Testing Intersectingness of Uniform Families

Ishay Haviv
The Academic College of Tel Aviv-Yaffo, Tel Aviv 61083, Israel

Michal Parnas
The Academic College of Tel Aviv-Yaffo, Tel Aviv 61083, Israel

Abstract

A set family $\mathcal{F}$ is called intersecting if every two members of $\mathcal{F}$ intersect, and it is called uniform if all members of $\mathcal{F}$ share a common size. A uniform family $\mathcal{F} \subseteq \binom{[n]}{k}$ of k-subsets of $[n]$ is ε-far from intersecting if one has to remove more than $\varepsilon \cdot \binom{n}{k}$ of the sets of $\mathcal{F}$ to make it intersecting. We study the property testing problem that given query access to a uniform family $\mathcal{F} \subseteq \binom{[n]}{k}$, asks to distinguish between the case that $\mathcal{F}$ is intersecting and the case that it is ε-far from intersecting. We prove that for every fixed integer r, the problem admits a non-adaptive two-sided error tester with query complexity $O(\frac{\ln n}{\varepsilon})$ for $\varepsilon \geq \Omega((\frac{k}{n})^r)$ and a non-adaptive one-sided error tester with query complexity $O(\frac{\ln k}{\varepsilon})$ for $\varepsilon \geq \Omega((\frac{k^2}{n})^r)$. The query complexities are optimal up to the logarithmic terms. For $\varepsilon \geq \Omega((\frac{k^2}{n})^2)$, we further provide a non-adaptive one-sided error tester with optimal query complexity of $O(\frac{1}{\varepsilon})$. Our findings show that the query complexity of the problem behaves differently from that of testing intersectingness of non-uniform families, studied recently by Chen, De, Li, Nadimpalli, and Servedio (ITCS, 2024).

2012 ACM Subject Classification Theory of computation → Streaming, sublinear and near linear time algorithms; Mathematics of computing → Combinatorial algorithms

Keywords and phrases Intersecting family, Uniform family, Property testing

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2024.35

Category RANDOM

Related Version *Full Version*: `http://arxiv.org/abs/2404.11504`

Funding *Ishay Haviv*: Research supported in part by the Israel Science Foundation (grant No. 1218/20).

Acknowledgements We would like to thank the anonymous reviewers for their useful comments.

1 Introduction

A set family $\mathcal{F}$ is called intersecting if for every two sets $F_1, F_2 \in \mathcal{F}$, it holds that $F_1 \cap F_2 \neq \emptyset$. The study of intersecting families plays a central role in the area of extremal combinatorics with a particular attention dedicated to the uniform case, where all the sets of the family share a common size. One of the most influential results in this context is the Erdős–Ko–Rado theorem [4], proved in 1938 and published in 1961, which states that for integers n and k with $n \geq 2k$, the maximum size of an intersecting family of k-subsets of $[n] = \{1, 2, \ldots, n\}$ is $\binom{n-1}{k-1}$, attained by the families of all k-subsets that include a fixed element. Another prominent result, proved by Lovász [10] in 1978 settling a conjecture of Kneser [9] from 1955, asserts that for $n \geq 2k$, the family $\binom{[n]}{k}$ of all k-subsets of $[n]$ cannot be covered by fewer than $n - 2k + 2$ intersecting families. This result is tight, as follows by considering, for each $i \in [n - 2k + 1]$, the family of k-subsets of $[n]$ that include i, and the family of k-subsets of $[n] \setminus [n - 2k + 1]$. A more recent result, proved by Dinur and Friedgut [3] in 2009, provides a structural characterization for large intersecting families of k-subsets of $[n]$ when k is sufficiently smaller than n. It says, roughly speaking, that every such family is

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2024).
Editors: Amit Kumar and Noga Ron-Zewi; Article No. 35; pp. 35:1–35:15
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

approximately contained in an intersecting junta, that is, an intersecting family $\mathcal{J}$ over $[n]$, such that the membership of a set F in $\mathcal{J}$ depends only on $F \cap J$ for a fixed set $J \subseteq [n]$, where the size of J is determined by the precision of the containment (see Theorem 7).

In this paper, we investigate intersecting uniform families from the computational perspective of property testing. This field delves into the amount of data required for distinguishing objects that satisfy a prescribed property from those that significantly deviate from satisfying it. The objective is to design a randomized algorithm for this task, called a (two-sided error) tester, that succeeds with high constant probability and minimizes the query complexity, i.e., the number of queries to the input object. If the tester accepts objects that satisfy the given property with probability 1, we say that its error is one-sided. The tester is said to be non-adaptive if its queries are determined independently of the answers provided for prior queries. For a thorough introduction to the field of property testing, the reader is referred to, e.g., [8].

For integers n and k with $n \geq 2k$ and for a real $\varepsilon \in [0,1)$, we say that a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting if one has to remove more than $\varepsilon \cdot \binom{n}{k}$ of its sets to make it intersecting. In the property testing problem $\textsc{Intersecting}_{n,k,\varepsilon}$, we are given access to a family $\mathcal{F} \subseteq \binom{[n]}{k}$, represented by an indicator function $f : \binom{[n]}{k} \to \{0,1\}$, and the goal is to distinguish between the case that $\mathcal{F}$ is intersecting and the case that it is ε-far from intersecting. Note that the Erdős–Ko–Rado theorem [4] implies that every intersecting family $\mathcal{F} \subseteq \binom{[n]}{k}$ includes at most a k/n fraction of the sets of $\binom{[n]}{k}$. This gives impetus to studying the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem for a proximity parameter $\varepsilon = \varepsilon(n,k)$ that diminishes faster than the ratio k/n (see the discussion at the end of Section 3).

Our interest in testing intersectingness of uniform families is sparked and inspired by a recent paper of Chen, De, Li, Nadimpalli, and Servedio [2], who introduced and explored the analogue problem for general (non-uniform) families of sets. In their setting, the input consists of an indicator function $f : \{0,1\}^n \to \{0,1\}$ of a family $\mathcal{F}$ of subsets of $[n]$ (of any size), identified with their characteristic vectors in $\{0,1\}^n$, and the goal is to decide whether $\mathcal{F}$ is intersecting or ε-far from intersecting. Here, since the size of the domain is 2^n, a family is said to be ε-far from intersecting if more than $\varepsilon \cdot 2^n$ of its sets should be removed to make it intersecting. Chen et al. [2] proved that this problem admits a non-adaptive one-sided error tester with query complexity $\mathrm{poly}(n^{\sqrt{n \cdot \log(1/\varepsilon)}}, \frac{1}{\varepsilon})$. They further established a nearly matching lower bound, showing that the query complexity of every non-adaptive one-sided error tester for the problem is $2^{\Omega(\sqrt{n \cdot \log(1/\varepsilon)})}$, whenever $\varepsilon \in [2^{-n}, \varepsilon_0]$ for some constant $\varepsilon_0 > 0$. For non-adaptive two-sided error testers, they obtained a lower bound of $2^{\Omega(n^{1/4}/\sqrt{\varepsilon})}$ on the query complexity, assuming that $\varepsilon \in [1/\sqrt{n}, \varepsilon_0]$ for some constant $\varepsilon_0 > 0$. As will be shortly described, the results of the present paper highlight a significant difference between the behavior of the query complexity of testing intersectingness in the uniform and non-uniform settings.

1.1 Our Contribution

This paper studies the query complexity of the $\textsc{Intersecting}_{n,k,\varepsilon}$ testing problem. We offer nearly matching upper and lower bounds for various settings of the integers n and k and the proximity parameter $\varepsilon = \varepsilon(n,k)$. Let us mention already here that all of our upper bounds are proved via efficient testers, whose running time is polynomial in n. We further note that the results presented below are applicable for all integers n and k for which the conditions on ε allow it to be smaller than 1. The precise and formal statements are given in the upcoming technical sections.

Our first result furnishes a non-adaptive two-sided error tester for the case of $\varepsilon \geq \Omega((k/n)^r)$ for a fixed constant r. Its analysis borrows the structural characterization of large intersecting uniform families due to Dinur and Friedgut [3]. Note that the multiplicative factors hidden by the $O(\cdot)$ and $\Omega(\cdot)$ notations might depend on r.

▶ **Theorem 1** (Two-Sided Error Tester). *For every fixed integer r, for all integers n and k with $n \geq 2k$ and for any real $\varepsilon \geq \Omega((\frac{k}{n})^r)$, there exists a non-adaptive two-sided error tester for* $\textsc{Intersecting}_{n,k,\varepsilon}$ *with $O(\frac{\ln n}{\varepsilon})$ queries.*

In fact, we prove Theorem 1 in a stronger form, adopting the concept of tolerant property testing, introduced by Parnas, Ron, and Rubinfeld [12]. Strengthening the standard notion of property testing, a tolerant tester is required to accept not only objects that satisfy the given property, but also those that are close to satisfying it (and as usual, to reject objects that significantly deviate from the property). Accordingly, the tester given in Theorem 1 is shown to accept with high probability any family that can be made intersecting by removing relatively few of its sets (see Theorem 6). Note that a result of Tell [13] implies that a two-sided error is essentially inherent in tolerant property testing.

We next turn our attention to designing one-sided error testers for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem, wherein an intersecting family must be accepted with probability 1. A natural non-adaptive tester for this purpose, termed the canonical tester, selects m random sets from $\binom{[n]}{k}$, uniformly and independently, and checks whether they include two sets that demonstrate the non-intersectingness of the input family, namely, two disjoint sets within the family. This raises the combinatorial question, which might be of independent interest, of determining the smallest number $m = m(n, k, \varepsilon)$ of random sets from $\binom{[n]}{k}$ that guarantee with high probability a pair of disjoint sets that lie in a family $\mathcal{F} \subseteq \binom{[n]}{k}$, assuming that $\mathcal{F}$ is ε-far from intersecting. As our main technical contribution, we address this question for the case where $\varepsilon \geq \Omega((k^2/n)^r)$ for a fixed constant r. Our analysis yields the following result.

▶ **Theorem 2** (One-Sided Error Tester). *For every fixed integer r, for all integers n and k with $n \geq 2k$ and for any real $\varepsilon \geq \Omega((\frac{k^2}{n})^r)$, there exists a non-adaptive one-sided error tester for* $\textsc{Intersecting}_{n,k,\varepsilon}$ *with $O(\frac{\ln k}{\varepsilon})$ queries.*

Let us emphasize that the tester provided by Theorem 2 surpasses that of Theorem 1 in two respects: its error is one-sided, and its query complexity is lower, replacing the $\ln n$ term by $\ln k$. On the other hand, Theorem 2 requires ε to satisfy $\varepsilon \geq \Omega((k^2/n)^r)$, and is thus applicable only for $n \geq \Omega(k^2)$, whereas the two-sided error tester of Theorem 1 is applicable already for $n \geq \Omega(k)$.

For the special case of $r = 2$, we offer a notably simple analysis of the canonical tester, enabling us to enhance the query complexity achieved in Theorem 2 by getting rid of the logarithmic term. This gives the following result.

▶ **Theorem 3** (One-Sided Error Tester; $r = 2$). *For all integers n and k with $n \geq 2k$ and for any real $\varepsilon \geq \Omega((\frac{k^2}{n})^2)$, there exists a non-adaptive one-sided error tester for* $\textsc{Intersecting}_{n,k,\varepsilon}$ *with $O(\frac{1}{\varepsilon})$ queries.*

We further consider the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem for integers n and k satisfying $n = \alpha \cdot k$ for an arbitrary constant $\alpha \geq 2$. Interestingly, the canonical tester fails in this case, because its random samples are unlikely to include even a single pair of disjoint sets, unless the number of samples is exponential in n. Nevertheless, we show that for all constants $\alpha \geq 2$ and $\varepsilon \in (0, 1)$, the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem with $n = \alpha \cdot k$ admits a non-adaptive one-sided error tester with constant query complexity. The proof employs a result of Friedgut and Regev [7], and the details are given in Section 4.2 (see Theorem 17).

Our final result supplies a lower bound on the query complexity of the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem. Note that the lower bound applies even to adaptive two-sided error testers. While the result can be derived from a general result of [5], our proof in the full version of the paper explicitly presents hard instances of the problem.

▶ **Theorem 4** (Lower Bound). *For all integers n and k with $n \geq 2k$ and for any real $\varepsilon = \varepsilon(n,k)$ with $\binom{n}{k}^{-1} \leq \varepsilon < \frac{1}{2}$, the query complexity of every tester for $\textsc{Intersecting}_{n,k,\varepsilon}$ is $\Omega(\frac{1}{\varepsilon})$.*

It is noteworthy that Theorem 4 implies that the query complexities achieved by our testers for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem are nearly tight. Specifically, the query complexity obtained in Theorem 3 is tight up to a multiplicative constant, while those of Theorems 1 and 2 are tight up to multiplicative logarithmic terms. An intriguing task for further research would be to decide whether these logarithmic terms can be avoided. More ambitiously, it would be interesting to determine the query complexity of the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem for general values of n, k, and $\varepsilon = \varepsilon(n,k)$.

1.2 Proof Techniques

We provide here a high-level description of the ideas applied in the proofs of Theorems 1 and 2. Let us start with the proof of Theorem 1, which gives a non-adaptive two-sided error tester for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem with query complexity $O(\frac{\ln n}{\varepsilon})$, where $\varepsilon \geq \Omega((\frac{k}{n})^r)$ for a fixed integer r. Given access to a family $\mathcal{F} \subseteq \binom{[n]}{k}$, our tester attempts to decide whether $\mathcal{F}$ is approximately contained in an intersecting j-junta over $[n]$ for some integer j that depends solely on r. To do so, the tester picks random sets from $\binom{[n]}{k}$, uniformly and independently, and checks whether they lie in $\mathcal{F}$. These samples are used to estimate, for each intersecting j-junta $\mathcal{J}$ over $[n]$, the fraction of sets in $\binom{[n]}{k}$ that lie in $\mathcal{F} \setminus \mathcal{J}$. Our analysis shows that $O(\frac{\ln n}{\varepsilon})$ samples suffice for these estimations to be pretty accurate with high probability. Then, if those estimations indicate that $\mathcal{F}$ is approximately contained in some intersecting j-junta over $[n]$, the tester predicts that $\mathcal{F}$ is close to intersecting. Otherwise, relying on the aforementioned structural result of Dinur and Friedgut [3], the tester deduces that $\mathcal{F}$ is far from intersecting with high probability. Let us stress that the tester is two-sided error, because even if $\mathcal{F}$ is intersecting and is essentially aligned with some intersecting j-junta $\mathcal{J}$ over $[n]$, the random samples might wrongly suggest, with low probability, that $\mathcal{F}$ significantly deviates from $\mathcal{J}$. As previously noted, Theorem 1 is proved with respect to tolerant property testing. For the precise statement and argument, the reader is referred to Section 3.

We next outline the approach applied in the proof of Theorem 2, which establishes a non-adaptive one-sided error tester for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem with query complexity $O(\frac{\ln k}{\varepsilon})$, where $\varepsilon \geq \Omega((\frac{k^2}{n})^r)$ for a fixed integer r. As mentioned earlier, the proof of this result relies on the canonical tester, which selects random sets from $\binom{[n]}{k}$, uniformly and independently, and checks if they include two sets that violate the intersectingness of the given family. Therefore, our goal is to show that if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting, then a collection of $O(\frac{\ln k}{\varepsilon})$ random sets from $\binom{[n]}{k}$ includes with high probability two disjoint sets within $\mathcal{F}$. Consider a family $\mathcal{F} \subseteq \binom{[n]}{k}$, and assume that $\mathcal{F}$ is ε-far from intersecting. Notice that this in particular implies that $|\mathcal{F}| > \varepsilon \cdot \binom{n}{k}$, as otherwise, one could remove at most $\varepsilon \cdot \binom{n}{k}$ of its sets to make it intersecting.

Let us first suppose that the chosen random sets include, for each set $A \subseteq [n]$ of size at most $r-1$, a set $F_A \in \mathcal{F}$ with $F_A \cap A = \emptyset$. A key observation in our approach is that the number of sets in $\binom{[n]}{k}$ that intersect all of those sets F_A does not exceed $k^r \cdot \binom{n-r}{k-r}$. To see this, observe that every set that intersects them all includes an element $j_1 \in F_\emptyset$, an element $j_2 \in F_{\{j_1\}}$, an element $j_3 \in F_{\{j_1,j_2\}}$, and so on, up to an element $j_r \in F_{\{j_1,\ldots,j_{r-1}\}}$, where the r elements $j_1, \ldots, j_r$ are distinct. It thus follows that there exists a collection of at most k^r subsets of $[n]$ of size r, such that every set in $\binom{[n]}{k}$ that intersects all the sampled sets that lie in $\mathcal{F}$ contains at least one of the sets of the collection. This implies that the number of those sets does not exceed $k^r \cdot \binom{n-r}{k-r} \leq (\frac{k^2}{n})^r \cdot \binom{n}{k}$. Now, by combining our assumption on ε with the fact that $|\mathcal{F}| > \varepsilon \cdot \binom{n}{k}$, one can show that a random set from $\binom{[n]}{k}$ lies in $\mathcal{F}$ and is disjoint from at least one of the previously sampled sets that lie in $\mathcal{F}$ with non-negligible probability. Therefore, a few additional random sets from $\binom{[n]}{k}$ are expected to provide with high probability the desired witness for the non-intersectingness of $\mathcal{F}$.

The scenario discussed above, however, is somewhat oversimplified. It definitely might be the case that for some set $A \subseteq [n]$ of size at most $r-1$, the sets of $\mathcal{F}$ that are disjoint from A are quite rare, and as such, our random samples are not expected to include them. Following the terminology of [3], we say that such a set A captures $\mathcal{F}$ (see Definition 13). To deal with this situation, we show that if a set A of size at most $r-1$ captures $\mathcal{F}$, then it admits two disjoint subsets $B, C \subseteq A$, such that the sub-families $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{F}$, defined as the restrictions of $\mathcal{F}$ to the sets whose intersection with A is, respectively, B and C, are far from being cross-intersecting, that is, one has to remove plenty of sets from $\mathcal{F}_1$ and $\mathcal{F}_2$ so that every set of $\mathcal{F}_1$ will intersect every set of $\mathcal{F}_2$. Since B and C are disjoint, this essentially allows us to ignore from now on the elements of A, and to analyze the probability that the random samples include two disjoint sets, one from $\mathcal{F}_1$ and one from $\mathcal{F}_2$. By a slight variant of the key observation described above, we wish our samples to include, for each set $A' \subseteq [n] \setminus A$ of size at most $r-1$, a set $F_{A'} \in \mathcal{F}_2$ with $F_{A'} \cap A' = \emptyset$. If the samples include such sets, then it can be shown that a random set is expected to lie in $\mathcal{F}_1$ and be disjoint from some previously chosen set of $\mathcal{F}_2$ with non-negligible probability. However, if some set $A' \subseteq [n] \setminus A$ of size at most $r-1$ captures $\mathcal{F}_2$, then a suitable set $F_{A'}$ is unlikely to be found. In this case, we repeat the above process and further refine $\mathcal{F}_1$ and $\mathcal{F}_2$, fixing the intersections of their sets with A' to some disjoint subsets and keeping them far from cross-intersecting. It might be needed to repeat this process multiple times, but a crucial component of our argument shows that r iterations suffice. Indeed, after this number of iterations, it turns out that the refined family $\mathcal{F}_2$ is already too small to still satisfy the invariant that $\mathcal{F}_1$ and $\mathcal{F}_2$ are far from cross-intersecting.

To summarize, our analysis of the canonical tester shows that if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting for a sufficiently large $\varepsilon = \varepsilon(n,k)$, then there exist a set $A \subseteq [n]$ and two disjoint subsets $B, C \subseteq A$, such that (a) the restrictions $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{F}$ to the sets whose intersections with A are B and C, respectively, are far from cross-intersecting, and (b) no subset of $[n] \setminus A$ of size at most $r-1$ captures $\mathcal{F}_2$ (see Lemma 15). This allows us to show that $O(\frac{\ln k}{\varepsilon})$ sets chosen randomly from $\binom{[n]}{k}$ include with high probability a collection of sets of $\mathcal{F}_2$, such that relatively few sets of $\mathcal{F}_1$ intersect them all. Therefore, a few additional samples from $\binom{[n]}{k}$ are expected to include a set of $\mathcal{F}_1$ that is disjoint from one of the previously picked sets of $\mathcal{F}_2$, resulting in the desired witness for non-intersectingness (see Lemma 16). For the precise and full argument, the reader is referred to Section 4.1.2.

1.3 Related Work

The Erdős–Ko–Rado theorem [4] implies that for all integers n and k with $n \geq 2k$, every family $\mathcal{F} \subseteq \binom{[n]}{k}$ whose size exceeds $\binom{n-1}{k-1}$ is not intersecting. One may thus ask whether such a family $\mathcal{F}$ must include a set that is disjoint from many of the sets of $\mathcal{F}$. This question was recently investigated by Frankl and Kupavskii [6] and by Chau, Ellis, Friedgut, and

Lifshitz [1], who provided a positive answer in a strong sense. Namely, it was shown in [1] that for any $\delta > 0$, there exists some $\alpha > 0$, such that for all integers n and k with $n \geq \alpha \cdot k$, every family $\mathcal{F} \subseteq \binom{[n]}{k}$ of size $|\mathcal{F}| = \binom{n-1}{k-1} + 1$ includes a set that is disjoint from at least $(\frac{1}{2} - \delta) \cdot \binom{n-k-1}{k-1}$ of the sets of $\mathcal{F}$.

We point out that the combinatorial question arising in our analysis of the canonical tester for the $\text{INTERSECTING}_{n,k\varepsilon}$ problem is similar in spirit to the question studied in [6, 1]. Indeed, the analysis requires us to show that if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting, then a relatively small collection of random sets, chosen uniformly and independently from $\binom{[n]}{k}$, includes with high probability two disjoint sets that lie in $\mathcal{F}$. For such a statement, one has to show not only the existence of a set in $\mathcal{F}$ that is disjoint from many of the sets of $\mathcal{F}$, but that many of the sets of $\mathcal{F}$ satisfy this property. Yet, a crucial difference between the two concepts is that in ours, the family $\mathcal{F}$ is not assumed to exceed the Erdős–Ko–Rado threshold but rather to be ε-far from intersecting.

1.4 Outline

The rest of the paper is organized as follows. In Section 2, we provide some definitions and tools that will be used throughout the paper. In Section 3, we present and analyze our two-sided error tester for the $\text{INTERSECTING}_{n,k,\varepsilon}$ problem and confirm Theorem 1. Finally, in Section 4, we present and analyze the one-sided error canonical tester for the $\text{INTERSECTING}_{n,k,\varepsilon}$ problem and confirm Theorems 2 and 3. We further present there another one-sided error tester, appropriate for integers n and k with $n = \Theta(k)$. The proof of Theorem 4 can be found in the full version of the paper.

2 Preliminaries

Throughout the paper, we omit all floor and ceiling signs, whenever these are not crucial.

2.1 Intersecting Families

For integers n and k, let $\binom{[n]}{k}$ denote the family of all k-subsets of $[n] = \{1, \ldots, n\}$. A family $\mathcal{F} \subseteq \binom{[n]}{k}$ is called intersecting if for every two sets $F_1, F_2 \in \mathcal{F}$, it holds that $F_1 \cap F_2 \neq \emptyset$. For a real $\varepsilon \in [0, 1]$, we say that $\mathcal{F}$ is ε-close to intersecting if it is possible to make $\mathcal{F}$ intersecting by removing at most $\varepsilon \cdot \binom{n}{k}$ of its sets. Otherwise, we say that $\mathcal{F}$ is ε-far from intersecting.

For two families $\mathcal{F}_1, \mathcal{F}_2 \subseteq \binom{[n]}{k}$, the pair $(\mathcal{F}_1, \mathcal{F}_2)$ is called cross-intersecting if for every two sets $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$, it holds that $F_1 \cap F_2 \neq \emptyset$. For a real $\varepsilon \in [0, 1]$, we say that the pair $(\mathcal{F}_1, \mathcal{F}_2)$ is ε-close to cross-intersecting if it is possible to make $(\mathcal{F}_1, \mathcal{F}_2)$ cross-intersecting by removing at most $\varepsilon \cdot \binom{n}{k}$ of the sets of $\mathcal{F}_1$ and $\mathcal{F}_2$. Otherwise, we say that $(\mathcal{F}_1, \mathcal{F}_2)$ is ε-far from cross-intersecting. Note that if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting, then the pair $(\mathcal{F}, \mathcal{F})$ is ε-far from cross-intersecting.

A function $f : \binom{[n]}{k} \to \{0, 1\}$ is called intersecting, ε-close to intersecting, and ε-far from intersecting, if the family $f^{-1}(1)$ of the sets of $\binom{[n]}{k}$ that are mapped by f to 1 is, respectively, intersecting, ε-close to intersecting, and ε-far from intersecting. For reals $\varepsilon_1, \varepsilon_2 \in [0, 1)$ with $\varepsilon_1 \leq \varepsilon_2$, let $\text{INTERSECTING}_{n,k,\varepsilon_1,\varepsilon_2}$ denote the tolerant property testing problem that given query access to a function $f : \binom{[n]}{k} \to \{0, 1\}$, asks to distinguish between the case that f is ε_1-close to intersecting and the case that f is ε_2-far from intersecting. We will be particularly interested in the case of $\varepsilon_1 = 0$, hence for a real $\varepsilon \in [0, 1)$, we let $\text{INTERSECTING}_{n,k,\varepsilon}$ denote the $\text{INTERSECTING}_{n,k,0,\varepsilon}$ problem.

2.2 Chernoff–Hoeffding Bound

We will need the following version of the Chernoff–Hoeffding bound.

▶ **Theorem 5.** *For an integer m and a real $p \in (0,1)$, let $X_1, \dots, X_m$ be independent binary random variables satisfying* $\Pr[X_i = 1] = p$ *and* $\Pr[X_i = 0] = 1 - p$ *for all $i \in [m]$, and put* $\overline{X} = \frac{1}{m} \cdot \sum_{i=1}^{m} X_i$. *Then, for any $\mu \geq 0$,*

1. *if $p \leq \mu$, then* $\Pr\left[\ \overline{X} \geq 2\mu\ \right] \leq e^{-m \cdot \mu/3}$, *and*
2. *if $p \geq \mu$, then* $\Pr\left[\ \overline{X} \leq \mu/2\ \right] \leq e^{-m \cdot \mu/8}$.

Note that the assertion of Theorem 5 for $p = \mu$ follows from standard statements of the Chernoff–Hoeffding bound (see, e.g, [11, Theorem 2.3]). The cases of $p < \mu$ and $p > \mu$ stem by monotonicity.

3 Two-Sided Error Tester

In this section, we present and analyze a tolerant non-adaptive two-sided error tester for the $\text{Intersecting}_{n,k,\varepsilon_1,\varepsilon_2}$ problem and prove the following result. Its special case with $\varepsilon_1 = 0$ yields Theorem 1.

▶ **Theorem 6.** *For every integer $r \geq 2$, there exist constants $c_1 = c_1(r)$ and $c_2 = c_2(r)$ for which the following holds. For all sufficiently large integers n and k with $n \geq 2k$ and for all reals $\varepsilon_1, \varepsilon_2 \in [0,1)$ with $\varepsilon_2 \geq 4 \cdot \varepsilon_1 + c_1 \cdot (\frac{k}{n})^r$, there exists a tolerant non-adaptive two-sided error tester for $\text{Intersecting}_{n,k,\varepsilon_1,\varepsilon_2}$ with at most $c_2 \cdot \frac{\ln n}{\varepsilon_2}$ queries, running time polynomial in n, and success probability at least $2/3$.*

A crucial ingredient in the proof of Theorem 6 is the following theorem due to Dinur and Friedgut [3]. Here, a family $\mathcal{J}$ of subsets of $[n]$ is called a j-junta over $[n]$ if there exists a set $J \subseteq [n]$ of size $|J| = j$ such that the membership of a set F in $\mathcal{J}$ depends only on $F \cap J$.

▶ **Theorem 7** ([3]). *For every integer $r \geq 2$, there exist constants $j = j(r)$ and $a = a(r)$ for which the following holds. For all integers n and k with $n > 2k$ and $k > j$, if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is intersecting, then there exists an intersecting j-junta $\mathcal{J}$ over $[n]$ such that $|\mathcal{F} \setminus \mathcal{J}| \leq a \cdot \binom{n-r}{k-r}$.*

We are ready to prove Theorem 6.

Proof of Theorem 6. Fix an integer $r \geq 2$, and let $j = j(r)$ and $a = a(r)$ be the constants given in Theorem 7. Let n and k be sufficiently large integers. Since the theorem trivially holds for $n = 2k$ (with an appropriate c_1), it may be assumed that $n > 2k$ and $k > j$. For an integer m and for reals $\varepsilon_1, \varepsilon_2 \in [0,1)$ with $\varepsilon_1 \leq \varepsilon_2$, consider the following tester for the $\text{Intersecting}_{n,k,\varepsilon_1,\varepsilon_2}$ problem.

Input: Query access to a function $f : \binom{[n]}{k} \to \{0,1\}$.

1. Pick m sets $G_1, \dots, G_m$ uniformly and independently at random from $\binom{[n]}{k}$.
2. Query f for the value of $f(G_i)$ for each $i \in [m]$.
3. For each intersecting j-junta $\mathcal{J}$ over $[n]$, let $\alpha_{\mathcal{J}}$ denote the fraction of the chosen sets G_i that are mapped by f to 1 and do not lie in $\mathcal{J}$, that is,
$$\alpha_{\mathcal{J}} = \frac{1}{m} \cdot |\{i \in [m] \mid f(G_i) = 1 \text{ and } G_i \notin \mathcal{J}\}|.$$
4. If there exists an intersecting j-junta $\mathcal{J}$ over $[n]$ with $\alpha_{\mathcal{J}} \leq \frac{\varepsilon_2}{2}$, then accept, and otherwise reject.

The above tester is clearly non-adaptive. For the analysis of its success probability, suppose that

$$\varepsilon_2 \geq 4 \cdot \left(\varepsilon_1 + a \cdot \left(\frac{k}{n}\right)^r\right) \quad \text{and} \quad m \geq \frac{12 \cdot (j \cdot \ln n + 2^j + 2)}{\varepsilon_2}. \tag{1}$$

Let $f : \binom{[n]}{k} \to \{0,1\}$ be an input function, and consider the family $\mathcal{F} = f^{-1}(1)$. Our goal is to prove that if $\mathcal{F}$ is ε_1-close to intersecting then the tester accepts f with probability at least 2/3, and that if $\mathcal{F}$ is ε_2-far from intersecting then the tester rejects f with probability at least 2/3.

For each intersecting j-junta $\mathcal{J}$ over $[n]$, let $p_{\mathcal{J}}$ denote the fraction of the sets of $\binom{[n]}{k}$ that are mapped by f to 1 and do not lie in $\mathcal{J}$, that is,

$$p_{\mathcal{J}} = \frac{1}{\binom{n}{k}} \cdot \left|\left\{G \in \binom{[n]}{k} \;\middle|\; f(G) = 1 \text{ and } G \notin \mathcal{J}\right\}\right|.$$

We shall prove that, with high probability, the values of $p_{\mathcal{J}}$ are well approximated by the quantities $\alpha_{\mathcal{J}}$ that our tester computes. Let $\mathcal{E}$ denote the event that for every intersecting j-junta $\mathcal{J}$ over $[n]$, it holds that

1. if $p_{\mathcal{J}} \leq \frac{\varepsilon_2}{4}$, then $\alpha_{\mathcal{J}} < \frac{\varepsilon_2}{2}$, and
2. if $p_{\mathcal{J}} \geq \varepsilon_2$, then $\alpha_{\mathcal{J}} > \frac{\varepsilon_2}{2}$.

For a fixed intersecting j-junta $\mathcal{J}$ over $[n]$, apply Item 1 of Theorem 5 with $\mu = \frac{\varepsilon_2}{4}$ to obtain that if $p_{\mathcal{J}} \leq \frac{\varepsilon_2}{4}$, then $\alpha_{\mathcal{J}} \geq \frac{\varepsilon_2}{2}$ with probability at most $e^{-m \cdot \varepsilon_2/12}$. Further, apply Item 2 of Theorem 5 with $\mu = \varepsilon_2$ to obtain that if $p_{\mathcal{J}} \geq \varepsilon_2$, then $\alpha_{\mathcal{J}} \leq \frac{\varepsilon_2}{2}$ with probability at most $e^{-m \cdot \varepsilon_2/8} \leq e^{-m \cdot \varepsilon_2/12}$. The number of intersecting j-juntas over $[n]$ is clearly bounded by $\binom{n}{j} \cdot 2^{2^j}$. Therefore, by the union bound, the probability that there exists an intersecting j-junta $\mathcal{J}$ over $[n]$ for which either $p_{\mathcal{J}} \leq \frac{\varepsilon_2}{4}$ and $\alpha_{\mathcal{J}} \geq \frac{\varepsilon_2}{2}$, or $p_{\mathcal{J}} \geq \varepsilon_2$ and $\alpha_{\mathcal{J}} \leq \frac{\varepsilon_2}{2}$ does not exceed

$$\binom{n}{j} \cdot 2^{2^j} \cdot e^{-m \cdot \varepsilon_2/12} \leq n^j \cdot 2^{2^j} \cdot e^{-m \cdot \varepsilon_2/12} \leq e^{-2} < \tfrac{1}{3},$$

where the second inequality holds by our assumption on m given in (1). Therefore, the event $\mathcal{E}$ occurs with probability at least 2/3.

It suffices to show now that if the event $\mathcal{E}$ occurs, then the answer of our tester is correct. Suppose first that $\mathcal{F}$ is ε_1-close to intersecting. By definition, there exists an intersecting family $\mathcal{F}' \subseteq \binom{[n]}{k}$ such that $|\mathcal{F} \setminus \mathcal{F}'| \leq \varepsilon_1 \cdot \binom{n}{k}$. By Theorem 7, using $n > 2k$ and $k > j$, there exists an intersecting j-junta $\mathcal{J}$ over $[n]$ such that $|\mathcal{F}' \setminus \mathcal{J}| \leq a \cdot \binom{n-r}{k-r} \leq a \cdot (\frac{k}{n})^r \cdot \binom{n}{k}$. It thus follows that this $\mathcal{J}$ satisfies

$$|\mathcal{F} \setminus \mathcal{J}| \leq \varepsilon_1 \cdot \binom{n}{k} + |\mathcal{F}' \setminus \mathcal{J}| \leq \varepsilon_1 \cdot \binom{n}{k} + a \cdot \left(\frac{k}{n}\right)^r \cdot \binom{n}{k} \leq \frac{\varepsilon_2}{4} \cdot \binom{n}{k},$$

where the last inequality relies on our assumption on ε_1 and ε_2 in (1). This implies that $p_{\mathcal{J}} \leq \frac{\varepsilon_2}{4}$, and since the event $\mathcal{E}$ occurs, it follows that $\alpha_{\mathcal{J}} < \frac{\varepsilon_2}{2}$, hence our tester accepts f. Next, suppose that $\mathcal{F}$ is ε_2-far from intersecting. This implies, for each intersecting j-junta $\mathcal{J}$ over $[n]$, that $p_{\mathcal{J}} > \varepsilon_2$, as otherwise, one could remove at most $\varepsilon_2 \cdot \binom{n}{k}$ of the sets of $\mathcal{F}$ to obtain a sub-family of the intersecting family $\mathcal{J}$. Since the event $\mathcal{E}$ occurs, each such $\mathcal{J}$ satisfies $\alpha_{\mathcal{J}} > \frac{\varepsilon_2}{2}$, hence our tester rejects f.

Finally, let m be the smallest integer satisfying the condition in (1). Observe that the running time of our tester is polynomial in m and in the number of intersecting j-juntas over $[n]$, where j depends only on r. It follows that the running time is polynomial in n and $1/\varepsilon_2$, so by our assumption on ε_2 in (1), it is polynomial in n. This completes the proof. ◀

▶ Remark 8. The assumption on ε_1 and ε_2 in Theorem 6 can be weakened, by an almost identical proof, to $\varepsilon_2 \geq (1+\delta)\cdot\varepsilon_1 + c_1 \cdot (\frac{k}{n})^r$ with an arbitrary $\delta > 0$. For simplicity of presentation, we omit the details.

We conclude this section with the observation that for all integers n and k with $n \geq 2k$ and for all reals $\varepsilon_1, \varepsilon_2 \in [0,1)$ with $\varepsilon_2 \geq 4 \cdot (\varepsilon_1 + \frac{k}{n})$, there exists an efficient tolerant non-adaptive two-sided error tester for $\textsc{Intersecting}_{n,k,\varepsilon_1,\varepsilon_2}$ with $O(\frac{1}{\varepsilon_2})$ queries and high success probability. To see this, consider the tester that given access to a function $f : \binom{[n]}{k} \to \{0,1\}$ picks $O(\frac{1}{\varepsilon_2})$ sets uniformly and independently at random from $\binom{[n]}{k}$, queries f on them, and computes the fraction α of the chosen sets that are mapped by f to 1. If $\alpha \leq \frac{\varepsilon_2}{2}$, then the tester accepts, and otherwise it rejects.

For correctness, let p denote the fraction of the sets of $\binom{[n]}{k}$ that are mapped by f to 1. If f is ε_1-close to intersecting, then f can be made intersecting by flipping at most $\varepsilon_1 \cdot \binom{n}{k}$ of its values. By the Erdős–Ko–Rado theorem [4], the fraction of the sets of $\binom{[n]}{k}$ in an intersecting family does not exceed $\frac{k}{n}$, hence $p \leq \varepsilon_1 + \frac{k}{n} \leq \frac{\varepsilon_2}{4}$. In this case, Theorem 5 yields that with high probability, it holds that $\alpha \leq \frac{\varepsilon_2}{2}$ and the tester accepts. On the other hand, if f is ε_2-far from intersecting, then it holds that $p > \varepsilon_2$, so using again Theorem 5, it follows that with high probability, we have $\alpha > \frac{\varepsilon_2}{2}$ and the tester rejects.

4 One-Sided Error Tester

In this section, we present non-adaptive one-sided error testers for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem. We start with the canonical tester, which is used to prove Theorems 2 and 3. We then offer another non-adaptive one-sided error tester, appropriate for integers n and k with $n = \Theta(k)$.

4.1 Canonical Tester

Consider the following tester for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem.

Canonical Tester (n,k,m)
Input: Query access to a function $f : \binom{[n]}{k} \to \{0,1\}$.
1. Pick m sets $G_1, \ldots, G_m$ uniformly and independently at random from $\binom{[n]}{k}$.
2. Query f for the value of $f(G_i)$ for each $i \in [m]$.
3. If for every two indices $i_1, i_2 \in [m]$ with $G_{i_1} \cap G_{i_2} = \emptyset$, it holds that either $f(G_{i_1}) = 0$ or $f(G_{i_2}) = 0$, then accept, and otherwise reject.

The description of Canonical Tester immediately implies that it is non-adaptive and that it accepts every intersecting function with probability 1 and is thus one-sided error. In what follows, we analyze the number of queries m needed to reject with high probability any function ε-far from intersecting for $\varepsilon \geq \Omega((\frac{k^2}{n})^r)$, where r is a fixed integer. We start with the case of $r = 2$, for which we offer a simple analysis that yields a bound of $O(\frac{1}{\varepsilon})$ on the query complexity. We then consider the case of a general r, for which we obtain a bound of $O(\frac{\ln k}{\varepsilon})$ on the query complexity.

4.1.1 The case $r = 2$

We prove the following result, which confirms Theorem 3.

▶ **Theorem 9.** *For all integers n and k with $n \geq 2k$ and for any real $\varepsilon \in [0,1)$ with $\varepsilon \geq 2 \cdot (\frac{k^2}{n})^2$, there exists a non-adaptive one-sided error tester for* Intersecting$_{n,k,\varepsilon}$ *with $O(\frac{1}{\varepsilon})$ queries, running time polynomial in n, and success probability at least $2/3$.*

The proof of Theorem 9 requires two simple lemmas. The first one, given below, may be viewed as an analogue of a lemma of [2] for uniform families. Here, for a collection $\mathcal{M}$ of pairs of sets, we say that the pairs of $\mathcal{M}$ are pairwise disjoint if no set lies in more than one pair of $\mathcal{M}$.

▶ **Lemma 10.** *For integers n and k with $n \geq 2k$ and for a real $\varepsilon \in [0,1)$, if a family $\mathcal{F} \subseteq \binom{[n]}{k}$ is ε-far from intersecting, then there exists a collection of more than $\frac{\varepsilon}{2} \cdot \binom{n}{k}$ pairwise disjoint pairs (C, D) of sets of $\mathcal{F}$ satisfying $C \cap D = \emptyset$.*

Proof. Let $\mathcal{F} \subseteq \binom{[n]}{k}$ be a family ε-far from intersecting, and consider a maximal collection $\mathcal{M}$ (with respect to inclusion) of pairwise disjoint pairs (C, D) of sets of $\mathcal{F}$ satisfying $C \cap D = \emptyset$. The maximality of $\mathcal{M}$ implies that $\mathcal{F}$ can be made intersecting by removing all the $2 \cdot |\mathcal{M}|$ sets that lie in the pairs of $\mathcal{M}$. Since $\mathcal{F}$ is ε-far from intersecting, it follows that $2 \cdot |\mathcal{M}| > \varepsilon \cdot \binom{n}{k}$. This completes the proof. ◀

Using Lemma 10, we obtain the following.

▶ **Lemma 11.** *For integers n and k with $n \geq 2k$ and for a real $\varepsilon \in [0,1)$, let $\mathcal{F} \subseteq \binom{[n]}{k}$ be a family ε-far from intersecting with $|\mathcal{F}| > k^2 \cdot \binom{n-2}{k-2}$. Then, more than $\frac{\varepsilon}{2} \cdot \binom{n}{k}$ of the sets of $\mathcal{F}$ are disjoint from at least $\frac{1}{2} \cdot (|\mathcal{F}| - k^2 \cdot \binom{n-2}{k-2})$ of the sets of $\mathcal{F}$.*

Proof. Let $\mathcal{F} \subseteq \binom{[n]}{k}$ be a family ε-far from intersecting. By Lemma 10, there exists a collection $\mathcal{M}$ of more than $\frac{\varepsilon}{2} \cdot \binom{n}{k}$ pairwise disjoint pairs (C, D) of sets of $\mathcal{F}$ satisfying $C \cap D = \emptyset$. Fix a pair $(C, D) \in \mathcal{M}$. The disjointness of C and D implies that every set in $\binom{[n]}{k}$ that intersects both C and D includes some element of C and a different element of D. It follows that the number of sets in $\binom{[n]}{k}$ that intersect both C and D does not exceed $k^2 \cdot \binom{n-2}{k-2}$. Therefore, either C or D is disjoint from at least $\frac{1}{2} \cdot (|\mathcal{F}| - k^2 \cdot \binom{n-2}{k-2})$ of the sets of $\mathcal{F}$. Since the pairs of $\mathcal{M}$ are pairwise disjoint, it follows that more than $\frac{\varepsilon}{2} \cdot \binom{n}{k}$ sets of $\mathcal{F}$ are disjoint from at least $\frac{1}{2} \cdot (|\mathcal{F}| - k^2 \cdot \binom{n-2}{k-2})$ of the sets of $\mathcal{F}$, as required. ◀

We are ready to prove Theorem 9.

Proof of Theorem 9. For integers n, k, m and for a real $\varepsilon \in [0,1)$, consider the Canonical Tester (n, k, m) for the Intersecting$_{n,k,\varepsilon}$ problem. As previously mentioned, this tester is non-adaptive and one-sided error. To establish its correctness, it suffices to prove that if a function $f : \binom{[n]}{k} \to \{0,1\}$ is ε-far from intersecting, then the m sets picked by the tester include with probability at least $2/3$ two disjoint sets that lie in $f^{-1}(1)$. Indeed, this implies that the tester rejects such an f with the desired probability.

For a real $\varepsilon \geq 2 \cdot (\frac{k^2}{n})^2$, let $f : \binom{[n]}{k} \to \{0,1\}$ be a function ε-far from intersecting. Consider the family $\mathcal{F} = f^{-1}(1)$, and note that $\mathcal{F}$ is ε-far from intersecting. In particular, it holds that $|\mathcal{F}| > \varepsilon \cdot \binom{n}{k}$, because $\mathcal{F}$ can be made intersecting by removing all of its sets. This implies that

$$|\mathcal{F}| - k^2 \cdot \binom{n-2}{k-2} > \varepsilon \cdot \binom{n}{k} - k^2 \cdot \left(\frac{k}{n}\right)^2 \cdot \binom{n}{k} \geq \frac{\varepsilon}{2} \cdot \binom{n}{k}, \tag{2}$$

where the second inequality relies on our assumption on ε. We say that a set of $\binom{[n]}{k}$ is useful if it lies in $\mathcal{F}$ and is disjoint from at least $\frac{\varepsilon}{4} \cdot \binom{n}{k}$ of the sets of $\mathcal{F}$. By Lemma 11, combined with (2), the number of useful sets is greater than $\frac{\varepsilon}{2} \cdot \binom{n}{k}$. Hence, a random set picked uniformly from $\binom{[n]}{k}$ is useful with probability at least $\varepsilon/2$.

Now, consider the m sets chosen by CANONICAL TESTER. The probability that no useful set is picked throughout the first $4/\varepsilon$ choices does not exceed $(1-\varepsilon/2)^{4/\varepsilon} \leq e^{-2} < 1/6$. Therefore, with probability at least $5/6$, these choices include a useful set A. Since A is useful, it is disjoint from at least $\frac{\varepsilon}{4} \cdot \binom{n}{k}$ of the sets of $\mathcal{F}$. Therefore, once such a set A is chosen, the probability that a random set picked uniformly from $\binom{[n]}{k}$ lies in $\mathcal{F}$ and is disjoint from A is at least $\varepsilon/4$. It follows that the probability that no such set is picked throughout the next $8/\varepsilon$ choices does not exceed $(1-\varepsilon/4)^{8/\varepsilon} \leq e^{-2} < 1/6$. By the union bound, for $m = 12/\varepsilon$, the m sets picked by our tester include with probability at least $2/3$ two disjoint sets that lie in $\mathcal{F}$. Observe that for this choice of m, the running time of our tester is polynomial in n and $1/\varepsilon$. By our assumption on ε, the running time is polynomial in n, as desired. ◀

4.1.2 General r

We proceed by analyzing the more nuanced case in which $\varepsilon \geq \Omega((k^2/n)^r)$ for a general integer r. We prove the following result, which confirms Theorem 2.

▶ **Theorem 12.** *For every integer $r \geq 1$, there exist constants $c_1 = c_1(r)$ and $c_2 = c_2(r)$ for which the following holds. For all integers n and k with $n \geq 2k$ and for any real $\varepsilon \in [0,1)$ with $\varepsilon \geq c_1 \cdot (\frac{k^2}{n})^r$, there exists a non-adaptive one-sided error tester for* INTERSECTING$_{n,k,\varepsilon}$ *with at most $c_2 \cdot \frac{\ln k}{\varepsilon}$ queries, running time polynomial in n, and success probability at least* $2/3$.

The proof of Theorem 12 requires a few lemmas, which involve the following definition.

▶ **Definition 13.** *For integers n and k with $n \geq 2k$, a family $\mathcal{F} \subseteq \binom{[n]}{k}$, a set $A \subseteq [n]$, and a set $B \subseteq A$, we let $\mathcal{F}(A_{\downarrow B})$ denote the family of sets of $\mathcal{F}$ whose intersection with A is B, that is,*

$$\mathcal{F}(A_{\downarrow B}) = \{F \in \mathcal{F} \mid F \cap A = B\}.$$

We say that the set A ε-captures $\mathcal{F}$ if the number of sets of $\mathcal{F}$ that are disjoint from A is smaller than $\varepsilon \cdot \binom{n}{k}$, equivalently, $|\mathcal{F}(A_{\downarrow \emptyset})| < \varepsilon \cdot \binom{n}{k}$.

The following lemma shows that if two families are far from cross-intersecting, then for every small set A, it is possible to restrict the families to disjoint intersections with A, so that the obtained restrictions are still far from cross-intersecting. The proof can be found in the full version of the paper.

▶ **Lemma 14.** *For an integer $r \geq 0$, integers n and k with $n \geq 2k$, and a real $\varepsilon \in [0,1)$, let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \binom{[n]}{k}$ be two families such that the pair $(\mathcal{F}_1, \mathcal{F}_2)$ is ε-far from cross-intersecting, and let $A \subseteq [n]$ be a set of size $|A| \leq r$. Then, there exist two sets $B, C \subseteq A$ with $B \cap C = \emptyset$ for which the pair $(\mathcal{F}_1(A_{\downarrow B}), \mathcal{F}_2(A_{\downarrow C}))$ is $\frac{\varepsilon}{3^r}$-far from cross-intersecting. Moreover, if A $\frac{\varepsilon}{3^r}$-captures $\mathcal{F}_2$, then the guaranteed set C is not empty.*

As a consequence of Lemma 14, we obtain the following.

▶ **Lemma 15.** *For integers $r, t \geq 0$, integers n and k with $n \geq 2k$, and a real $\varepsilon \in [0,1)$, let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \binom{[n]}{k}$ be families such that the pair $(\mathcal{F}_1, \mathcal{F}_2)$ is ε-far from cross-intersecting. Then, there exist sets $A \subseteq [n]$ and $B, C \subseteq A$ with $B \cap C = \emptyset$, such that*

1. *the pair $(\mathcal{F}_1(A_{\downarrow B}), \mathcal{F}_2(A_{\downarrow C}))$ is $\frac{\varepsilon}{3^{r \cdot t}}$-far from cross-intersecting, and*
2. *either there is no subset of $[n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r \cdot t}}$-captures $\mathcal{F}_2(A_{\downarrow C})$, or $|C| \geq t$.*

Moreover, if $\varepsilon \geq 3^{r \cdot t} \cdot (\frac{k}{n})^t$, then the guaranteed set C satisfies $|C| < t$.[1]

[1] Note that by combining the second item of the lemma with the "moreover" part, it follows that if $\varepsilon \geq 3^{r \cdot t} \cdot (\frac{k}{n})^t$, then there is no subset of $[n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r \cdot t}}$-captures $\mathcal{F}_2(A_{\downarrow C})$.

Proof. Fix an integer $r \geq 0$. We start by proving the first part of the lemma, i.e., the existence of sets $A \subseteq [n]$ and $B, C \subseteq A$ with $B \cap C = \emptyset$ satisfying Items 1 and 2. To do so, we apply induction on t. The result clearly holds for $t = 0$, as follows from the choice $A = B = C = \emptyset$ (for which we have $\mathcal{F}_1 = \mathcal{F}_1(A_{\downarrow B})$, $\mathcal{F}_2 = \mathcal{F}_2(A_{\downarrow C})$, and $|C| \geq 0$).

Now, take $t \geq 1$, and assume that the result holds for $t-1$. To prove it for t, let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \binom{[n]}{k}$ be families such that the pair $(\mathcal{F}_1, \mathcal{F}_2)$ is ε-far from cross-intersecting. By the induction hypothesis, there exist sets $A \subseteq [n]$ and $B, C \subseteq A$ with $B \cap C = \emptyset$, such that the families $\mathcal{H}_1 = \mathcal{F}_1(A_{\downarrow B})$ and $\mathcal{H}_2 = \mathcal{F}_2(A_{\downarrow C})$ satisfy that

1. the pair $(\mathcal{H}_1, \mathcal{H}_2)$ is $\frac{\varepsilon}{3^{r\cdot(t-1)}}$-far from cross-intersecting, and
2. either there is no subset of $[n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r\cdot(t-1)}}$-captures $\mathcal{H}_2$, or $|C| \geq t-1$.

By $\frac{\varepsilon}{3^{r\cdot t}} \leq \frac{\varepsilon}{3^{r\cdot(t-1)}}$, it follows from Item 1 that the pair $(\mathcal{H}_1, \mathcal{H}_2)$ is $\frac{\varepsilon}{3^{r\cdot t}}$-far from cross-intersecting. Therefore, if there is no subset of $[n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r\cdot t}}$-captures $\mathcal{H}_2$, then the result holds for t with the same sets A, B, C, and we are done. Otherwise, there exists a set $A' \subseteq [n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r\cdot t}}$-captures $\mathcal{H}_2$. In particular, there exists a subset of $[n] \setminus A$ of size at most r that $\frac{\varepsilon}{3^{r\cdot(t-1)}}$-captures $\mathcal{H}_2$, hence by Item 2 above, it follows that $|C| \geq t-1$.

We apply Lemma 14 with the pair $(\mathcal{H}_1, \mathcal{H}_2)$, which is $\frac{\varepsilon}{3^{r\cdot(t-1)}}$-far from cross-intersecting, and with the set A'. Letting $B', C' \subseteq A'$ be the sets with $B' \cap C' = \emptyset$ guaranteed by the lemma, it follows that the pair $(\mathcal{H}_1(A'_{\downarrow B'}), \mathcal{H}_2(A'_{\downarrow C'}))$ is $\tilde{\varepsilon}$-far from cross-intersecting for $\tilde{\varepsilon} = \frac{1}{3^r} \cdot \frac{1}{3^{r\cdot(t-1)}} = \frac{1}{3^{r\cdot t}}$. Moreover, since A' $\tilde{\varepsilon}$-captures $\mathcal{H}_2$, it follows from the lemma that C' is not empty.

To complete the argument, define $A'' = A \cup A'$, $B'' = B \cup B'$, and $C'' = C \cup C'$. Recalling that $A \cap A' = \emptyset$, $B \cap C = \emptyset$, and $B' \cap C' = \emptyset$, we obtain that $B'' \cap C'' = \emptyset$ and that

$$|C''| = |C| + |C'| \geq (t-1) + 1 = t.$$

We further have $\mathcal{H}_1(A'_{\downarrow B'}) = \mathcal{F}_1(A''_{\downarrow B''})$ and $\mathcal{H}_2(A'_{\downarrow C'}) = \mathcal{F}_2(A''_{\downarrow C''})$. Hence, the sets A'', B'', C'' satisfy the required properties with respect to $\mathcal{F}_1, \mathcal{F}_2$ and t. This completes the proof of the first part of the lemma.

We finally show that if $\varepsilon \geq 3^{r\cdot t} \cdot (\frac{k}{n})^t$, then the guaranteed set C satisfies $|C| < t$. To establish the contrapositive statement, let A, B, C be the sets guaranteed by the lemma for two families $\mathcal{F}_1, \mathcal{F}_2 \subseteq \binom{[n]}{k}$, and suppose that $|C| \geq t$. Since the pair $(\mathcal{F}_1(A_{\downarrow B}), \mathcal{F}_2(A_{\downarrow C}))$ is $\frac{\varepsilon}{3^{r\cdot t}}$-far from cross-intersecting, it follows that $|\mathcal{F}_2(A_{\downarrow C})| > \frac{\varepsilon}{3^{r\cdot t}} \cdot \binom{n}{k}$. On the other hand, each set of $\mathcal{F}_2(A_{\downarrow C})$ contains the set C, hence $|\mathcal{F}_2(A_{\downarrow C})| \leq \binom{n-t}{k-t} \leq (\frac{k}{n})^t \cdot \binom{n}{k}$. By combining the two inequalities, we obtain that $\varepsilon < 3^{r\cdot t} \cdot (\frac{k}{n})^t$. This completes the proof. ◀

The following lemma constitutes a key ingredient in our analysis of the canonical tester. Its proof can be found in the full version of the paper.

▶ **Lemma 16.** *For every integer $r \geq 1$, there exists a constant $c = c(r)$, such that for all integers n and k with $n \geq 2k$ and for any real $\varepsilon \in [0,1)$ with $\varepsilon \geq 2 \cdot (\frac{k^2}{n})^r$, the following holds. Let $\mathcal{F} \subseteq \binom{[n]}{k}$ be a family, and suppose that there exist sets $A \subseteq [n]$ and $B, C \subseteq A$ with $B \cap C = \emptyset$, such that the pair $(\mathcal{F}(A_{\downarrow B}), \mathcal{F}(A_{\downarrow C}))$ is ε-far from cross-intersecting, and there is no subset of $[n] \setminus A$ of size at most $r-1$ that ε-captures $\mathcal{F}(A_{\downarrow C})$. Then, if at least $c \cdot \frac{\ln k}{\varepsilon}$ sets are chosen uniformly and independently at random from $\binom{[n]}{k}$, then the probability that they include two disjoint sets that lie in $\mathcal{F}$ is at least $2/3$.*

We are ready to prove Theorem 12.

Proof of Theorem 12. For integers n, k, m and for a real $\varepsilon \in [0,1)$, consider the Canonical Tester (n,k,m) for the $\text{Intersecting}_{n,k,\varepsilon}$ problem. As previously mentioned, this tester is non-adaptive and one-sided error. Fix an integer $r \geq 1$, let $c = c(r)$ be the constant given in Lemma 16, and suppose that

$$\varepsilon \geq 2 \cdot 3^{r^2} \cdot \left(\frac{k^2}{n}\right)^r \quad \text{and} \quad m \geq c \cdot 3^{r^2} \cdot \frac{\ln k}{\varepsilon}. \tag{3}$$

It suffices to prove that if a function $f : \binom{[n]}{k} \to \{0,1\}$ is ε-far from intersecting, then the m sets picked by our tester include with probability at least $2/3$ two disjoint sets that lie in $f^{-1}(1)$. Indeed, this implies that the tester rejects such an f with the desired probability.

Let $f : \binom{[n]}{k} \to \{0,1\}$ be a function ε-far from intersecting. Set $\mathcal{F} = f^{-1}(1)$, and note that the family $\mathcal{F}$ is ε-far from intersecting, hence the pair $(\mathcal{F},\mathcal{F})$ is ε-far from cross-intersecting. Apply Lemma 15 with $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{F}$ and with $t = r$, to obtain that there exist sets $A \subseteq [n]$ and $B, C \subseteq A$ with $B \cap C = \emptyset$, such that for $\tilde{\varepsilon} = \frac{\varepsilon}{3^{r^2}}$, it holds that

1. the pair $(\mathcal{F}(A_{\downarrow B}), \mathcal{F}(A_{\downarrow C}))$ is $\tilde{\varepsilon}$-far from cross-intersecting, and
2. either there is no subset of $[n] \setminus A$ of size at most r that $\tilde{\varepsilon}$-captures $\mathcal{F}_2(A_{\downarrow C})$, or $|C| \geq r$.

Moreover, our assumption on ε in (3) clearly implies that $\varepsilon \geq 3^{r^2} \cdot (\frac{k}{n})^r$, hence it follows from the lemma that $|C| < r$. We thus derive from Item 2 above that there is no subset of $[n] \setminus A$ of size at most r that $\tilde{\varepsilon}$-captures $\mathcal{F}(A_{\downarrow C})$. This obviously implies that there is no subset of $[n] \setminus A$ of size at most $r-1$ that $\tilde{\varepsilon}$-captures $\mathcal{F}(A_{\downarrow C})$.

Now, by our assumptions in (3), it holds that $\tilde{\varepsilon} \geq 2 \cdot (\frac{k^2}{n})^r$ and $m \geq c \cdot \frac{\ln k}{\tilde{\varepsilon}}$. Therefore, we can apply Lemma 16 with the family $\mathcal{F}$, the sets A, B, C, the integer r, and the real $\tilde{\varepsilon}$, to obtain that with probability at least $2/3$, the m sets picked by our tester include two disjoint sets that lie in $\mathcal{F}$, as required.

Finally, let m be the smallest integer satisfying the condition in (3). Observe that the running time of our tester is polynomial in n and m, hence it is polynomial in n and $1/\varepsilon$. By our assumption on ε in (3), the running time is polynomial in n, and the proof is completed. ◀

4.2 The case $n = \Theta(k)$

We turn our attention now to the $\text{Intersecting}_{n,k,\varepsilon}$ problem, where the integers n and k satisfy $n = \alpha \cdot k$ for an arbitrary constant $\alpha \geq 2$. We first observe that for this range of parameters, the canonical tester is not effective, even for a constant ε. To see this, observe that two random sets, chosen uniformly and independently from $\binom{[n]}{k}$, are disjoint with probability $\binom{n-k}{k}/\binom{n}{k}$. For $n = \alpha \cdot k$, with $\alpha \geq 2$ being a constant, this probability decreases exponentially in n. Therefore, a collection of random sets from $\binom{[n]}{k}$ is unlikely to include even a single pair of disjoint sets, unless their number grows exponentially in n. As a result, the canonical tester does not provide a useful upper bound on the query complexity of the $\text{Intersecting}_{n,k,\varepsilon}$ problem in this setting.

To overcome this difficulty, we consider a slightly different tester for the $\text{Intersecting}_{n,k,\varepsilon}$ problem, which picks random pairs of disjoint sets from $\binom{[n]}{k}$ and checks if at least one of the pairs demonstrates a violation of intersectingness for the tested function. This tester allows us to prove the following result.

▶ **Theorem 17.** *For all reals $\alpha \geq 2$ and $\varepsilon \in (0,1)$, there exists some $c = c(\alpha,\varepsilon)$, such that for all sufficiently large integers n and k with $n = \alpha \cdot k$, there exists a non-adaptive one-sided error tester for $\text{Intersecting}_{n,k,\varepsilon}$ with c queries and success probability at least $2/3$.*

Theorem 17 is obtained using the following result, that was proved (in a generalized form) by Friedgut and Regev [7].

▶ **Theorem 18** ([7]). *For all reals $\alpha > 2$ and $\varepsilon \in (0,1)$, there exist a real $\delta = \delta(\alpha, \varepsilon) > 0$ and an integer $j = j(\alpha, \varepsilon)$, such that for all sufficiently large integers n and k with $n = \alpha \cdot k$, the following holds. Suppose that $\mathcal{F} \subseteq \binom{[n]}{k}$ is a family, such that for every intersecting j-junta $\mathcal{J}$ over $[n]$, it holds that $|\mathcal{F} \setminus \mathcal{J}| > \varepsilon \cdot \binom{n}{k}$. Then, a random unordered pair $\{A, B\}$ of two disjoint sets of $\binom{[n]}{k}$, chosen uniformly from all such pairs, satisfies $A \in \mathcal{F}$ and $B \in \mathcal{F}$ with probability at least δ.*

Proof of Theorem 17. Fix $\alpha \geq 2$ and $\varepsilon \in (0,1)$, and let n and k be sufficiently large integers with $n = \alpha \cdot k$. For an integer m, consider the following tester for the $\textsc{Intersecting}_{n,k,\varepsilon}$ problem.

> Input: Query access to a function $f : \binom{[n]}{k} \to \{0,1\}$.
> 1. Pick m unordered pairs $\{A_1, B_1\}, \dots, \{A_m, B_m\}$ of sets of $\binom{[n]}{k}$ with $A_i \cap B_i = \emptyset$ for all $i \in [m]$ uniformly and independently at random.
> 2. Query f for the values of $f(A_i)$ and $f(B_i)$ for each $i \in [m]$.
> 3. If for each $i \in [m]$, it holds that $f(A_i) = 0$ or $f(B_i) = 0$, then accept, and otherwise reject.

The above tester is clearly non-adaptive. Since it accepts every intersecting function with probability 1, its error is one-sided. It suffices to show that there exists $m = m(\alpha, \varepsilon)$, such that if a function $f : \binom{[n]}{k} \to \{0,1\}$ is ε-far from intersecting, then our tester rejects it with probability at least 2/3. Let $f : \binom{[n]}{k} \to \{0,1\}$ be a function ε-far from intersecting.

Consider first the simple case of $\alpha = 2$. Here, the fact that f is ε-far from intersecting implies that more than $\varepsilon \cdot \binom{n}{k}$ of the $\frac{1}{2} \cdot \binom{n}{k}$ unordered pairs of disjoint sets of $\binom{[n]}{k}$ violate the intersectingness. For $m = 1/\varepsilon$, the probability that our tester accepts f is at most $(1-2\varepsilon)^m = (1-2\varepsilon)^{1/\varepsilon} \leq e^{-2}$. Hence, with the complement probability, which exceeds 2/3, our tester rejects f, as desired.

Next, suppose that $\alpha > 2$, let $\delta = \delta(\alpha, \varepsilon) > 0$ and $j = j(\alpha, \varepsilon)$ be the constants given in Theorem 18, and put $\mathcal{F} = f^{-1}(1)$. Since $\mathcal{F}$ is ε-far from intersecting, it follows that every intersecting j-junta $\mathcal{J}$ over $[n]$ satisfies $|\mathcal{F} \setminus \mathcal{J}| > \varepsilon \cdot \binom{n}{k}$. By Theorem 18, at least δ fraction of the unordered pairs of disjoint sets of $\binom{[n]}{k}$ violate the intersectingness. Letting $m = 2/\delta$, the probability that our tester accepts f is at most $(1-\delta)^m = (1-\delta)^{2/\delta} \leq e^{-2}$. As before, it follows that with probability at least 2/3, our tester rejects f, and we are done. ◀

References

1. Hou Tin Chau, David Ellis, Ehud Friedgut, and Noam Lifshitz. On the maximum degree of induced subgraphs of the Kneser graph. *arXiv*, abs/2312.06370, 2023. arXiv:2312.06370.
2. Xi Chen, Anindya De, Yuhao Li, Shivam Nadimpalli, and Rocco A. Servedio. Testing intersecting and union-closed families. In *Proc. of the 15th Innovations in Theoretical Computer Science Conference (ITCS'24)*, pages 33:1–33:23, 2024.
3. Irit Dinur and Ehud Friedgut. Intersecting families are essentially contained in juntas. *Comb. Probab. Comput.*, 18(1–2):107–122, 2009.
4. Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quart. J. Math.*, 12(1):313–320, 1961.
5. Eldar Fischer. A basic lower bound for property testing. *arXiv*, abs/2403.04999, 2024. arXiv:2403.04999.
6. Peter Frankl and Andrey Kupavskii. Maximal degrees in subgraphs of Kneser graphs. *arXiv*, abs/2004.08718, 2020. arXiv:2004.08718.

7 Ehud Friedgut and Oded Regev. Kneser graphs are like Swiss cheese. *Discrete Analysis*, 2:1–18, 2018.

8 Oded Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017.

9 Martin Kneser. Aufgabe 360. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 58(2):27, 1955.

10 László Lovász. Kneser's conjecture, chromatic number, and homotopy. *J. Comb. Theory, Ser. A*, 25(3):319–324, 1978.

11 Colin McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

12 Michal Parnas, Dana Ron, and Ronitt Rubinfeld. Tolerant property testing and distance approximation. *J. Comput. Syst. Sci.*, 72(6):1012–1042, 2006.

13 Roei Tell. A note on tolerant testing with one-sided error. In *Computational Complexity and Property Testing – On the Interplay Between Randomness and Computation*, pages 173–177. Springer, 2020.