any $v, w \in V$, we write v, w in coordinates of F^n with respect to the basis B, i.e. for $v = \sum i = 1^n a_i e_i$ and $w = \sum_{i=1}^n b_i e_i$, $v = (a_1, ..., a_n)^T$ and $w = (b_1, ..., b_n)^T$. Then we have

$$s(v, w) = v^T G w^\alpha \tag{3.1.7}$$

where $w^\alpha = (b_1^\alpha, b_2^\alpha, ..., b_n^\alpha)^T$. Indeed, this is true since

$$\begin{aligned} s(v,w) &= s(\sum_{i=1}^n a_i e_i, \sum_{i=1}^n b_i e_i) \\ &= \sum_{i=1}^n a_i s(e_i, \sum_{i=1}^n b_j e_j) \\ &= (a_1, ..., a_n)^T \begin{bmatrix} s(e_1,e_1) & ... & s(e_1,e_n) \\ ... & ... & ... \\ s(e_n,e_1) & ... & s(e_n,e_n) \end{bmatrix} \begin{bmatrix} b_1^\alpha \\ . \\ . \\ . \\ b_n^\alpha \end{bmatrix} \\ &= v^T G w^\alpha. \end{aligned} \tag{3.1.8}$$

Also, bilinearity follows from linear algebra and properties of field automorphisms.

Now for unitary polarities, i.e. when s if Hermitian, we have $s(v, w) = s(w, v)^\alpha$ for all $v, w \in V$ and α is an automorphism of F with order 2. So we have

$$\begin{aligned} \sum_{i=1}^n b_i s(e_i, \sum_{j=1}^n a_j e_j) &= (\sum_{i=1}^n a_i s(e_i, \sum_{j=1}^n b_j e_j))^\alpha \\ \sum_{i=1}^n \sum_{j=1}^n b_i a_j^\alpha s(e_i, e_j) &= (\sum_{i=1}^n \sum_{j=1}^n a_i b_j^\alpha s(e_i, e_j))^\alpha \\ &= \sum_{i=1}^n \sum_{j=1}^n b_i a_j^\alpha s(e_j, e_i)^\alpha, \end{aligned} \tag{3.1.9}$$

which implies that $G_{ij} = G_{ji}^\alpha$ when the polarity if unitary.

Definition 3.1.38. Analogous to the definition of Hermitian matrices of $M_n(\mathbb{C})$, we define a matrix G on $M_n(F)$ with an order-2 automorphism α to be *Hermitian* if $G_{ij} = G_{ji}^\alpha$ for all $i, j \in [n]$. We also denote $G^\alpha := \left[G_{ij}^\alpha\right]_{i,j\in[n]}$.

So by the above definitions, the associated Gram matrices of unitary polarities are Hermitian, They are also nonsingular by the fact that s is nondegenerate.

By properties of finite field, a finite field F has an involutional (or involutary) automorphism, i.e. an automorphism of order 2, if and only if $F \cong \mathbb{F}_{q^2}$ for some prime power q. When $F \cong \mathbb{F}_{q^2}$, the field automorphism $x \mapsto x^q$ will be the associated companion automorphism for any Hermitian form s by properties of cyclic groups. In this case, for any n-dimensional vector space V over $\mathbb{F}_{q^2}$ and a Hermitian form s with companion automorphismα, and for all $v, w \in V$, $s(v, w) = v^T G w^q$. We also note that for the basis $\{e_1, ..., e_n\}$, $s(e_i, e_i) \in \mathbb{F}_q \subseteq F_{q^2}$ for all $i \in [n]$, which implies $s(v, v) \in \mathbb{F}_q$ for all $v \in V$.

Definition 3.1.39. For $PG(n-1, F)$ ($n \geq 3$) with P to be the set of points, H to be the set of hyperplanes, and ρ to be a polarity, for any $p \in P, h \in H$, we call the hyperplane P^ρ the *polar hyperplane* of p and the point h^ρ the *pole* of h.

Now, using the associated Gram matrix G of the nondegenerate sesquilinear form$s$ that coupled with the polarity ρ, write $p = \langle v \rangle \subseteq F^n$, $p^\rho = \{\langle w \rangle \in P : s(v, w) = 0\} = \{\langle w \rangle \in P : v^T G w = 0\} = \{\langle w \rangle \in P : w^T G v^\alpha = 0\}$ by reflexivity of s. So p^ρ has homogeneous dual coordinates G_{v^α}. Similarly, for all $h \in H$ such that h has homogeneous dual coordinates $[a_1, ..., a_n]^T =: \langle y \rangle$, if $\langle v \rangle = h^\rho$, then $s(v, w) = 0$ for all $w \in h$. Let $w = (x_1, ..., x_n)^T$, then $\sum_{i=1}^n a_i x_i = 0 \to y^T w = 0 \to y^T G^{-1} G w^\alpha = 0 \to \langle v \rangle$ has homogeneous coordinates $(y^\alpha)^T G^{-1}$.

Definition 3.1.40. Let ρ be a polarity of some projective geometry Π. Then for any point $p \in \Pi$, p is called *absolute* if $p \in p^\rho$ and *nonabsolute* if not; for any $H \subseteq \Pi$ a hyperplane, H is called *absolute* if $H^\rho \in H$ and *nonabsolute* if not.

3.1.6 Hermitian Curves and Unitals

Definition 3.1.41 (Hermitian variety and Hermitian curve)**.** Let ρ be a unitary polarity of the classical projective geometry $PG(n-1, F)$ where $n \geq 3$ and F is some field. Then we call the set of absolute points of ρ a *nondegenerate Hermitian variety.* When $n = 3$, we call the set of absolute points of $PG(2, F)$ a *nondegenerate Hermitian curve.*

Remark 3.1.42. Note that Hermitian varieties can be empty. Take the example of $F = \mathbb{C}$ and α to be the complex conjugation isomorphism. Take the standard basis of $\mathbb{C}^n$ associated with the inner product $\langle \cdot, \cdot \rangle$ where for all $x, y \in \mathbb{C}^n$, $x = (x_1, ..., x_n)^T$, $y = (y_1, .., y_n)^T$, $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$. Let the Gram matrix $G = I_n$, which associates an α-sesquilinear form s where $s(x, y) = \langle x, y \rangle$ is the inner product defined above. Observe that for any $W \subseteq \mathbb{C}^n$ a subspace, $W^\perp$ consists of all element of $\mathbb{C}^n$ orthogonal to W (with respect to the inner product). So if x is an absolute point, then $\sum_{i=1}^n x_i \bar{x}_i = 0 \to \sum_{i=1}^n |x_i|^2 = 0 \to x = 0$, which implies that for all $n \geq 3$, in $PG(n-1, \mathbb{C})$, the Hermitian variety is empty.

However, we will see later that all Hermitian varieties over $PG(2, q^2)$ where q is a prime power is nonempty.

Now we capture the projective equivalence of nondegenerate Hermitian varieties $\mathcal{H}(n-1, q^2)$ in $PG(n-1, q^2)$ with P to be the set of points. Consider any two Hermitian varieties $\mathcal{H}_1, \mathcal{H}_2$ induced by distinct Hermitian forms s_1, s_2 with associated gram matrix G_1, G_2. Now for all $p \in P$ with homogeneous coordinates $[x_1, ..., x_n]^T$, let $\mathcal{H}_1 = \{p \in P : p = [x_1, ..., x_n]$ and $(x_1, ..., x_n) G (x_1^p, ..., x_n^p)^T = 0\}$, $\mathcal{H}_2 = \{p \in P : p = [y_1, ..., y_n]$ and $(y_1, ..., y_n) G_2 (y_1^p, ..., y_n^p)^T = 0\}$. Since G_1, G_2 are nontrivial Hermitian, $G_2 = U^\alpha G_1 U$ for some unitary matrix U (i.e. U is invertible and $U_{-1} = U^\alpha$). So $\mathcal{H}_2 = \{p \in P : p = [y_1, ..., y_n]$ and $(y_1, ..., y_n) U^\alpha G_1 U (y_1^\alpha, ..., y_n^\alpha)^T = 0\}$, which implies for all $p = [x_1, ..., x_n]^T \in \mathcal{H}_1$, $(x_1, ..., x_n)^T = U(y_1, ..., y_n)^T$ for some $p_2 \in \mathcal{H}_2$ with homogeneous coordinates $[y_1, ..., y_n]$. Note that under the vector space $\mathbb{F}_{q^2}^n$, U is a nonsingular linear transformation, which implies $U \subseteq PGL(n, \mathbb{F}_{q^2})$ from the earlier definition of homography subgroup, or subgroup of projectivities. In conclusion, we have the following proposition (one can see more details in [9] regarding homography subgroups and its relation with all semilinear transformations):

Proposition 3.1.43. Any nondegenerate Hermitian variety $\mathcal{H}$ of $PG(n-1, q^2)$ $(n \geq 3)$ can be mapped to any other nondegenerate Hermitian variety $\mathcal{H}'$ of $PG(n-1, q^2)$ by some homography. Hence, for any nondegenerate Hermitian variety $\mathcal{H}$ of $PG(n-1, q^2)$, we say $\mathcal{H}$ is *uniquely determined up to projective equivalence.*

We have the following 2 theorems, theorem 3.1.44 and 3.1.45, as balckboxes, before inspecting on Hermitian curves on $PG(n-1, q^2)$ where $n \geq 3$:

Theorem 3.1.44. If $\mathcal{P} = (P, L)$ is any finite projective plane, i.e. $|P|, |L| < \infty$, then there exists some inter ger $n \geq 2$ such that:

- $\forall p \in P$, $|\{l \in L : p \in l\}| = n + 1$,
- $\forall l \in L$, $|l| = n + 1$,
- $|P| = |L| = n^2 + n + 1$.

Theorem 3.1.45 (Hughes and Piper [11])**.** For any projective plane $\mathcal{P} = (P, L)$ of order $N = n^2$ for some $n \in \mathbb{N}$, let ρ be a polarity. Then $\mathcal{H}$, the set of absolute points in P, has size at most n^3+1. Moreover, if $|H| = n^3 + 1$, then for each $l \in L$, $|l \cap H| = 1$ or $n + 1$.

Theorem 3.1.46. A nondegenerate Hermitian curve $\mathcal{H}$ in $PG(2, q^2)$ (where q is a prime power) has precisely $q^3 + 1$ points.

Proof. By the proposition 3.1.43, we have that all Hermitian curves in $PG(2, q^2)$ are projectively equivalent. Therefore, it's enough to consider the case where the polarity ρ is induced by a Hermitian form s with the companion automorphism α being the automorphism $x \mapsto x^q$ and the associated Gram matrix to be the identity matrix. Then $\mathcal{H}$ has the equation

$$X_1^{q+1} + X_2^{q+1} + X_3^{q+1} = 0, \tag{3.1.10}$$

which has solutions $[0, 1, z]^T : z^{q+1} = -1$; $[1, y, 0]^T : y^{q+1} = -1$; $[1, y, z]^T : y^{q+1} \neq -1$ and $z^{q+1} = -1 - y^{q+1} \neq 0$. Now note that $(\mathbb{F}_{q^2}, \times) \cong C_{q^2-1}$, the cyclic group of order $q^2 - 1$ and the mapping $x \mapsto x^{q+1}$ restricted to $\mathbb{F}_{q^2}^{\times}$ is surjective from $\mathbb{F}_{q^2}^{\times}$ to the subfield $\mathbb{F}_q^{\times} \subseteq \mathbb{F}_{q^2}^{\times}$ since $(x^{q+1})^{q-1} = 1$ for all $x \in \mathbb{F}_{q^2}$ and for $v \in \mathbb{F}_q^{\times}$ which generates $\mathbb{F}_{q^2}^{\times}$, v^{q+1} has order $q - 1$, so v^{q+1} generates $\mathbb{F}_q^{\times}$. Also, this mapping is precisely $(q+1)$-to-1 since the equation $X^{q+1} = a$ ($a \in \mathbb{F}_q$) has no more than $q + 1$ solutions, which by pigeon hold principle, has precisely $q + 1$ solutions. So we have in total $2(q + 1) + (q + 1)(q^2 - q + 1) = (q^2 - q + 1)(q + 1) = q^3 + 1$ solutions. □

Now by theorem 3.1.45, we immediately have the following theorem:

Theorem 3.1.47. If $\mathcal{H}$ is a nondegenerate Hermitian curve on $PG(2, q^2)$, then for every line l of $PG(2, q^2)$< $|l \cap H| = 1$ or $q + 1$.

Remark 3.1.48. There are $q^2(q^2 - q + 1)$ secant lines of $\mathcal{H}$ by double counting:

$$\text{number of lines} = \frac{(q^3 + 1)\frac{q^3}{q}}{q + 1} = q^2(q^2 - q + 1). \tag{3.1.11}$$

We now complete the definition of unitals through the definition of designs:

Definition 3.1.49 (design)**.** Let $t, v, k, \lambda \in \mathbb{N}$ and $t < k < v$. A t-(v, k, λ) *design* is an ordered pair (V, B) where V is a set of *points*, $|V| = v$, and B is a set of subsets of V of size k, called *blocks*. Moreover, (V, B) satisfies that for any $V' \subseteq V$ where $|V'| = t$, $|V \cap B| = \lambda$, i.e. every t points are contained in exactly λ blocks.

Example 3.1.50. For a prime power q, any nondegenerate Hermitian curve $\mathcal{H} \subseteq PG(2, q^2)$ is a 2-$(q^3+1, q+1, 1)$ design.

Definition 3.1.51. Let $n \geq 3$ be an integer, q be a prime power. We call any 2-$(n^3+1, n+1, 1)$ design a *unital of order* n, and we will call nondegenerate Hermitian curves on $PG(2, q^2)$ *Hermitian unitals.*

One of the most important properties of Hermitian unitals will be addressed in the next theorem. We first define an O'Nan configuration on the projective plane $PG(2, \mathbb{F}_q')$ (where q' is a prime power) to be a pair of sets (P_0, L_0) where $P_0 \subseteq P$, $L_0 \subseteq L$, $P_0 = \{a, b, c, d, e, f\}$, $L_0 = \{l_1, l_2, l_3, l_4\}$ such that (P_0, L_0) satisfies the following structure:

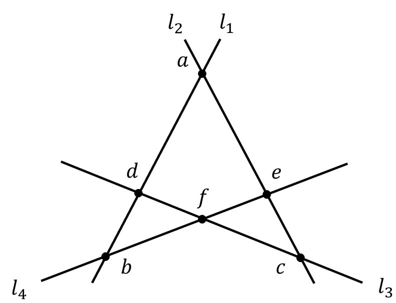

Figure 1: The O'Nan Configuration

The following theorem was first proved by O'Nan in 1971 (see [16]). Mattheus and Verstraëte [13] found a linear algebraic proof of the theorem, which will be reconstructed in this paper.

Theorem 3.1.52. For any Hermitian unital $\mathcal{H}$ on $PG(2, q^2)$ where q is a prime power, $\mathcal{H}$ does not contain the O'Nan configuration.

Proof. Fix a representation of homogeneous coordinates a_0 of a. Choose a representation of homogeneous coordinates b_0 of b such that $d_0 = a_0 + b_0$, for some d_0 a representation of d. Choose a representation of homogeneous coordinates c_0 of c such that $e_0 = a_0 + c_0$ for e_0 satisfying $\langle e_0 \rangle = e$. Then for some $\lambda_1, \lambda_2 \in \mathbb{F}_{q^2}^{\times}$, since b linearly independent with $\text{Span}\{a, c\}$,

$$\begin{aligned} \langle b_0 + \lambda_1 e_0 \rangle &= f = \langle \lambda_2 c_0 + d_0 \rangle \\ \langle b_0 + \lambda_1(a_0 + c_0) \rangle &= \langle \lambda_2 c_0 + a_0 + b_0 \rangle \\ \langle \lambda_1 a_0 + \lambda_1 c_0 \rangle &= \langle \lambda_2 c_0 + a_0 + b_0 \rangle \\ \langle \lambda_1 a_0 + \lambda_1 c_0 \rangle &= \langle \lambda_2 c_0 + a_0 \rangle \\ \langle a_0 + c_0 \rangle = \langle a + \lambda_2 c_0 \rangle & \end{aligned} \tag{3.1.12}$$

So $\lambda_2 = 1$, which implies $\langle b_0 + (a_0 + c_0) \rangle = \langle b_0 + \lambda_1(a_0 + c_0) \rangle$. So $\lambda_1 = 1$. Hence $f = \langle a_0 + b_0 + c_0 \rangle = a + b + c$. Now we choose a representation of homogeneous coordinates f_0 of f such that $f_0 = a_0 + b_0 + c_0$. By projective equivalence of Hermitian unitals, we consider the Hermitian unital $\mathcal{H}$ induced by the α-sesquilinear form s with the associated Gram matrix to be the identity matrix, and α is the automorphism $x \mapsto x^q$. Construct matrices

$$A := \begin{bmatrix} a_0^T \\ b_0^T \\ c_0^T \end{bmatrix} \text{ and } B := \begin{bmatrix} a_0^\alpha & b_0^\alpha & c_0^\alpha \end{bmatrix}. \tag{3.1.13}$$

Note that both A and B are nonsingular as $\{a_0, b_0, c_0\}$ are linearly independent, so is a_0^α, b_0^α, c_0^α. next note that

$$AB = \begin{bmatrix} a_0^T \\ b_0^T \\ c_0^T \end{bmatrix} \begin{bmatrix} a_0^\alpha & b_0^\alpha & c_0^\alpha \end{bmatrix} = \begin{bmatrix} s(a,a) & s(a,b) & s(a,c) \\ s(b,a) & s(b,b) & s(b,c) \\ s(c,a) & s(c,b) & s(c,c) \end{bmatrix} \tag{3.1.14}$$

By definition of Hermitian variety on $PG(2, q^2)$, $s(d, d) = s(e, e) = s(f, f) = 0$, which implies that $s(a, b) = -s(b, a)$, $s(a, c) = -s(c, a)$, and $s(a, b) + s(a, c) + s(b, a) + s(b, c) + s(c, a) + s(c, b) = 0$, which implies $s(b, c) = -s(c, b)$. So

$$AB = \begin{bmatrix} 0 & s(a,b) & s(a,c) \\ -s(a,b) & 0 & s(b,c) \\ -s(a,c) & -s(b,c) & 0 \end{bmatrix} \tag{3.1.15}$$

Note that $\det(A, b) = -s(a, b)s(b, c)s(a, c) + s(a, c)s(a, b)s(b, c) = 0$, which contradicts the nonsingularity of AB. Thus, the O'Nan configuration is forbidden in $\mathcal{H}$ for any $\mathcal{H}$ being a Hermitian unital in $PG(2, q^2)$ where q is a prime power. □

3.2 Probabilistic Methods

This section will introduce three common theorems of probabilistic methods from [1] that will be used in Mubayi and Verstraëte's results [15] on the upper bound.

Theorem 3.2.1 (Lovasz Local Lemma)**.** Let $A_1, A_2, .., A_n$ be events in some probability space, and $D = (V, E)$ be a dependency digraph of $A_1, A_2, ..., A_n$, i.e. $V = \{1, 2, ..., n\}$ and for any $i, j \in V$, the directed edge $(i, j) \in E$ if and only if $i \neq j$ and A_i depends on A_j. Suppose there exists $x_1, ..., x_n \in [0, 1)$ such that $\mathbb{P}(A_i) \leq x_i \cdot \prod_{j:(i,j)\in E}(1 - x_j)$ for all $i \in [n]$. Then

$$\mathbb{P}(\bigcap_{i=1}^{n} \bar{A}_i) \geq \prod_{i=1}^{n}(1 - x_i). \tag{3.2.1}$$

Proof. We claim that for all $s \in [n]$, $0 \leq |S| \leq n - 1$, for any $i \notin S$,

$$\mathbb{P}(A_0 \mid \bigcap_{j\in S} \bar{A}_j) \leq x_i. \tag{3.2.2}$$

Note that if the claim is true, then since for any events A, B, C, $\mathbb{P}(A \cap B|C) = \mathbb{P}(A|C)\mathbb{P}(B|A \cap C)$,

$$\begin{aligned} \mathbb{P}(\bigcap_{i=1}^{n} \bar{A}_i) &= \mathbb{P}(\bar{A}_i)\mathbb{P}(\bigcap_{i=1}^{n} \bar{A}_i \mid \bar{A}_1) \\ &= \mathbb{P}(\bar{A}_i)\prod_{i=1}^{n}\mathbb{P}(\bar{A}_i \mid \bigcap_{j=1}^{i-1} \bar{A}_j) \\ &= (1 - \mathbb{P}(\bar{A}_1))\prod_{i=1}^{n}(a - \mathbb{P}(\bar{A}_i \mid \bigcap_{j=1}^{i-1} \mathbb{P}(\bar{A}_j))) \qquad (3.2.3) \\ &\geq (1 - x_1)\prod_{i=2}^{n}(1 - x_i) \\ &= \prod_{i=1}^{n}(1 - x_i). \end{aligned}$$

Lemma then follows. Now we will prove the claim by induction. If $|S| = 0$, then $P(A_i) \leq x_i$ since for all $j \neq i$ we have $0 \leq x_j < 1$. Suppose the claim holds for all S' such that $|S'| \leq s-1$ for some s satisfying $s \in 1, ..., n-1$. By strong induction hypothesis, for any S, i defined in the claim where $|S| = s$, let $S_1 = \{j \in S : (i,j) \in E\}$, $S_2 = S \backslash S_1$. Now,

$$\begin{aligned}\mathbb{P}(A_i \mid \bigcap_{j\in S} \bar{A}_j) &= \frac{\mathbb{P}(A_i \cap \bigcap_{j\in S} \bar{A}_j)}{\mathbb{P}(\bigcap_{j\in S} \bar{A}_j)} \\ &= \frac{\mathbb{P}(A_i \cap \bigcap_{j\in S_1} \bar{A}_j \mid \bigcap_{k\in S_2} \bar{S}_k)\mathbb{P}(\bigcap_{k\in S_2} \bar{A}_k)}{\mathbb{P}(\bigcap_{j\in S_1} \bar{S}_j \mid \bigcap_{k\in S_2} \bar{A}_k)\mathbb{P}(\bigcap_{k\in S_2} \bar{A}_k)} \\ &= \frac{\mathbb{P}(A_i \cap \bigcap_{j\in S_1} \bar{A}_j \mid \bigcap_{k\in S_2} \bar{A}_k)}{\mathbb{P}(\bigcap_{j\in S_1} \bar{A}_j \mid \bigcap_{s\in S_2} \bar{A}_k)}.\end{aligned} \tag{3.2.4}$$

Denote $S_1 := \{j_1, ..., j_r\}$ for some r. Then note that

$$\begin{aligned}\mathbb{P}(A_i \cap \bigcap_{j\in S_1} \bar{A}_j \mid \bigcap_{k\in S_2} \bar{A}_k) &\leq \mathbb{P}(A_i \mid \bigcap_{k\in S_2} \bar{A}_k) \\ &\leq \mathbb{P}(A_i) \\ &\leq \prod_{j:(i,j)\in E} (1 - x_j) \\ &\leq x_i \prod_{i=1}^{r} (1 - x_{j_i}).\end{aligned} \tag{3.2.5}$$

Also note that

$$\begin{aligned}\mathbb{P}(\bigcap_{j\in S_1} \bar{A}_j \mid \bigcap_{k\in S_2} \bar{A}_k) &= \mathbb{P}(\bar{A}_{j_1} \mid \bigcap_{k\in S_2} \bar{A}_k) \cdot \prod_{i=1}^{r} \mathbb{P}(\bar{A}_{j_i} \mid \bigcap_{l=1}^{i-1} \bar{A}_{j_l} \cap \bigcap_{k\in S_2} \bar{A}_k) \\ &= (1 - \mathbb{P}(A_{j_1} \mid \bigcap_{k\in S_2} \bar{A}_k)) \cdot \prod_{i=1}^{r}(1 - \mathbb{P}(A_{j_i} \mid \bigcap_{l=1}^{i-1} \bar{A}_{j_l} \cap \bigcap_{k\in S_2} \bar{A}_k)) \\ &\geq \prod_{i=1}^{r}(1 - x_{j_i}).\end{aligned} \tag{3.2.6}$$

Altogether, claim follows. □

We have the following setup for Janson's Inequality. let Ω be a finite set, and R be a random subset of Ω such that for all $r \in \Omega$, $\mathbb{P}(r \in R) = p_r$. Let $A_1, ..., A_n$ be subsets of Ω for some $n \in \mathbb{N}$, and B_i be the event that $A_i \subseteq R$. Let $X_i := 1_{B_i}$ be the indicator random variable of B_i, and let $X := \sum_{i=1}^{n} B_i$. We then define $\sim$ to be the relation where for all $i, j \in [n]$, $i \sim j$ represents $i \neq j$ and $A_i \cap A_j \neq \varnothing$. So when $i \neq j$ and $i \nsim j$, the evenets B_i and B_j are independent. Now let $\Delta := \sum_{i\sim j} \mathbb{P}(B_i \cap B_j)$, set $M := \prod_{i=1}^{n} \mathbb{P}(\bar{B}_i)$, $\mu = E[X] = \sum_{i\in I} \mathbb{P}(B_i)$.

Theorem 3.2.2 (Janson's Inequality)**.** Let $\{B_i\}_{i=1}^{n}$, M, μ defined on the above setup, and let $\epsilon = \max_{i\in[n]} \mathbb{P}(B_i)$ and assume $\delta \leq \mu$. Then,

$$M \leq \mathbb{P}(\bigcap_{i=1}^{n} \bar{B}_i) \leq M e^{\frac{1}{1-\epsilon} \cdot \frac{\delta}{2}} \tag{3.2.7}$$

and

$$\mathbb{P}(\bigcap_{i=1}^{n} \bar{B}_i) \leq e^{-\mu+\frac{\delta}{2}}. \tag{3.2.8}$$

Proof. In this setting, for any $J \subseteq I = \{1, ..., n\}$, $i \in I$, we will use the inequalities $\mathbb{P}(\bar{B}_i \mid \bigcap_{j \neq i, j \in J} \bar{B}_j \subseteq \mathbb{P}(B_i)$ and $\mathbb{P}(B_i \mid B_k \cap \bigcap_{j \neq i,k, j \in J} \bar{B}_j) \leq \mathbb{P}(B_i \mid B_k)$, both of which follow from the setup of $\{B_i\}_{i=1}^{n}$. Now for the lower bound, note that $\mathbb{P}(\bar{B}_i \mid \bigcap_{i \neq i, j \in iJ} \bar{B}_j) \geq \mathbb{P}(\bar{B}_j)$ so we have

$$\begin{aligned}\mathbb{P}(\bigcap_{i \in I} \bar{B}_i) &= \mathbb{P}(\bar{B}_i) \cdot \prod_{i=2}^{n} \mathbb{P}(\bar{B}_i \mid \bigcap_{k=1}^{i-1} \bar{B}_k) \\ &\geq \prod_{i=1}^{n} \mathbb{P}(\bar{B}_i) = M.\end{aligned} \tag{3.2.9}$$

For the upper bound, for any $i \geq 2$, $k \in [i-1]$, we renumber $\bar{B}_1, ..., \bar{B}_{i-1}$ such that for some $d \geq 1$, $\bar{B}_1, ..., \bar{B}_d$ are dependent with $\bar{B}_i$ and $\bar{B}_{d+1}, ..., \bar{B}_{i-1}$ are not. Then we have

$$\begin{aligned}\mathbb{P}(B_i \mid \bigcap_{k=1}^{i-1} \bar{B}_k) &= \mathbb{P}(B_i \mid \bigcap_{k=1}^{d} \bar{B}_k \cap \bigcap_{k=d+1}^{i-1} \bar{B}_l) \\ &= \mathbb{P}(B_i)\mathbb{P}(\bigcap_{k=1}^{d} \bar{B}_k \mid B_i \cap \bigcap_{k=d+1}^{i-1} \bar{B}_l) \\ &= \mathbb{P}(B_i)(1 - \mathbb{P}(\bigcap_{k=1}^{d} B_k \mid B_i \cap \bigcap_{k=d+1}^{i-1} \bar{B}_l)) \\ &\geq \mathbb{P}(B_i)(1 - \sum_{k=1}^{d} \mathbb{P}(B_k \mid B_i)) \\ &= \mathbb{P}(B_i) - \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i).\end{aligned} \tag{3.2.10}$$

So since $\mathbb{P}(B_i) \leq \epsilon$, we have

$$\begin{aligned}\mathbb{P}(\bar{B}_i \mid \bigcap_{k=1}^{i-1} \bar{B}_k) &\leq \mathbb{P}(\bar{B}_i) + \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i) \\ &\leq \mathbb{P}(\bar{B}_i)(1 + \frac{1}{1-\epsilon} \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i)) \\ &\leq \mathbb{P}(\bar{B}_i) e^{\frac{1}{1-\epsilon} \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i)}.\end{aligned} \tag{3.2.11}$$

So

$$\begin{aligned}\mathbb{P}(\bigcap_{i=1}^{n} \bar{B}_i) &\leq \mathbb{P}(\bar{B}_1) \prod_{i=1}^{n} \mathbb{P}(\bar{B}_i \mid \bigcap_{k=1}^{i-1} \bar{B}_k) \\ &\leq \mu e^{\frac{1}{1-\epsilon} \cdot \frac{\delta}{2}}.\end{aligned} \tag{3.2.12}$$

To prove (2), note that

$$\mathbb{P}(\bar{B}_i \mid \bigcap_{k=1}^{i-1} \bar{B}_k) \leq 1 - \mathbb{P}(B_i) + \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i) \leq e^{-\mathbb{P}(B_i) + \sum_{k=1}^{d} \mathbb{P}(B_k \cap B_i)}. \tag{3.2.13}$$

(2) then follows. □

Remark 3.2.3. If $\Delta \geq 2\mu$ then the above theorem will have no meaning. So we have the following inequality, called the *extended Janson's inequality*: under the setup of Janson's Inequality, when $\Delta \geq 2\mu$,

$$\mathbb{P}(\bigcap_{i \in [n]} \bar{B}_i) \leq e^{-\frac{\mu^2}{2\delta}}. \tag{3.2.14}$$

We state the above theorem as an extension of our current topic - this will not be used in Mubayi and Verstraëte's proof in the next section. The details of this extended Janson's inequality can be found in [1].

The following theorem originates from the appendix of [1] and will use as a blackbox.

Theorem 3.2.4 (Chernoff bound [1] [15])**.** Let X be a binomial random variable with mean μ. Then for any $\epsilon \in [0, 1]$,

$$\mathbb{P}(X > (1 + \epsilon)\mu) \leq e^{-\frac{\epsilon^2 \mu}{4}} \text{ and} \tag{3.2.15}$$

$$\mathbb{P}(x < (1 - \epsilon)\mu) \leq e^{-\frac{\epsilon^2 \mu}{2}}. \tag{3.2.16}$$

4 Upper Bound on $f(n, K_s, K_{s+1})$ for $s \geq 3$: Mubayi and Verstraëte's Results

Theorem 4.0.1. For any $s \geq 3$, $f(n, K_s, K_{s+1}) = O(\sqrt{n} \log n)$.

4.1 Setups

Definition 4.1.1 (Blowup of a graph [20])**.** For a graph G and an integer r where $r \geq 2$, an r-*blowup* of G is a graph $G\chi$ where χ contains color classes $\chi := \{c_1, ..., c_r\}$ such that $V(G\chi) = \bigsqcup_{i=1}^{r} X_{c_i}$ for each $X_{c_i} \subseteq V(G)$ and for any pair of vertices u, v, $(u, v) \in E(G\chi)$ if and only if $(u, v) \in E(G)$ and $u \in X_{c_i}, v \in X_{c_j}$ for some $i \neq j$. In other words, we assign each vertex in G a color from χ, keep the edges whose endpoints belong to different color classes, and delete all other edges.

We state the following proposition as a black box as it arises from Janson's inequality. Detailed proof can be seen in the appendix of [15]

Proposition 4.1.2. Let $G_{n,\rho}$ be a random graph on n vertices, i.e. $|V(G)| = n$ and for any $v_1, v_2, \in V(G)$, $\mathbb{P}((v_1, v_2) \in E) = \rho$. For any $s \geq 3$, let $n \geq 2^{40s}$ and $\rho = (\frac{8s}{n})^{\frac{2}{s}}$, and let χ be the associated color class of an s-blowup ($s \geq 3$) of $G_{n,\rho}$ with each color class having at least $\frac{n}{2s}$ vertices. Let $G_{n,\rho}(\chi)$ be the graph after the blowup. Then,

$$\mathbb{P}(K_s \nsubseteq G_{n,\rho}(\chi) \leq e^{-2^{2s-4}n}. \tag{4.1.1}$$

Definition 4.1.3 (s-fan)**.** For $s \geq 3$, an s-fan is a set of s pairwise intersecting lines $s-1$ of which are concurrent at a point say p. When $s \geq 4$, p is unique and is called the *point of concurrency* of the s-fan, and the line that doesn't cross p is called the *base line* of the s-fan.

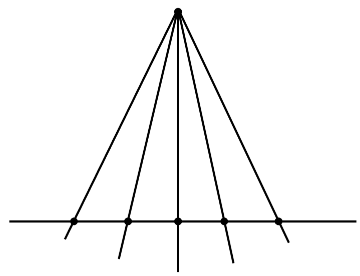

Figure 2: Example: A 6-fan

For any $s \geq 3$, choose some prime power q such that $q^2 >> s$, and consider a Hermitian unital $\mathcal{H}_q = (P, L)$ in $PG(2, q^2)$. Then we have the following lemma:

Lemma 4.1.4. If s lines in $\mathcal{H}_q$ pairwise intersect, then they are either concurrent with some point of $\mathcal{H}_q$ or they form an s-fan.

Proof. When $s = 3$, this is trivial. When $s = 4$, for the pairwise intersecting lines $l_1, l_2, l_3, l_4 \in \mathcal{H}_q$, if there does not exist $p \in l_1 \cup l_2 \cap l_3 \cap l_4$ such that p is contained in at least 3 of $l_1, ..., l_4$, then they form an O'Nan configuration, which is a contradiction. For $s \geq 5$, if we have $l_1, ..., l_5$ pairwise intersect and $\nexists p \in l_1 \cup ... \cup l_5$ such that p is contained in at least 4 of $l_1, ..., l_5$, then by the forbidden O'Nan configuration, $\exists p \in l_1 \cup ... \cup l_5$ such that p is contained in 3 lines, say l_1, l_2 and l_3. Let l_4 be such that $l_4 \cap l_1 = p_1$, $l_4 \cap l_2 = p_2$, $l_4 \cap l_3 = p_3$, all different from p. Now for l_5, note that it must not contain p_1, p_2 or p_3 since if not, WLOG, say $p_1 \in l_5$, then l_2, l_3, l_4 and l_5 form an O'Nan configuration (this can be realized by deleting l_1). So l_5 must cross l_1, l_2, l_4 in 4 points different from p_1, p_2, p_3. But l_5 coupled with any three other lines will give an O'Nan configuration, which is a contradiction. $\square$

4.2 The K_{s+1}-free Process

Definition 4.2.1 (Intersection graph)**.** We define the *intersection graph* G as follows:

$$V(G) = L, \quad E(G) = \{(l_1, l_2) : l_1, l_2 \in L, l_1 \cap l_2 \neq \varnothing\}. \tag{4.2.1}$$

Let $n = |V(G)|$ (we will keep this notation throughout the following sections). Note that G is an edge-disjoint union of K_{q^2}'s. To make G K_{s+1}-free while containing sufficiently many K_s's, we will first sample points in $\mathcal{H}_q$ and edges in G such that for any K_{s+1} in G_χ, the graph after the sampling through random blowups, K_{s+1} is an $(s+1)$-fan in $\mathcal{H}_q$. We next make a random graph $G_\rho = G_{n,\rho}$ where $\rho = (\frac{8s}{n})^{\frac{2}{s}}$. Let $H = G_\chi \cap G_\rho$, i.e. $V(H) = V(G_\chi) = V(G_\rho)$ and $E(H) = E(G_\chi) \cap E(G_\rho)$. We will then argue that with positive probability, H is K_{s+1}-free and for any large enough set X, $H[X]$ contains K_s.

4.2.1 Randomly sampling points in $\mathcal{H}_q$

Lemma 4.2.2. For any $s \geq 3$, $a \geq 128$ and any (large) prime power $q \geq a \log q$, there exists a partial linear space $\mathcal{H} = \mathcal{H}_{a,q,s} \subseteq \mathcal{H}_q$, i.e. $\mathcal{H} = (P_{\mathcal{H}}, L_{\mathcal{H}})$ where $P_{\mathcal{H}} \subseteq P$, $L_{\mathcal{H}} = \{l \cap P_{\mathcal{H}} : l \in L\}$, such that the following hold:

(1). $|L_H| = q^2(q^2 - q + 1)$ and any $s + 1$ pairwise intersecting lines are either concurrent on some point p or is an s-fan.

(2). $\frac{aq^2 \log q}{2} \leq |P_{\mathcal{H}}| \leq 2aq^2 \log q$.

(3). $\forall l \in L_{\mathcal{H}}$, $\frac{a \log q}{2} \leq |l| \leq 2a \log q$.

(4). The number of $(s + 1)$-fans in $\mathcal{H}$ containing a pair of lines ini $L_{\mathcal{H}}$ is at most $k = (2a \log q)^s$.

Proof. We uniformly randomly select points in $\mathcal{H}_q$ with probability $p_0 = \frac{a \log q}{q+1}$, so $p_0(q^3 + 1) \leq aq^2 \log q$. Note that (1) follows from (2) and (3). For (2), when q is large, by the Chernoff bound in the form of theorem 3.2.4, we choose $\epsilon = \frac{1}{2} - \delta$ for some small δ such that $\epsilon \in (\frac{3}{8}, \frac{5}{8})$ and $(1 - \epsilon)p_0(q^3 + 1) \geq \frac{aq^2 \log q}{2}$ and $(1 + \epsilon)p_0(q^3 + 1) \leq 2aq^2 \log q$. We have the following inequality:

$$\mathbb{P}(|P_{\mathcal{H}}| > (1 + \epsilon)p_0(q^3 + 1)) \leq e^{-\frac{9}{256}aq^2 \log q} \leq \frac{1}{6} \tag{4.2.2}$$

and

$$\mathbb{P}(|P_{\mathcal{H}}| < (1 - \epsilon)p_0(q^3 + 1)) \leq e^{-\frac{9}{128}aq^2 \log q} < \frac{1}{6}. \tag{4.2.3}$$

, so (2) fails with probability less tan $\frac{1}{3}$. For (3), similarly,

$$\mathbb{P}(|l| \geq \frac{1}{2}a \log q \leq e^{-\frac{1}{16}a \log q} < \frac{1}{6} \tag{4.2.4}$$

and

$$\mathbb{P}(|l| \leq 2a \log q) \leq e^{-2a \log q} < \frac{1}{6}. \tag{4.2.5}$$

So (2) or (3) fails with probability strictly less than $\frac{2}{3}$. With (2) and (3), for (4), if l_1, l_2 meet at the points of concurrency at the $(s + 1)$-fan, then there are at most $(2a \log q)^2$ choices of base lines, and for each chosen base line, there are at most $(2a \log q)^{s-2}$ choices for the remaining $s - 2$ lines crossing p. If one of l_1 and l_2 is the base line for the $(s + 1)$-fan, note that there are at most $2a \log q$ choices for the point of concurrency, and for each point of concurrency, there are at most $(2a \log q)^{s-1}$ choices for the remaining $s - 1$ lines. So (4) follows. □

Now for a subset $X \subseteq L_{\mathcal{H}}$, for any $p \in P_{\mathcal{H}}$, let $X_p := \{l \in L_{\mathcal{H}} : p \in l\}$ and for any $b \geq 1$, let $P_X = P_{X,b} := \{p \in P_H H : |X_p| \geq b\}$. We have the following lemma:

Lemma 4.2.3. For $b \geq 1$, $a \geq 128$, $q \geq a \log q$, and any $X \in L_{\mathcal{H}}$ where $\mathcal{H}$ is defined in the previous lemma,

$$\sum_{p \in P_X} |X_p| > \frac{1}{2}(a \log q)|X| - 2ab^2 \log q. \tag{4.2.6}$$

Proof. Note that

$$\sum_{p \in P_{\mathcal{H}} \setminus P_X} |X_p| < b|P_{\mathcal{H}}| \leq 2abq^2 \log q. \tag{4.2.7}$$

Also note that by double counting and property (2) of the previous lemma,

$$\sum_{p \in P_{\mathcal{H}}} |X_p| \geq \frac{1}{2}|X| \cdot a \log q. \tag{4.2.8}$$

The result follows. □

4.2.2 The blowup G_χ and the graph H

Let G be the intersection graph of $\mathcal{H}$ in lemma 4.2.2. Note that G is an edge-disjoint union of cliques K_p where $p \in P_{\mathcal{H}}$ and $K_p = K_{q^2}$, the complete graph with q^2 vertices. Now we s-blowup each K_p with the associated color classes $\chi = \{c_1, ..., c_s\}$ by considering $(p, \{l \in \mathcal{H} : p \in l\})$. For any $p \in P_{\mathcal{H}}$, we uniformly independently assign a color c from χ to each line l that intersects p. We then s-blowup each K_p by definition 4.1.1 to obtain a graph G_χ where each K_{s+1} in G_χ corresponds to an $(s+1)$-fan in $\mathcal{H}$, i.e. we've eliminated all K_{s+1}'s induced by $s+1$ concurrent lines. Now we let $b \geq 2^{40s}$ and $\rho = (\frac{8s}{b})^{\frac{2}{s}}$ and define $G_\rho = G_{n,\rho}$ to be the random graph on $V(G_\rho) = E(\mathcal{H})$ with edge probability ρ. We next let $H = G_\chi \cap G_\rho$ and for any $X \subseteq L_{\mathcal{H}} = V(H)$ and a point $p \in P_{\mathcal{H}}$, we fix a family $\Pi_p(x) = \Pi_p$ of $r_p(x) = \lfloor \frac{|x_p|}{b} \rfloor$ disjoint subsets each having size b. Then for any p, for any $Y_p \in \Pi_p$, we say Y_p is *bad* if $K_s \nsubseteq Y_p$. We say X_p is *bad* if all $Y_p \in \Pi_p$ are bad and X is *bad* if all X_p's are bad. We let A_{X_p} be the event that X_p is bad and $A_X$be the event that X is bad. We also let A_Y be the event that Y is bad. Note that $A_X = \bigcap_{p \in P_{\mathcal{H}}} A_{X_p}$ and if A_X does not occur, then X must contain a K_s.

Lemma 4.2.4. Let $s \geq 3$, $b \geq 2^{40s}$ and $\rho = (\frac{8s}{b})^{\frac{2}{s}}$. Then for any $X \subseteq V(H)$,

$$\mathbb{P}(A_X) \leq e^{-\frac{1}{32s} \sum_{p \in P_x} |X_p|}. \tag{4.2.9}$$

Proof. We first note that

$$\mathbb{P}(A_X) = \prod_{p \in P_{\mathcal{H}}} \mathbb{P}(A_{X,p}) = \prod_{p \in P_{\mathcal{H}}} \prod_{Y \in \Pi_p} \mathbb{P}(A_Y). \tag{4.2.10}$$

So it remains to find an upper bound for $P(A_Y)$. Now note that for any $Y \in \Pi_p$, $|Y| = b$ and for each color c of the blowup, the probability that it appears at most $\frac{b}{2s}$ times is less than $e^{-\frac{b}{8s}}$. So by the Chernoff bound,

$$\mathbb{P}(\text{some color } c \text{ in } Y \text{ appears at most } \frac{b}{2s} \text{ times}) \leq se^{-\frac{b}{8s}}. \tag{4.2.11}$$

For the blowup with coloring χ such that each color in Y appears $\frac{b}{2s}$ times, by proposition 4.1.2,

$$\mathbb{P}(K_s \subseteq G_{b,\rho}) \leq e^{-2^{2s-4}b}. \tag{4.2.12}$$

Since there are at most s^b colorings on Y,

$$\begin{aligned} \mathbb{P}(A_Y) &\leq se^{-\frac{b}{8s}} + s^b e^{-2^{2s-4}b} \\ &= e^{-\frac{b}{8s} + \log s} + e^{-2^{2s-4}b + b\log s} \\ &\leq e^{-\frac{b}{16s}} + e^{-\frac{b}{16s}} \end{aligned} \tag{4.2.13}$$

since $2^{2s-4}b \geq \frac{17}{16}b\log s$ and $b\log s \geq \frac{b}{s}$. So

$$\mathbb{P}(A_Y) \leq 2e^{-\frac{b}{16s}} \leq e^{-\frac{b}{24s}}. \tag{4.2.14}$$

Altogether,

$$\begin{aligned}\mathbb{P}(A_X) &= \prod_{p\in P_X}\prod_{Y\in\Pi_p}\mathbb{P}(A_Y)\\ &\leq \exp(-\sum_{p\in P_X}\lfloor\frac{|X_p|}{b}\rfloor\frac{b}{32s})\\ &\leq \exp(-\sum_{p\in P_X}\frac{2}{3}\frac{|X_p|}{b}\frac{b}{32s})\\ &= \exp(-\frac{1}{32s}\sum_{p\in P_X}|X_p|).\end{aligned} \tag{4.2.15}$$

□

4.3 Proof of Theorem 4.0.1

Let G be the intersection graph defined on lemma 4.2.2, $b = 2^{40s}a\log q$ and $\rho = (\frac{8s}{b})^{\frac{2}{s}}$. Let $\mathcal{K} := \{K \subseteq V(G) : K$corresponds to an $(s+1)$-fan. For any $K \in \mathcal{K}$, we say K is *bad* if $H[K]$ is an $(s+1)$-clique. Let A_K be the event that K is bad. Let $\mathcal{X}$ be the set $\{X \subseteq V(G) : |X| = 8bq^2\}$. Then H is K_{s+1}-free and does not contain any K_s-free set of size $8bq^2$ if none of A_X or A_K occurs over all $X \in \mathcal{X}$, $K \in \mathcal{K}$. For any large n, by Bertrand's Postulate, choose $q \in [\frac{1}{2}n^{\frac{1}{4}}, 2n^{\frac{1}{4}}]$ such that $n = cq^2(q^2 - q + 1)$ for some constant $c \in [0, 32]$ when q is large. Then if none of A_X of A_K occurs, we have $f_{s,s+1}(n) \leq 8bq^2 = O(\sqrt{n}\log n)$, which is essentially our result. So it remains to show that none of A_X or A_K occurs with positive probability, which will be proved through the Lovasz Local Lemma.

We first check the dependencies. Since A_X is an edge-disjoint union of A_{x_p}'s and each A_{X_p} is a disjoint union of A_Y's, we let

$$\hat{E}[X] = \bigsqcup_{p\in P_X}\bigsqcup_{Y\in\Pi_p} E(G[Y]) \tag{4.3.1}$$

to be all edges that will make A_X and A_K dependent for some $K \in \mathcal{K}$. Note that

$$\begin{aligned}|\hat{E}[X]| &= \sum_{p\in P_X}\sum_{Y\in\Pi_p}|E(G[Y])|\\ &= \sum_{p\in P_X}\lfloor\frac{|X_p|}{b}\rfloor\cdot\binom{b}{2}\\ &\leq \frac{b}{2}\sum_{p\in P_X}|X_p|.\end{aligned} \tag{4.3.2}$$

Note that by lemma 4.2.2 (4), each edge in $\hat{E}[X]$ is contained in at most $k = (2a\log q)^s$ K_{s+1}'s that are induced by $(s+1)$-fans, so the event A_X is dependent on at most

$$\lambda := k|\hat{E}[X]| \leq k\cdot\frac{b}{2}\cdot\sum_{p\in P_X}|X_p| \tag{4.3.3}$$

A_K's. Moreover, for each pair of $(s+1)$-cliques KK' induced by $(s+1)$-fans in $\mathcal{H}$, A_K and A'_K are dependent if and only if $V(K) \cap V(K') \neq \varnothing$. So A_K is dependent in at most

$$\kappa := \binom{s+1}{2} k \leq bk \tag{4.3.4}$$

$A_{k'}$'s. Note that A_X is at most dependent on λ A_K's and all other $A_{X'}$'s (where $X' \in \mathcal{X}$). Also note that A_K is dependent on at most κ $A_{k'}$;s and all A_X's. We now let $N = |\mathcal{X}|$. By Lovasz Local Lemma, it's enough to show the following lemma:

Lemma 4.3.1. For $\delta = \frac{1}{N+1}$, $\gamma = \frac{1}{64sbk}$, $\rho = (\frac{8s}{b})^{\frac{2}{s}}$, $k = (2a \log q)^s$, $b = 2^{40s} a \log q$, for any $K \in \mathcal{K}$, $X \in \mathcal{X}$,

(1). $\mathbb{P}(A_k) \leq \gamma(1-\gamma)^{\kappa}(1-\delta)^N$, and

(2). $\mathbb{P}(A_X) \leq \delta(1-\delta)^N(1-\gamma)^{\lambda}$.

Proof. FOr (1), note that

$$(1-\delta)^N = (1 - \frac{1}{N+1})^N \geq \frac{1}{2e} \tag{4.3.5}$$

and

$$(1-\gamma)^{\kappa} \geq 1 - \kappa\gamma \geq 1 - \frac{1}{32s} \geq \frac{1}{2}. \tag{4.3.6}$$

So it's enough to show that $\frac{4e\mathbb{P}(A_K)}{\gamma} < 1$.

$$\begin{aligned}
\frac{4e\mathbb{P}(A_K)}{\gamma} &= 256sbk\mathbb{P}(A_K) \\
&= 256sbk\rho^{\binom{s+1}{2}} \\
&= 256sbk(\frac{8s}{b})^{\frac{2}{s}\cdot\frac{s(s+1)}{2}} \\
&= \frac{256 \cdot 8^{s+1} \cdot s^{s+2} \cdot k}{b^s} \\
&= \frac{32 \cdot 8^{s+2} s^{s+2}}{2^{40s^2}} \\
&\leq (\frac{32s}{2^{40\frac{s^2}{s+2}}})^{s+2} < 1
\end{aligned} \tag{4.3.7}$$

since $\frac{32s}{2^{40\frac{s^2}{s+2}}} < 1$ for all $s \geq 3$. For (2), note that $1-\gamma \geq \exp(-2\gamma)$ as $\gamma < \frac{1}{2}$. Also since $(1-\delta)^N \geq \frac{1}{2e}$, it's enough to show that

$$\mathbb{P}(A_X) \leq \exp(-\log(N+1) - 2\gamma\lambda - 1 - \log 2). \tag{4.3.8}$$

By lemma 4.2.4, $\mathbb{P}(A_X) \leq \exp(-\frac{1}{32s}\sum_{p \in P_X}|X_p|)$, so it's enough to show that $\exp(-\frac{1}{32s}\sum_{p \in P_X}|X_p|) \leq \exp(-\log(N+1) - 2\gamma\lambda - 1 - \log 2)$. Note that when q is large, we have

$$\begin{aligned}
\log(N+1) &= \log(\binom{q^2(q^2-q+1)}{8bq^2} + 1) \\
&\leq \log(\binom{q^4}{8bq^2} + 1) \\
&\leq \log(q^{4\cdot 8bq^2}) - 1 - \log 2 = 32bq^2 \log q
\end{aligned} \tag{4.3.9}$$

and also $\gamma\lambda \leq \frac{1}{64s}\sum_{p\in P_x}|X_p|$. So it remains to show that

$$\frac{1}{32s}\sum_{p\in P_X}|X_p| \geq 32bq^2\log q + \frac{1}{64s}\sum_{p\in P_X}|X_p|, \tag{4.3.10}$$

i.e.,

$$\frac{1}{64s}\sum_{p\in P_X}|X_p| \geq 32bq^2\log q. \tag{4.3.11}$$

Note that by lemma 4.2.3,

$$\begin{aligned}\frac{1}{64s}\sum_{p\in P_X}|X_p| &> \frac{1}{64s}(\frac{1}{2}a(\log q)|X| - 2abq^2\log q)\\ &= \frac{1}{64s}(4abq^2\log q - 2abq^2\log q)\\ &= \frac{1}{32s}(2abq^2\log q)\end{aligned} \tag{4.3.12}$$

Now using $a = 2^{10}s$, (2) follows, which proves theorem 4.0.1. □

5 Bound on $f(n, F, G)$ for triangle-free F and $G = K_3$

This subsection will introduce the results from Ruzsa-Szemerédi theorem on hypergraphs [18] [21].

Theorem 5.0.1 (Verstraëte)**.**

$$f(n, F, K_3) = \sqrt{n}(\log n)^{O(\sqrt{\log n})}. \tag{5.0.1}$$

5.1 A Ruzsa-Szemerédi-type Argument

Ruzsa and Szemerédi in the 1970s [18] first gave a bound of the number of triples on the integer set $[n]$ satisfying that there are no six points that induces three triples on $[n]$. They connected their bound with Behrend 's lower bound [4]. In the world of hypergraphs, realizing triples to be hyperedges, one can show that through Ruzsa and Szemeredi's method, one can obtain a lower bound for the number of hyperedges on all hypergraphs that keeps the hypergraph linear, i.e. two hyperedges can intersect at at most one vertex, and triangle-free, i.e. no three hyperedges pairwise intersect. Verstraëte [21] generalized the method from Ruzsa-Szemerédi and Behrend and in principle can obtain a lower bound for the r-uniform hypergraphs for any well-defined r which will be reconstructed in this section. However, in this section, I will only show the case where $r = \Theta(\log n)$ to obtain the main theorem, i.e. theorem 5.0.1.

Definition 5.1.1. For an r-uniform hypergraph $\mathcal{H}$ ($r \in \mathbb{N}$), we call $\mathcal{H}$ *linear* if for any two distinct hyperedges in H intersect at at most one vertex. For any triple $\{e_1, e_2, e_3\} \subseteq E(\mathcal{H})$, we call it a *loose triangle* if any pair of the triple intersect at exactly one vertex.

Theorem 5.1.2 (Main theorem of this subsection). For some large enough N, if $v(F) = v_f$ for some fixed $v_f \in \mathbb{N}$, then there exists an N-vertex $r = \lceil\frac{4}{\log\frac{v_f}{v_f-1}}\log N\rceil$-uniform hypergraph $\mathcal{H}$ that is linear and loose-triangle-free.

To elucidate the statement of the above theorem, we have the following remark:

Remark 5.1.3. The setup of the theorem is that for some large enough N, we can find some set $\Gamma \subseteq [n]$ where $N = \binom{r+1}{2}n$ (so $\log N = \log n$) and

$$|\Gamma| \geq \frac{n}{e^{c\sqrt{\log n \log\log n}}}. \tag{5.1.1}$$

For all $x \in [n]$, we create hyperedges over the r levels $L_1, ..., L_r$, where L_k is a copy of $[kn]$ for all $k \in [r-1]$. In this setup, a hyperedge is made up of $(x \in L_1, x+a \in L_2, x+2a \in L_3, ..., x+(r-1)a \in L_r)$, where $a \in \Gamma$. It is linear since if $\exists x,\ y \in N,\ x \neq y,\ a_x, a_y \in \Gamma,\ c_1, c_2 \in [r-1],\ c_1 \neq c_2$, such that

$$\begin{cases} x + c_1 a_x = y + c_1 a_y \\ x + c_2 a_x = y + c_2 a_y, \end{cases} \tag{5.1.2}$$

then, taking differences of the two above equations, we get that $a_x = a_y$, which implies that $x = y$, which is a contradiction. In the following sub-sub section, we will argue that the set Γ that we find together with the construction of hyperedges above will give us a loose-triangle-free linear hypergraph $\mathcal{H}$.

5.1.1 A Behrend-Type Construction

The following construction is a generalization of the construction of 3-term arithmetic progression-free triple systems by Felix Behrend in 1946 [4].

The setup: Consider $\mathbb{R}^d$ for some d that will be set later. Consider the integer grid on $\mathbb{R}^d$. Let $s \geq 2$ also be some parameter that will be set up later. let $k \leq d(s-1)^2$ be some nonnegative integer. Consider the set $F_k(d, s) \in \mathbb{Z}$ where $\forall x \in F_k(d, s)$, $x = a_1 + a_2(rs-1) + ... + a_d(rs-1)^d$ where the d-tuple $(a_1, .., a_d)$ has the properties that $a_i \in [s-1]$ for all $i \in [d]$. Let A to be the collection of these d-tuples, then we have that $A_2^2 := \{||a||_2^2 : a \in A\} \subseteq [d(s-1)^2] \cup \{0\}$. Since we have s^d possible values under the restriction of the n-tuples, and A_2^2 takes no more than $d(s-1)^2+1$ possible values, there exists some $k \in [d(s-1)^2]$ such that

$$|\{a \in A : ||a||_2^2 = k\}| \geq \frac{s^d}{d(s-1)^2+1} > \frac{s^{d-2}}{d}. \tag{5.1.3}$$

Now let $\Gamma = \{a \in A : ||a||_2^2 = k\}$ for such k. Since

$$a_1 + (rs-1)a_2 + ... + a_d(rs-1)^{d-1} \leq (rs-1)^{d-1} \cdot 2 \cdot (s-1) \leq (rs-1)^d, \tag{5.1.4}$$

we let $n = (rs)^d$, so $s = \frac{n^{1/d}}{r}$ for some fixed $r = r(n) = \Theta(\log n)$. Following from the injectivity of n-ary expansions, we now have a linear injective function

$$h : (\mathbb{Z}^+)^d \to \mathbb{Z}^+, \qquad h((a_1, ..., a_d)) = a_1 + (rs-1)a_2 + ... + a_d(rs-1)^{d-1}. \tag{5.1.5}$$

Note that since h is linear, and since we selected integer points on the intersection of $\mathbb{Z}^d$ and $\sqrt{k}S^{d-1}$ (i.e. the unit sphere on $\mathbb{R}^d$ but with radius k), for any $a, b \in \Gamma$, l_{ab}, the line generated by a and b does not intersect any other vertex of Γ, which implies that there does not exist any $i, j \in \mathbb{N}$ and $b_1, b_2, b_3 \in \Gamma$ such that $ib_1 + jb_2 = (i+j)b_3$ and $ih(b_1) + jh(b_2) = (i+j)h(b_3)$.

Moreover,

$$\begin{aligned} |\Gamma| &\geq \frac{1}{d}s^{d-2} \\ &= \frac{n^{\frac{d-2}{d}}}{r^{d-2}d} \\ &= \frac{n^{1-\frac{2}{d}}}{dr^{d-2}} \\ &= n \cdot \frac{1}{dn^{\frac{2}{d}}r^{d-2}}, \end{aligned} \tag{5.1.6}$$

which implies

$$|\Gamma| \geq \max_d (n \cdot \frac{1}{dn^{\frac{2}{d}}r^{d-2}}). \tag{5.1.7}$$

To get a large lower bound for $|\Gamma|$, we now start to minimize $g(d) := dn^{\frac{2}{d}}r^{d-2}$. Note that $g(d)$ obtains minimum when

$$\begin{aligned} dn^{\frac{3}{d}} &= r^{d-2} \\ d^d n^2 &= r^{d^2-2d} \\ d^{d/2}n &= r^{(d/2)-d} \\ \frac{d}{2}\log d + \log n &= (\frac{d^2}{2} - d)\log r \\ \frac{\log r + 2d\log d + 2\log n}{\log r} &= d^2 - 2d + 1 = (d-1)^2. \end{aligned} \tag{5.1.8}$$

So

$$\begin{aligned} d - 1 &= \sqrt{\frac{\log r + 2d\log d + 2\log n}{\log r}} \\ &\leq \sqrt{\frac{\log n + 2d\log n + 2\log n}{\log r}} \\ &= \frac{(2d+3)\log n}{\log r} \\ &= \sqrt{2d+3}\sqrt{\frac{\log n}{\log r}}, \end{aligned} \tag{5.1.9}$$

which implies

$$2d + 3 \leq 3\sqrt{2d+3}\sqrt{\frac{\log n}{\log r}}, \tag{5.1.10}$$

which gives

$$d \leq c_1 \frac{\log n}{\log r} \tag{5.1.11}$$

for some $c_1 \in \mathbb{R}$. Note that since $s = \frac{n^{1/d}}{r} >> d$,we have $s^2 > d$, which implies

$$|\Gamma| > \max_d \frac{n^{1-4/d}}{r^{d-4}}. \tag{5.1.12}$$

Similar to the above minimization process, minimizing $n^{4/d} r^{d-4}$ gives $d = c_2\sqrt{\log n / \log r}$, which implies that for some constants α_1, α_2, c,

$$\begin{aligned} |\Gamma| &\geq \frac{n^{1-4/d}}{r^{d-4}} \\ &= n \cdot \frac{1}{n^{\alpha_1\sqrt{\frac{\log r}{\log n}}} \cdot r^{\alpha_2\sqrt{\frac{\log n}{\log r}}}} \\ &= \frac{n}{e^{c\sqrt{\log r \log n}}}. \end{aligned} \tag{5.1.13}$$

Now, by the construction in Remark 5.1.3 with the Γ obtained from the above setup, we have the following lemma:

Lemma 5.1.4. There does not exist $x, y, z \in [n]$, $b_1, b_2, b_3 \in \Gamma$ and $c_1, c_2, c_3 \in [r-1]$ such that for the r-hyperedges $(x, x+b_1, ..., x+(r-1)b_1)$, $(y, y+b_2, ..., y+(r-1)b_2)$, $(z, z+b_3, ..., z+(r-1)b_3)$,

$$\begin{cases} y + c_1 b_2 = z + c_1 b_3; \\ x + c_2 b_1 = z + c_2 b_3; \\ x + c_3 b_1 = y + c_3 b_2. \end{cases} \tag{5.1.14}$$

Proof. For the sake of contradiction, suppose not. WLOG, assume $c_1 < c_2 < c_3$, y, z intersect at level c_1, x, z intersect at level c_2, and x, y intersect at level c_3. Then we have the following system of equations:

$$\begin{cases} y + c_1 b_2 = z + c_1 b_3; \\ x + c_2 b_1 = z + c_2 b_3; \\ x + c_3 b_1 = y + c_3 b_2. \end{cases} \tag{5.1.15}$$

Note that $y = z + c_1 b_3 - c_1 b_2$, so we have

$$\begin{cases} x + c_2 b_1 = z + c_2 b_3; \\ x + c_3 b_1 = z + c_1 b_3 - c_1 b_2 + c_3 b_2. \end{cases} \tag{5.1.16}$$

Now note that $x - z = c_2 b_3 - c_2 b_1$. So

$$\begin{aligned} c_2 b_3 - c_2 b_1 + c_3 b_1 &= c_1 b_3 - c_1 b_2 + c_3 b_2 \\ c_2 b_3 + (c_3 - c_2) b_1 &= c_1 b_3 + (c_3 - c_1) b_2 \\ (c_2 - c_1) b_3 + (c_3 - c_2) b_1 &= (c_3 - c_1) b_2. \end{aligned} \tag{5.1.17}$$

Letting $i = c_2 - c_1$, $j = c_3 - c_2$, since $i + j = c_3 - c_1$, we have

$$(i + j) b_2 = j b_1 + i b_3, \tag{5.1.18}$$

which is a contradiction. $\square$

Therefore, to conclude on Section 5.1, we constructed a graph $\mathcal{H}$ on $N = \binom{r+1}{2} n$ vertices, where $r = \Theta(\log n) = \Theta(\log N)$ with the number of hyperedges

$$|E(\mathcal{H})| \geq \frac{N}{e^{c\sqrt{\log N \log\log N}}} \cdot n = \frac{N}{e^{c\sqrt{\log N \log\log N}}} \cdot \frac{N}{e^{\Theta(\log\log N)}} = \frac{N^2}{e^{c'\sqrt{\log N \log\log N}}}. \tag{5.1.19}$$

5.2 The Triangle-Free Graph $\mathcal{G}$

Let $\mathcal{G}$ be the graph such that $V(\mathcal{G}) = E(\mathcal{H})$ and $\forall v_1, v_2 \in \mathcal{G}$, v_1, v_2 are connected by an edge iff the corresponding hyperedges $e_1, e_2 \in \mathcal{H}$ intersects on some $v\ inV(\mathcal{H})$. Now, note that $\mathcal{G}$ consists of $V(\mathcal{H})$ edge-disjoint cliques K_v for all $v \in V(\mathcal{H})$: Indeed, since 2 edge-non-disjoint cliques imply that there are $e_1, e_2 \in E(\mathcal{H})$ s.t. $e_1 \cap e_2 \supseteq \{u, v\}$ for some $u, v \in V(\mathcal{H})$. Also, each clique-triple K_u, K_v, K_w on $\mathcal{G}$ are not pairwise-vertex-intersecting since otherwise it would contradict the loose-triangle-forbidden property. Let $\{u_1, ..., u_{|V(F)|}\}$ be a labelling of vertices of F. We now partition each K_v into $|V(F)|$ parts $A_{v,1}, A_{v,2}, ..., A_{v,|V(F)|}$, where each part has size either $\lceil |V(K_v)|/|V(F)| \rceil$ or $\lfloor |V(K_v)|/|V(F)| \rfloor$. We delete all edges both of whose ends are in the same part, i.e. $K_v[Av, i]$ has no edges for all i. For any $u_{v,i} \in A_{v,i}$, $u_{v,j} \in A_{v,j}$, there is an edge connecting $u_{v,i}$ and $u_{v,j}$ if and only if there is an edge between v_i and v_j in F. We call this new graph $\mathcal{G}'$ Now uniformly randomly choose $I \subseteq V(G)$ with $|I| = t$ for some parameter t that will be set up later. Let $t_v = |I \cap K_v|$ for all K_v. Then, by the fact that each vertex in $\mathcal{G}$ is an r-edge in $\mathcal{H}$,

$$\begin{aligned}
\mathbb{P}(I \text{ does not contain } F) &= \prod_{v \in V(\mathcal{H})} \mathbb{P}(I \text{ does not contain } F \text{ in } K_v) \\
&= \prod_{v \in V(\mathcal{H})} (|V(F)| \cdot (1 - \frac{1}{|V(F)|})^{t_v}) \\
&= |V(F)|^{|V(\mathcal{H})|} \cdot (1 - \frac{1}{|V(F)|})^{\sum_{v \in V(\mathcal{H})} t_v} \\
&= |V(F)|^{|V(\mathcal{H})|} \cdot (1 - \frac{1}{|V(F)|})^{rt}.
\end{aligned} \tag{5.2.1}$$

Denote $v_f = |V(F)|$. By linearity of expectation,

$$\begin{aligned}
\mathbb{E}[I] &\leq \binom{V(\mathcal{G})}{t} |V(F)|^{|V(\mathcal{H})|} (1 - \frac{1}{|V(F)|})^{rt} \\
&\leq (\frac{eN^2}{t})^t |V(F)|^{|V(\mathcal{H})|} (1 - \frac{1}{|V(F)|})^{rt} \\
&= \frac{e^t n^{2t}}{t^t} v_f^N (\frac{v_f - 1}{v_f})^{-\frac{4t}{\log \frac{v_f - 1}{v_f}} \log N}
\end{aligned} \tag{5.2.2}$$

Since

$$\mathbb{P}(I \subseteq V(\mathcal{G}) \text{ is not } F\text{-free for all } I) = 1 - \mathbb{P}(I \subseteq V(\mathcal{G}) \text{ is } F\text{-free for some } I), \tag{5.2.3}$$

we need $\mathbb{P}(I \subseteq V(\mathcal{G})$ is F-free for some $I) < 1$. So we have

$$\begin{aligned}
\frac{e^t n^{2t}}{t^t} v_f^N (\frac{v_f - 1}{v_f})^{-\frac{4t}{\log \frac{v_f - 1}{v_f}} \log N} &< 1 \\
t + 2t \log N - t \log t + N \log v_f - 4t \log N &< 0 \\
t(1 + 2 \log N - 4 \log N - \log t) + N \log v_f &< 0.
\end{aligned} \tag{5.2.4}$$

The above inequality holds if we choose $t = N$. So if we set $n' = \frac{N^2}{e^{c' \sqrt{\log N} \log \log N}}$, we have $f(n', F, K_3) \leq \sqrt{n'} (\log n')^{O(\sqrt{\log n'})}$, concluding the proof of theorem 5.0.1.

References

[1] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley and Sons, Jan. 2016.

[2] S. Ball. Finite geometry and combinatorial applications, Apr. 2015.

[3] S. Barwick and G. Ebert. *Unitals in Projective Planes*. Springer, Dec. 2010.

[4] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proceedings of the National Academy of Sciences*, 32(12):331–332, Dec. 1946.

[5] G. Birkhoff and J. V. Neumann. The logic of quantum mechanics. *The Annals of Mathematics*, 37(4):823, Oct. 1936.

[6] A. Bishnoi. *Finite geometry and combinatorial applications*.

[7] P. Dembowski. *Finite Geometries*. Springer Berlin Heidelberg, 1968.

[8] G. B. Folland. *Real Analysis*. John Wiley and Sons, Apr. 1999.

[9] L. Giuzzi. *Hermitian varieties over finite fields*. 2000.

[10] T. Gowers and O. Janzer. Improved bounds for the erdős-rogers function. *Advances in Combinatorics*, Feb. 2020.

[11] D. Hughes and F. Piper. *Projective Planes*. Springer, Nov. 1983.

[12] D. J. Kleitman, J. Shearer, and D. Sturtevant. Intersections of k-element sets. *Combinatorica*, 1(4):381–384, Dec. 1981.

[13] S. Mattheus and J. Verstraete. The asymptotics of $r(4,t)$. 2023.

[14] R. McEliece. *The Theory of Information and Coding*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2002.

[15] D. Mubayi and J. Verstraete. On the order of erdos-rogers functions. 2024.

[16] M. E. O'Nan. Automorphisms of unitary block designs. *Journal of Algebra*, 20(3):495–511, Mar. 1972.

[17] A. Putman. The fundamental theorem of projective geometry.

[18] I. Ruzsa and E. Szemeredi. Triple systems with no six points carrying three triangles. *Combinatorica*, 18, Jan. 1976.

[19] J. B. Shearer. On the independence number of sparse graphs. *Random Structures and Algorithms*, 7(3):269–271, Oct. 1995.

[20] J. Verstraete. Pseudorandom ramsey grpahs.

[21] J. Verstraete. Unpublished.