K₄-INTERSECTING FAMILIES OF GRAPHS

AARON BERGER AND YUFEI ZHAO

ABSTRACT. Ellis, Filmus, and Friedgut proved an old conjecture of Simonovits and Sós showing that the maximum size of a triangle-intersecting family of graphs on n vertices has size at most $2^{\binom{n}{2}-3}$, with equality for the family of graphs containing some fixed triangle. They conjectured that their results extend to cross-intersecting families, as well to K_t -intersecting families. We prove these conjectures for $t \in \{3, 4\}$, showing that if \mathcal{F}_1 and \mathcal{F}_2 are families of graphs on n labeled vertices such that for any $G_1 \in \mathcal{F}_1$ and $G_2 \in \mathcal{F}_2$, $G_1 \cap G_2$ contains a K_t , then $|\mathcal{F}_1||\mathcal{F}_2| \leq 4^{\binom{n}{2}-\binom{t}{2}}$, with equality if and only if $\mathcal{F}_1 = \mathcal{F}_2$ consists of all graphs that contain some fixed K_t . We also establish a stability result. More generally, " $G_1 \cap G_2$ contains a K_t " can be replaced by " G_1 and G_2 agree on a non-(t-1)-colorable graph."

1. INTRODUCTION

We say that a family \mathcal{F} of graphs on a common vertex set is *triangle-intersecting* if, for every $G_1, G_2 \in \mathcal{F}, G_1 \cap G_2$ contains a triangle. A classic problem due to Simonovits and Sós from 1976 asks to determine the largest triangle-intersecting family of graphs on n labeled vertices. They conjectured that the maximizing family is obtained by taking all graphs containing some fixed triangle, which has size $2^{\binom{n}{2}-3}$. Triangle-intersecting families are a graph theoretic analogue of intersecting set families, which have been extensively studied, with the Erdős–Ko–Rado theorem [6] being a fundamental result. On the other hand, intersecting graph families are much less understood, and many standard techniques for studying intersecting set families (e.g., shifting) do not easily adapt to the graph setting.

The first significant progress on the Simonovits–Sós conjecture was due to Chung, Graham, Frankl, and Shearer [2], who introduced a powerful entropy lemma (now commonly known as Shearer's lemma) and used it to prove an upper bound of $2^{\binom{n}{2}-2}$, though it misses the conjecture by a factor of 2. A significant breakthrough was obtained by Ellis, Filmus, and Friedgut [4], who used Fourier analytic methods to prove the Simonovits–Sós conjecture along with several strengthenings.

We say that a family \mathcal{F} of graphs on a common vertex set is *H*-intersecting if for every $G_1, G_2 \in \mathcal{F}, G_1 \cap G_2$ contains *H* as a subgraph. Here is a natural extension of the above problem.

Problem 1.1. Given H and n, determine the maximum size of an H-intersecting family of graphs on n labeled vertices.

A family of graphs on a common set of vertices is called an *H*-umvirate if the family consists of all graphs containing some fixed copy of *H*. Clearly an *H*-umvirate family on *n* vertices is *H*-intersecting and has size $2^{\binom{n}{2}-e(H)}$. For a fixed *H*, is an *H*-umvirate family is the largest *H*intersecting family? It turns out that the answer can be no (see discussions at the end of this section), but it is conjectured [4] that the answer is yes whenever *H* is a clique.

Conjecture 1.2. Every K_t -intersecting family of graphs on n vertices has size at most $2^{\binom{n}{2}-\binom{t}{2}}$.

Berger was supported by NSF Graduate Research Fellowship DGE-1745302. Zhao was supported by NSF Award DMS-1764176, the MIT Solomon Buchsbaum Fund, and a Sloan Research Fellowship.

Ellis, Filmus, and Friedgut proved the result for t = 3. Applying tools developed by Friedgut [7], they also proved uniqueness and stability of the maximizer, showing that a triangle-intersecting family with size nearly the maximum size must be close to a triangle-unvirate.

Theorem 1.3 ([4, Theorem 1.4]). Let \mathcal{F} be a triangle-intersecting family of graphs on n labeled vertices.

- (Upper bound) $|\mathcal{F}| \leq 2^{\binom{n}{2}-3}$
- (Uniqueness) The upper bound is an equality if and only if \mathcal{F} is a triangle-unvirate.
- (Stability) There exists an absolute constant C > 0 so that for all $\epsilon > 0$, if $|\mathcal{F}| \ge (1 \epsilon)2^{\binom{n}{2}-3}$, then there exists a triangle-unvirate \mathcal{T} such that $|\mathcal{T} \triangle \mathcal{F}| \le C\epsilon 2^{\binom{n}{2}}$. (Here \triangle denotes the set symmetric difference.)

Results about intersecting families of sets are often extended to cross-intersecting families, and [4] conjectures that the same extensions hold here as well.

Definition 1.4 (Cross intersecting families). A pair $(\mathcal{F}_1, \mathcal{F}_2)$ of families of graphs on a common vertex set is said to be *cross-H-intersecting* if for any $G_1 \in \mathcal{F}_1, G_2 \in \mathcal{F}_2$ one has that $G_1 \cap G_2$ contains a subgraph isomorphic to H.

Conjecture 1 in [4] says that if $(\mathcal{F}_1, \mathcal{F}_2)$ is a cross-triangle-intersecting pair of family of graphs on *n* labeled vertices, then $|\mathcal{F}_1| |\mathcal{F}_2| \leq 4^{\binom{n}{2}-3}$, with equality if and only if $\mathcal{F}_1 = \mathcal{F}_2$ is a triangleunvirate. We prove this conjecture along with its K_4 -intersecting analogue. Our solution applies the framework set up in [4], and where they need to verify certain inequalities by hand via casework, we come up with systematic way to potentially verify the conjecture for any fixed *t* via a finite computation, though the complexity of the computation increases extremely quickly with *t*.

Theorem 1.5. Let $t \in \{3,4\}$. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross- K_t -intersecting pair of families graphs on n labeled vertices.

- (Upper bound) $|\mathcal{F}_1| |\mathcal{F}_2| \le 4^{\binom{n}{2} \binom{t}{2}}$
- (Uniqueness) The upper bound is an equality if and only if $\mathcal{F}_1 = \mathcal{F}_2$ is a K_t -unvirate.
- (Stability) There exists a constant $C_t > 0$ depending only on t so that for all $\epsilon > 0$, if $|\mathcal{F}_1| |\mathcal{F}_2| \ge (1-\epsilon)4^{\binom{n}{2}-\binom{t}{2}}$, then there exists a K_t -unvirate \mathcal{T} such that $|\mathcal{F}_1 \triangle \mathcal{T}|, |\mathcal{F}_2 \triangle \mathcal{T}| \le C_t \epsilon 2^{\binom{n}{2}}$.

Conjecture 1.6. Theorem 1.5 holds for all $t \ge 5$.

Theorem 1.3 was proved in [4] in a more general context of odd-cycle-agreeing families. Our results also hold in this generality. We write $G_1 \triangle G_2$ for the edge symmetric difference of two graphs, and \overline{G} for the edge-complement of a graph.

Definition 1.7. A pair $(\mathcal{F}_1, \mathcal{F}_2)$ of families of graphs on a common vertex is said to be *cross-t-chromatic-intersecting* if for any $G_1 \in \mathcal{F}_1$, $G_2 \in \mathcal{F}_2$, $G_1 \cap G_2$ is not (t-1)-colorable, and is said to be *cross-t-chromatic-agreeing* if for any $G_1 \in \mathcal{F}_1$, $G_2 \in \mathcal{F}_2$, $\overline{G_1 \triangle G_2}$ is not (t-1)-colorable (equivalently, there is some t-chromatic graph on which G_1 and G_2 agree).

A cross- K_t -intersecting pair is clearly cross-t-chromatic-agreeing. As stated below, Theorem 1.5 holds also for this stronger notion, and we conjecture that it holds for all t. Here the equality case needs to be changed from being a K_t -unvirate to an K_t -conjunction, which is defined to be a family of all graphs with a prescribed intersection with some given copy of K_t . In other words, \mathcal{F} is a K_t -conjunction if and only if there is some graph H on the same common vertex set such that $\{G \triangle H : G \in \mathcal{F}\}$ is a K_t -unvirate.¹

¹[4] uses " K_t -junta" to mean what we are calling K_t -conjunction here, which differs from the common usage of "junta" to mean any function determined by its restriction on a fixed set.

Theorem 1.8. Let $t \in \{3,4\}$. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-agreeing pair of families graphs on n labeled vertices.

- (Upper bound) $|\mathcal{F}_1| |\mathcal{F}_2| \le 4^{\binom{n}{2} \binom{t}{2}}$
- (Uniqueness) The upper bound is an equality if and only if F₁ = F₂ is a K_t-conjunction.
 (Stability) There exists a constant C_t > 0 depending only on t so that for all ε > 0, if $|\mathcal{F}_1||\mathcal{F}_2| \geq (1-\epsilon)4^{\binom{n}{2}-\binom{t}{2}}$, then there exists a K_t -conjunction \mathcal{T} such that $|\mathcal{F}_1 \triangle \mathcal{T}|, |\mathcal{F}_2 \triangle \mathcal{T}| \leq 1-\epsilon$ $C_{4\epsilon} 2^{\binom{n}{2}}$

The majority of the paper (Sections 3 and 4) is devoted to proving Theorem 1.8. In Section 5 we show how one can deduce Theorem 1.5 from Theorem 1.8, along with a p-biased version which we will introduce briefly now.

For a family \mathcal{F} of graphs on *n* labeled vertices, let $\mu_p(\mathcal{F}) = \sum_{G \in \mathcal{F}} p^{e(G)} (1-p)^{e(\overline{G})}$. Ellis, Filmus, and Friedgut [4] showed that Theorem 1.3 also holds for $p \leq 1/2$ when $|\mathcal{F}|$ is replaced by $\mu_p(\mathcal{F})$ with an upper bound of $\mu_p(\mathcal{F}) \leq p^3$, and the conclusion of (stability) modified accordingly. It was conjectured that these results extend to 1/2 , but this was disproved by Keller andLifschitz [9]. Whereas the *p*-biased proof was somewhat involved in [4], recent work of Ellis, Keller, and Lifshitz [5] enables one to take any of a wide variety of p = 1/2 intersecting family results as a black box and deduce p < 1/2 versions, including stability. This enables us to obtain the following corollary of Theorem 1.8 (see Section 5).

Corollary 1.9. Let $0 and let <math>t \in \{3, 4\}$. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-intersecting pair of families graphs on n labeled vertices.

- (Upper bound) μ_p(F₁)μ_p(F₂) ≤ p<sup>-2(^t₂)
 (Uniqueness) The upper bound is an equality if and only if F₁ = F₂ is a K_t-unvirate.
 </sup>
- (Stability) There exists a constant $C_{p,t} > 0$ depending only on p and t so that for all $\epsilon > 0$, if $|\mathcal{F}_1| |\mathcal{F}_2| \ge (1-\epsilon)p^{-2\binom{t}{2}}$, then there exists a K_t -umvirate \mathcal{T} such that $\mu_p(\mathcal{F}_i \setminus \mathcal{T}) \le C_{p,t} \epsilon^{\log_{(1-p)}(p)}$ for $i \in \{1, 2\}$.

Let us mention some additional related results. As noted at the end of of [4], the natural generalization of Conjecture 1.2 is false a general H, as an H-intersecting family could be a constant factor larger than an H-umvirate. Indeed, Noga Alon observed that for every fixed star forest H, the largest *H*-intersecting family on *n* vertices has size $(1 - o(1))2^{\binom{n}{2}-1}$ (divide the *n* vertices into two equal halves $A \cup B$ and take all graphs with at least n/4 + C vertices in A having degree $\geq n/4 + C$ to B; this family is H-intersecting for a large enough C, and it has size $(1 - o(1))2^{\binom{n}{2}-1}$. On the other hand, by pairing graphs with their complements, the size of an H-intersecting family is at most $2^{\binom{n}{2}-1}$). He further conjectured that there is some universal constant c > 0 such that for H that is not a star forest, the largest *H*-intersecting family on *n* vertices has size at most $(1-c)2^{\binom{n}{2}-1}$. It would suffice to prove this conjecture for the 3-edge-path P_3 . It had been conjectured [2] that the largest P_3 -intersecting family is a P_3 -umvirate, but this is false, as Christofides constructed a P_3 -intersecting family of graphs on 6 vertices with size $17 \cdot 2^8 > 2^{\binom{6}{2}-3}$. See [1, 11] for additional related work.

Organization. In Section 2 we explain the framework introduced in [4] to reduce the main result Theorem 1.8 to a certain linear program. The reduction here is essentially the same as in [4], though we need to state everything for cross-intersecting families instead of a single intersecting family as was done in [4]. In Section 3, we reduce the verification of dual constraints to a finite computation—this is where our arguments start to differ from [4]. In Section 4, we construct feasible dual solutions and verify their validity. In Section 5, we explain how to deduce Theorem 1.5 and Corollary 1.9 from Theorem 1.8. In Section 6, we conclude with some brief remarks on the different possible directions for generalization.

Acknowledgements. The first author thanks Pat Devlin for introducing him to this problem, and Jonathan Tidor for some helpful discussions. We thank Noam Lifshitz for pointing out to us that one can deduce Corollary 1.9 using [4].

2. Reduction to a linear program

In this section, we explain the solution framework introduced in [4] that reduces the problem to a certain linear program. This reduction is analogous to Delsarte's linear programming bound for error correcting codes [3].

2.1. Fourier analytic bound. Every graph on n labeled vertices is identified with its edge set indicator vector, viewed as an element of \mathbb{F}_2^N with $N = \binom{n}{2}$. Families of graphs correspond to indicator functions $f : \mathbb{F}_2^N \to \{0, 1\}$. We use the following standard conventions for Fourier analysis on \mathbb{F}_2^N , normalized with the uniform measure on the physical space and the counting measure on frequency space. For a function $f : \mathbb{F}_2^N \to \mathbb{R}$, its Fourier transform is given by

$$\widehat{f}(\lambda) = \mathop{\mathbb{E}}_{x \in \mathbb{F}_2^N} f(x) \cdot (-1)^{\lambda \cdot x}, \qquad \lambda \in \mathbb{F}_2^N$$

whereas the Fourier inversion formula is given by

$$f(x) = \sum_{\lambda \in \mathbb{F}_2^N} \widehat{f}(\lambda) \cdot (-1)^{\lambda \cdot x}, \qquad x \in \mathbb{F}_2^N.$$

Given two function $f, g: \mathbb{F}_2^N \to \mathbb{R}$, we denote their convolution by $f * g: \mathbb{F}_2^N \to \mathbb{R}$ by $(f * g)(x) = \mathbb{E}_y f(y)g(x-y)$. We have the identity $\widehat{f * g}(\lambda) = \widehat{f}(\lambda)\widehat{g}(\lambda)$.

The following claim can be viewed as a weighted version of Hoffman's bound on the independence number of a regular graph in terms of its eigenvalues. It is also known as a linear programming bound.

Proposition 2.1. Let $f, g : \mathbb{F}_2^N \to \{0, 1\}$ and let $\nu : \mathbb{F}_2^N \to \mathbb{R}$ satisfy $\mathbb{E}[\nu] = 1$ and $\langle f * g, \nu \rangle = 0$. Then

• (Upper bound)We have

$$\mathbb{E}[f] \mathbb{E}[g] \le \left(\frac{m}{1+m}\right)^2,\tag{1}$$

where $m = \max_{\lambda \neq 0} |\hat{\nu}(\lambda)|$ is the largest magnitude among nontrivial eigenvalues of ν .

- (Maximal families) If equality holds, then $\mathbb{E}[f] = \mathbb{E}[g]$, and $\widehat{f}(\lambda) = \widehat{g}(\lambda) = 0$ for all $\lambda \neq 0$ with $|\widehat{\nu}(\lambda)| < m$.
- (Stability) For all $\epsilon, \delta \in (0, 1]$, if $\mathbb{E}[f] \mathbb{E}[g] \ge (m/(1+m))^2 \epsilon$, then

$$\sum_{\substack{\lambda \neq 0 \\ |\hat{\nu}(\lambda)| \le (1-\delta)m}} \widehat{f}(\lambda)^2 \le \frac{(1+m)^4}{\delta(2-\delta)m^2} \epsilon.$$

Proof (Upper bound). Taking a Fourier transform, $0 = \langle f * g, \nu \rangle = \sum_{\lambda \in \mathbb{F}_2^N} \widehat{f}(\lambda) \widehat{g}(\lambda) \widehat{\nu}(\lambda)$. Since this sum equals 0, the contribution from $\lambda = 0$ must cancel with the rest of the summation:

$$\mathbb{E}[f] \mathbb{E}[g] = -\sum_{\lambda \neq 0} \widehat{\nu}(\lambda) \widehat{f}(\lambda) \widehat{g}(\lambda)$$

$$\leq \max_{\lambda \neq 0} \left[|\widehat{\nu}(\lambda)| \right] \sum_{\lambda \neq 0} |\widehat{f}(\lambda)| |\widehat{g}(\lambda)| \qquad (*)$$

$$\leq m \left(\sum_{\lambda \neq 0} \widehat{f}(\lambda)^2 \right)^{\frac{1}{2}} \left(\sum_{\lambda \neq 0} \widehat{g}(\lambda)^2 \right)^{\frac{1}{2}} \qquad [Cauchy-Schwarz]$$

$$= m(\mathbb{E}[f] - \mathbb{E}[f]^2)^{\frac{1}{2}} (\mathbb{E}[g] - \mathbb{E}[g]^2)^{\frac{1}{2}}$$
[Plancherel]
$$= m(\mathbb{E}[f] \mathbb{E}[g] - \mathbb{E}[f]^2 \mathbb{E}[g] - \mathbb{E}[f] \mathbb{E}[g]^2 + \mathbb{E}[f]^2 \mathbb{E}[g]^2)^{\frac{1}{2}}$$

$$\leq m(\mathbb{E}[f] \mathbb{E}[g] - 2 \mathbb{E}[f]^{3/2} \mathbb{E}[g]^{3/2} + \mathbb{E}[f]^2 \mathbb{E}[g]^2)^{\frac{1}{2}}$$
[AM-GM]
$$= m((\mathbb{E}[f] \mathbb{E}[g])^{1/2} - \mathbb{E}[f] \mathbb{E}[g]).$$

This is now a quadratic in $(\mathbb{E}[f]\mathbb{E}[g])^{1/2}$; solving for this quantity in terms of m gives the desired result.

Proof (Maximal families). For equality to hold, each inequality above must be tight. For the AM-GM step to be an equality, we must have $\mathbb{E}[f] = \mathbb{E}[g]$, assuming neither is identically zero. For the Cauchy–Schwarz step to be an equality, we must have $|\hat{f}(\lambda)| = \alpha |\hat{g}(\lambda)|$ for some α and all $\lambda \neq 0$. Since $\hat{f}(0) = \mathbb{E}[f] = \mathbb{E}[g] = \hat{g}(0)$, we deduce $\alpha = 1$. In particular, \hat{f} and \hat{g} have the same support, and so if (*) is also an equality, we must have $|\hat{\nu}(\lambda)| = m$ whenever $|\hat{f}(\lambda)|, |\hat{g}(\lambda)| > 0$, which completes the proof.

Proof (Stability). Let $D := \sum_{\lambda \neq 0, |\hat{\nu}(\lambda)| < m} \hat{f}(\lambda)^2$. In order to get a tighter bound on $\mathbb{E}[f]\mathbb{E}[g]$, we replace (*) with a tighter bound. Define

$$\widetilde{f}(\lambda) := \begin{cases} \widehat{f}(\lambda) & \lambda = 0 \text{ or } |\widehat{\nu}(\lambda)| > (1-\delta)m \\ \widehat{f}(\lambda) \cdot (1-\delta) & \text{otherwise.} \end{cases}$$

With this definition, we have

$$\begin{split} \mathbb{E}[f] \,\mathbb{E}[g] &= -\sum_{\lambda \neq 0} \widehat{\nu}(\lambda) \widehat{f}(\lambda) \widehat{g}(\lambda) \\ &\leq \sum_{\substack{\lambda \neq 0 \\ |\widehat{\nu}(\lambda)| > (1-\delta)m}} m |\widehat{f}(\lambda) \widehat{g}(\lambda)| + \sum_{\substack{\lambda \neq 0 \\ |\widehat{\nu}(\lambda)| \leq (1-\delta)m}} (1-\delta)m |\widehat{f}(\lambda) \widehat{g}(\lambda)| \\ &= m \sum_{\lambda \neq 0} |\widetilde{f}(\lambda)| |\widehat{g}(\lambda)| \\ &\leq m \Big(\sum_{\lambda \neq 0} \widetilde{f}(\lambda)^2 \Big)^{\frac{1}{2}} \Big(\sum_{\lambda \neq 0} \widehat{g}(\lambda)^2 \Big)^{\frac{1}{2}} \end{split}$$
 [Cauchy–Schwarz]

Let us pause for a moment to consider this $\sum_{\lambda \neq 0} \widetilde{f}(\lambda)^2$ term. We can rewrite this as

$$\sum_{\lambda \neq 0} \widetilde{f}(\lambda)^2 = \sum_{\substack{\lambda \neq 0 \\ |\widehat{\nu}(\lambda)| > (1-\delta)m}} \widehat{f}(\lambda)^2 + (1-\delta)^2 \sum_{\substack{\lambda \neq 0 \\ |\widehat{\nu}(\lambda)| \le (1-\delta)m}} \widehat{f}(\lambda)^2 = \sum_{\lambda \neq 0} \widehat{f}(\lambda)^2 - D\delta(2-\delta),$$

BERGER AND ZHAO

where we have plugged in the definition of D from the beginning of the proof. Since $\sum \hat{f}(\lambda)^2 \leq 1$, the RHS is at most $(1 - D\delta(2 - \delta)) \sum_{\lambda \neq 0} \hat{f}(\lambda)^2$. This multiplicative factor can now be pulled out and combined with the m, and the rest of the proof proceeds as before. To simplify notation, let $c^2 = 1 - D\delta(2 - \delta)$. Expanding out the final result, we obtain the upper bound

$$\mathbb{E}[f] \mathbb{E}[g] \le \left(\frac{cm}{1+cm}\right)^2 = \frac{m^2}{(1+m)^2} - \frac{m^2(1+cm)^2 - (1+m)^2(cm)^2}{(1+m)^2(1+cm)^2}$$
$$= \frac{m^2}{(1+m)^2} - \frac{m^2(1-c^2) + 2m^3(c-c^2)}{(1+m)^2(1+cm)^2}$$
$$\le \frac{m^2}{(1+m)^2} - \frac{m^2(1-c^2)}{(1+m)^4}$$
$$= \left(\frac{m}{1+m}\right)^2 - \frac{D\delta(2-\delta)m^2}{(1+m)^4}.$$

Thus $D\delta(2-\delta)m^2/(1+m)^4 \leq \epsilon$. Solving for D completes the proof.

2.2. A class of dual functions. In order to apply Proposition 2.1 to the indicator functions f, g of a cross-*t*-chromatic-agreeing pair of families, we need a suitable ν .

Lemma 2.2. Let f, g be indicator functions of a cross-t-chromatic-agreeing pair of families of graphs on n labeled vertices. Let $N = \binom{n}{2}$ and $\nu : \mathbb{F}_2^N \to \mathbb{R}$ be supported on graphs whose complements are (t-1)-colorable. Then $\langle f * g, \nu \rangle = 0$.

Proof. Since f, g are cross-t-chromatic-agreeing, for any x with f(x) = 1 and y with g(y) = 1, x + y is the edge-indicator vector of a graph whose complement is not (t - 1)-colorable. Hence, $\nu(x + y) = 0$. It follows that $\langle f * g, \nu \rangle = \mathbb{E}_{x,y}[f(x)g(y)\nu(x + y)] = 0$.

In view of Proposition 2.1, it therefore suffices to construct some function $\nu : \mathbb{F}_2^N \to \mathbb{R}$ that is supported on graphs whose complements are (t-1)-colorable, with $\mathbb{E}[\nu] = 1$ and $m = \max_{\lambda \neq 0} \left[|\hat{\nu}(\lambda)| \right] \le 1/(2^{\binom{t}{2}} - 1)$. For stability purposes we would also like to ensure that $|\hat{\nu}(\lambda)| < m - \delta$ for some fixed $\delta > 0$ whenever λ has more than $\binom{t}{2}$ edges.

This condition on the support of ν is complicated to state, but easy to work with, as it is linear. The following useful class of functions (and therefore, their linear combinations) satisfy this condition.

Lemma 2.3. Let $N = \binom{n}{2}$, let \mathcal{T} be an arbitrary distribution on (t-1)-colorable graphs, and let $\{f_G\}$ be a set of functions indexed by graphs on n vertices. If $\nu : \mathbb{F}_2^N \to \mathbb{R}$ satisfies

$$\widehat{\nu}(\lambda) = (-1)^{e(\lambda)} \mathop{\mathbb{E}}_{T \sim \mathcal{T}} f_T(T \cap \lambda),$$

then ν is supported on graphs whose complements are (t-1)-colorable.

Remark. This is a slight generalization of [4, Lemma 2.8], but the proof is identical.² We include a concise proof here for completeness.

Proof. The set of functions $\hat{\nu}$ such that ν is supported on graphs whose complements are (t-1)colorable forms a subspace of the space of functions $\mathbb{F}_2^N \to \mathbb{R}$. For (t-1)-colorable T, let ν_T be a
point mass at its complement \overline{T} . Then $\hat{\nu}_T(\lambda) = (-1)^{e(\lambda)}(-1)^{e(\lambda\cap T)}$ lies in this subspace. The set
of functions $\{\hat{\nu}_{T'}\}_{T'\subseteq T}$ forms a basis for the space of all functions of the form $(-1)^{e(\lambda)}f_T(T\cap\lambda)$,
where $f_T: \mathbb{F}_2^N \to \mathbb{R}$ is arbitrary. The lemma follows by linearity.

²To obtain this version, replace the words "OCC Spectrum" with "the Fourier transform of some ν supported on graphs whose complements are (t-1)-colorable," and instead of a distribution on bipartite graphs take a distribution on (t-1)-colorable graphs.

Although this gives us quite a large class of functions in which to search for an optimal $\hat{\nu}$, in practice we will only need the uniform distribution on complete (t-1)-partite graphs. In this case, T may be viewed as a uniform random 3-coloring of [n] and $T \cap \lambda$ is a random subgraph of λ given by choosing a uniform random 3-coloring of its vertices and deleting all monochromatic edges.

Definition 2.4. For a graph G on n labeled vertices, let $[q]^{V(G)}$ be the set of maps $\varphi \colon V(G) \to [q]$, viewed as q-colorings of V(G) (not necessarily proper). For each coloring $\varphi \colon V(G) \to [q]$, define G_{φ} to be the subgraph of G formed by deleting all monochromatic edges of G, and then deleting all isolated vertices from the result. Let G_q be the random graph G_{φ} given by choosing $\varphi \sim \text{Unif}([q]^{V(G)}).$

Proposition 2.5. Let $t \in \{3, 4\}$. There exists a set of unlabeled graphs $\{H\}$, coefficients $\{c_H\}$ and $\delta > 0$ so that for any G on n labeled vertices, we have that

$$\mu(G) := (-1)^{e(G)} \sum_{H} c_H \mathbb{P}[G_q \cong H]$$

$$\tag{2}$$

satisfies the following conditions.

- (a) $\mu(0) = 2^{\binom{t}{2}} 1.$ (b) $|\mu(G)| \leq 1$ for all $G \neq \emptyset$
- (c) $|\mu(G)| \leq 1 \delta$ whenever G has more than $\binom{t}{2}$ edges.

The constructions are presented in Section 4, along with verification of Proposition 2.5. Assuming the proposition, we are nearly in a position to prove the main result, Theorem 1.8. We first include two results from Boolean analysis.

Lemma 2.6 ([7, Lemma 2.8]). Let $f : \mathbb{F}_2^N \to \{0, 1\}$ be a monotone Boolean function with $\mathbb{E}[f] = 2^{-k}$ and $\widehat{f}(S) = 0$ whenever |S| > k. Then f is a k-unvirate.³

Theorem 2.7 ([10, Theorem 3]). For every k, there exists C > 0 and K such that for every $f: \mathbb{F}_2^N \to \{0,1\}$ there exists $g: \mathbb{F}_2^N \to \{0,1\}$ that depends on at most K coordinates and satisfies

$$\mathbb{P}_{x \in \{0,1\}^N}[f(x) \neq g(x)] \le C \sum_{|S| > k} \widehat{f}(S)^2.$$

To deduce uniqueness we will need to apply Lemma 2.6, for which we require monotonicity. To that end, we employ a shifting argument. For a family of graphs \mathcal{F} on n labeled vertices and $e \in {\binom{[n]}{2}}$, the compression of \mathcal{F} in direction e, denoted $C_e(\mathcal{F})$, is given by replacing each $G \in \mathcal{F}$ with $G \cup \{e\}$ whenever $e \notin E(G)$ and $G \cup \{e\}$ is not already in \mathcal{F} . It is easy to see from this definition that $|C_e(\mathcal{F})| = |\mathcal{F}|$.

Lemma 2.8. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-agreeing pair of families of graphs on n labeled vertices. Let $C = C_{e_1} \circ \cdots \circ C_{e_k}$ for some $k \in \mathbb{Z}$, $e_1, \ldots e_k \in \binom{[n]}{2}$. Then $(C(\mathcal{F}_1), C(\mathcal{F}_2))$ is also cross-t-chromatic-agreeing. Moreover, if $C(\mathcal{F}_1) = C(\mathcal{F}_2)$ is a K_t -conjunction, then $\mathcal{F}_1 = \mathcal{F}_2$ is a K_t -conjunction as well.

Proof. By induction, it suffices to prove both claims in the case $C = C_e$. To check the first claim, note that if $G'_1 \in C_e(\mathcal{F}_1)$ and $G'_2 \in C_e(\mathcal{F}_2)$, then there exist $G_1 \in \mathcal{F}_1$ and $G_2 \in \mathcal{F}_2$ such that G'_1 and G'_2 agree wherever G_1 and G_2 agree (and possibly elsewhere as well).

The second claim is nearly identical to [4, Lemma 2.7] (which was stated for an intersecting family rather than cross-intersecting families). Fix some copy T of K_t . Assume $C_e(\mathcal{F}_1) = C_e(\mathcal{F}_2)$ is

³Here f being an k-unvirate means that there is some fixed set of k coordinates so that f is 1 if and only if the input is 1 on each of these k coordinates. [7, Lemma 2.8] is stated for a biased distribution with p < 1/2, but with brief consideration it is easy to see the proof works for the unbiased case p = 1/2 as well, which is what we have stated here.

BERGER AND ZHAO

a *T*-conjunction. If *e* is outside *T*, C_e^{-1} must act trivially, and so $\mathcal{F}_1 = \mathcal{F}_2$ is the same conjunction. Otherwise, $e \in T$, and each choice of edges $H \subseteq {\binom{[n]}{2}} \setminus E(T)$ extends uniquely to a graph in \mathcal{F}_i , for each *i*. View the set of such *H* as a hypercube graph, with $H \sim H'$ if they differ in exactly one edge. Now consider the following two colorings of this hypercube. For *H* in the hypercube, let $\chi_1(H)$ be given by membership of *e* in the extension of *H* to a member of \mathcal{F}_1 , and let $\chi_2(H)$ be given by membership of *e* in the extension of $K_n \setminus E(T \cup H)$ to a member of \mathcal{F}_2 . If these two colorings χ_1 and χ_2 are constant and identical, then \mathcal{F}_1 and \mathcal{F}_2 are identical *T*-conjunctions. Otherwise, we can find a pair (H, H') that disagree on at most one edge and satisfy $\chi_1(H) \neq \chi_2(H')$. Then the pair $(H, K_n \setminus E(T \cup H'))$ extends to $(G, G') \in \mathcal{F}_1 \times \mathcal{F}_2$, whose agreement is contained in $K_t \setminus \{e\} \cup \{\leq 1 \text{ other edge}\}$. This set of edges in agreement cannot possibly have chromatic number at least *t*, which completes the contradiction.

Corollary 2.9. It suffices to prove (uniqueness) of Theorem 1.8 in the case when \mathcal{F}_1 and \mathcal{F}_2 are monotone.

Proof. Whenever C_e acts nontrivially on \mathcal{F}_i , it increases the total number of edges among graphs in \mathcal{F}_i by at least 1. This quantity is uniformly upper bounded, and so after some finite number of applications $C = C_{e_1} \circ \cdots \circ C_{e_k}$ one has that $C(\mathcal{F}_1)$ and $C(\mathcal{F}_2)$ are invariant under all C_e , and hence monotone. By the first half of Lemma 2.8, $C(\mathcal{F}_1)$ and $C(\mathcal{F}_2)$ are still K_t -chromatic-agreeing, so the monotone case of (uniqueness) tells us that they equal some K_t -conjunction. By the second half of 2.8, we conclude that \mathcal{F}_1 and \mathcal{F}_2 are identical K_t -conjunctions as well.

2.3. Proof of Theorem 1.8 assuming Proposition 2.5.

Proof (Upper bound). Let $(\mathcal{F}_1, \mathcal{F}_2)$ be t-chromatic agreeing. Let $\nu : \mathbb{F}_2^N \to \mathbb{R}$ satisfy $\hat{\nu} = \mu/(2^{\binom{t}{2}}-1)$, where μ is the function given by Proposition 2.5. Then $\mathbb{E}[\nu] = \hat{\nu}(0) = 1$, and $m = \max_{\lambda \neq 0} |\hat{\nu}(\lambda)| = 1/(2^{\binom{t}{2}}-1)$. By Lemmas 2.2 and 2.3, we see $\langle f * g, \nu \rangle = 0$. Applying Proposition 2.1, we obtain $\mathbb{E}[f_1] \mathbb{E}[f_2] \leq 2^{-2\binom{t}{2}}$.

Proof (Uniqueness). Taking $\mathcal{F}_1 = \mathcal{F}_2$ to be a K_t -conjunction achieves the upper bound. We need to show no other families may achieve the upper bound; by Corollary 2.9 it suffices to consider pairs $(\mathcal{F}_1, \mathcal{F}_2)$ that are monotone. By (maximal families) from Proposition 2.1 their indicator functions \hat{f}_1, \hat{f}_2 are supported on λ where $|\hat{\nu}(\lambda)|$ is maximized, and they satisfy $\mathbb{E}[f_1] = \mathbb{E}[f_2]$. Thus $\mathbb{E}[f_1]^2 = \mathbb{E}[f_2]^2 = \mathbb{E}[f_1] \mathbb{E}[f_2] = 2^{-2\binom{t}{2}}$. Since \hat{f}_1, \hat{f}_2 are supported only on λ with $|\lambda| \leq \binom{t}{2}$, we are in a position to apply Lemma 2.6. This tells us that f_1 depends only on some set of coordinates T_1 and f_2 depends only on some set of coordinates T_2 , where $|T_1| = |T_2| = \binom{t}{2}$. It is easy to choose $G_1 \in \mathcal{F}_1$ and $G_2 \in \mathcal{F}_2$ so that G_1 and G_2 disagree on all edges outside $T_1 \cap T_2$, so $T_1 \cap T_2$ is a graph on $\binom{t}{2}$ edges with chromatic number at least t. This can only happen when $T_1 = T_2 = T \cong K_t$, and now it is easy to see that to be cross-t-chromatic-agreeing \mathcal{F}_1 and \mathcal{F}_2 must be identical T-conjunctions. \Box

Proof (Stability). Applying (stability) from Proposition 2.1, we see $\sum_{|S| > \binom{t}{2}} \widehat{f_i}(S)^2 \leq C\epsilon$ for some constant C depending only on t and δ . Applying Theorem 2.7, we see that each f_i agrees with some g_i on all but a $C'\epsilon$ -fraction of inputs, where g_i is a function only of some set T_i of coordinates with $|T_i| \leq K_0$. Let V be the set of vertices incident to an edge of $E(T_1)$ or $E(T_2)$, so $|V| \leq 4K_0$. Then g_1 and g_2 may be viewed as functions $\overline{g_1}, \overline{g_2}$ on the set of graphs on V, and in particular this restriction does not change expectation. Assuming for the sake of contradiction that $\overline{g_1}, \overline{g_2}$ are not identical K_t -conjunctions, we see by (uniqueness) that $\mathbb{E}[g_1]\mathbb{E}[g_2] < 2^{-2\binom{t}{2}}$. Since there are a finite number of non-maximal families on $4K_0$ vertices, we may write

$$\mathbb{E}[g_1]\mathbb{E}[g_2] \le \max_{\substack{h_1,h_2 \text{ nonmaximal} \\ \text{on } \le 4K_0 \text{ vertices}}} \left\lfloor \mathbb{E}[h_1]\mathbb{E}[h_2] \right\rfloor =: 2^{-2\binom{l}{2}} - c, \text{ for some } c > 0.$$

If ϵ is sufficiently small that $2C'\epsilon + \epsilon < c$, then we have

$$\mathbb{E}[f_1]\mathbb{E}[f_2] \le \mathbb{E}[g_1]\mathbb{E}[g_2] + 2C'\epsilon \le 2^{-2\binom{t}{2}} - c + 2C'\epsilon < 2^{-2\binom{t}{2}} - \epsilon$$

which is a contradiction. Thus $\overline{g_1} = \overline{g_2}$ is the indicator function of some K_t -conjunction, and by extension so are f_1 and f_2 .

3. How to verify the dual linear constraints

In order to prove Proposition 2.5, once we have specified choices for c_H we need to be able to effectively bound $\mu(G) = (-1)^{e(G)} \sum c_H \mathbb{P}[G_q \cong H]$ for all graphs G. (Recall from Definition 2.4 that G_q is a random subgraph of G given by taking a random q-coloring of V(G) and deleting all monochromatic edges, and then all isolated vertices.) The nature of μ makes this easy to do for sufficiently large G: by choosing c_H to be supported on graphs with a bounded number of edges, it not too hard see that $\mu(G)$ must decay with e(G). The proof is then a balance between optimizing control on the decay of $\mu(G)$ and doing casework on small graphs for which the generic bounds are not strong enough. The casework can thankfully be offloaded to a computer but computation time quickly becomes a limiting factor, as the number of graphs on n vertices increases quickly with n. Our main tool to reduce to a finite computation is Proposition 3.5. We will begin with the derivation, and conclude with the statement of the bound. The derivation proceeds first by obtaining a result for connected graphs, and then bootstrapping that to all graphs.

The main idea of the proof is that if the coefficients c_H are supported only on graphs of a bounded size, we can bound $\mu(G)$ by a contribution from each of its induced subgraphs on a bounded number of vertices. For a graph G on n labeled vertices, let $\kappa(G)$ be the number of connected components of G. We begin with a lemma.

Lemma 3.1. Let q be a positive integer, and let $G'' \subseteq G' \subseteq G$ be labeled graphs, where the containments $V(G'') \subseteq V(G) \subseteq V(G)$ may be strict. Furthermore, assume G is connected. Then

$$\#\left\{\varphi'\in[q]^{V(G')}:G'_{\varphi'}=G''\right\}\geq q^{\kappa(G')-1}\cdot\#\left\{\varphi\in[q]^{V(G)}:G_{\varphi'}=G''\right\}.$$

This lemma presents the somewhat counterintuitive fact that when the ambient graph G is connected, a subgraph G' may have more colorings which yield G'' than G itself does.

Proof. Let $\Phi_G := \{\varphi \in [q]^{V(G)} : G'_{\varphi'} = G''\}$, and $\Phi_{G'} := \{\varphi' \in [q]^{V(G')} : G'_{\varphi'} = G''\}$. Restriction of colorings from V(G) to V(G') gives a map $\rho : \Phi_G \to \Phi_{G'}$. First, we claim ρ is injective. Let $\varphi \in \Phi_G$, so $G_{\varphi} = G'' \subseteq G'$. Then by definition, all edges in $E(G) \setminus E(G')$ are monochromatic in φ . Since G is connected, this means φ is uniquely determined by its colors on V(G'), which proves the claim.

In fact, a more careful application of the same argument shows that the image of Φ_G under restriction ρ is in fact much smaller than $\Phi_{G'}$. Let $\varphi' \in \Phi_{G'}$. Take some connected component of G'and permute its colors via a cyclic shift, so $\varphi'(v) \mapsto \varphi'(v) + 1 \pmod{q}$ for v in this component. After this operation, we still have $G'_{\varphi'} = G''$. This motivates the following definition. For $\varphi', \varphi'' \in \Phi_{G'}$, say $\varphi' \sim \varphi''$ if one may obtain φ'' by starting from the coloring φ' and applying a sequence of cyclic shifts to the colors of connected components in G'. This is a well-defined equivalence relation, and each equivalence class has size $q^{\kappa(G')}$. We claim that the image $\rho(\Phi_G)$ contains at most q elements of this equivalence class. As before, for any coloring $\varphi \in \Phi_G$, the edges in $E(G) \setminus E(G')$ must be monochromatic. Since G is connected, this means that φ is uniquely determined once we specify the equivalence class of its restriction to V(G') and the color of a single vertex v_0 . Since there are q choices for the color of v_0 , the claim follows.

Concluding, $q|\Phi_{G'}| \leq q^{\kappa(G')}|\Phi_G|$, which completes the proof.

For a graph G on n labeled vertices, let $\binom{G}{n_0}$ be the set of labeled subgraphs of G induced by all choices of n_0 vertices from V(G).

Lemma 3.2. Fix q > 0, a list $\{H\}$ of unlabeled graphs on at most n_0 vertices and $\{c_H\}$ a list of coefficients. Then for any connected G on $n \ge n_0$ labeled vertices,

$$\sum_{H} |c_{H}| \cdot \mathbb{P}[G_{q} \cong H] \leq \sum_{G' \in \binom{G}{n_{0}}} \frac{1}{q^{\kappa(G')-1}} \sum_{H} \frac{|c_{H}|}{q^{n-n_{0}} \binom{n-v(H)}{n_{0}-v(H)}} \mathbb{P}[G'_{q} \cong H].$$
(3)

Although complicated, this bound is appealing for three reasons. First, when $\mu(G) = \sum_{H} c_{H} \cdot \mathbb{P}[G_{\varphi} \cong H]$, the LHS is an upper bound for $|\mu(G)|$ by the triangle inequality. Second, plugging in even the very crude bound $\sup |c_{H}|/q^{n-n_{0}}$ for the inner summation, this gives us the bound $|\mu(G)| \leq \sup |c_{H}| \cdot n^{n_{0}}/q^{n-n_{0}}$ which decays exponentially in n. Third, each summand of the outer summation is a function only of n and some graph G' on n_{0} vertices. Thus given n, we can obtain a bound on this quantity through direct computation, iterating through all colorings of all graphs on n_{0} vertices. In fact, this bound will decay for n sufficiently large, which will allow us to remove the n dependence entirely and obtain a good uniform bound on μ for all graphs on $n > n_{0}$ vertices.

Proof of Lemma 3.2. We employ a double-counting argument, summing Lemma 3.1 over all induced subgraphs $G' \subseteq G$ on n_0 vertices. For fixed H with $c_H \neq 0$, we have

$$\begin{aligned} |c_{H}| \cdot \mathbb{P}[G_{q} \cong H] &= \sum_{G \supseteq G'' \cong H} \frac{|c_{H}|}{q^{n}} \cdot \# \Big\{ \varphi \in [q]^{V(G)} : G_{\varphi} = G'' \Big\} \\ &= \sum_{G' \in \binom{G}{n_{0}}} \sum_{G' \supseteq G'' \cong H} \frac{|c_{H}|}{q^{n}} \cdot \frac{\# \{ \varphi \in [q]^{V(G)} : G_{\varphi} = G'' \}}{\# \{ G' \in \binom{G}{n_{0}} : G'' \subseteq G' \}} \\ &\leq \sum_{G' \in \binom{G}{n_{0}}} \sum_{G' \supseteq G'' \cong H} \frac{|c_{H}|}{q^{n}} \cdot \frac{q^{1-\kappa(G')} \cdot \# \{ \varphi' \in [q]^{V(G')} : G'_{\varphi'} = G'' \}}{\binom{n-\nu(G'')}{n_{0}-\nu(G'')}} \\ &= \sum_{G' \in \binom{G}{n_{0}}} \frac{1}{q^{\kappa(G')-1}} \frac{|c_{H}|}{q^{n-n_{0}} \binom{n-\nu(H)}{n_{0}-\nu(H)}} \mathbb{P}[G'_{q} \cong H] \end{aligned}$$

Sum this inequality over all H and we obtain the lemma.

Now we eliminate the dependence of Lemma 3.2 on n. Take (3) and upper bound the contents of the outer summation by taking its maximum over $G' \subseteq K_{n_0}$. The summation turns into a multiplicative factor of $\binom{n}{n_0}$, and so we may upper bound the RHS by

$$\max_{G' \subseteq K_{n_0}} \left[\frac{1}{q^{k(G')-1}} \sum_H |c_H| \cdot \mathbb{P}[G'_q \cong H] \cdot \frac{\binom{n}{n_0}}{q^{n-n_0}\binom{n-v(H)}{n_0-v(H)}} \right].$$
(4)

The final term is complicated, but decays with n, and can be replaced by a uniform upper bound. Let

$$DC_{q,n_0}(x) := \max_{\substack{n \in \mathbb{Z} \\ n > n_0}} \frac{\binom{n}{n_0}}{q^{n-n_0} \binom{n-x}{n_0-x}}.$$
(5)

The letters DC stand for "double counting." This expression changes by a multiplicative factor of (n+1)/(q(n-x)) when n increases to n+1. So for $q \ge 2$ it is already decreasing for $n \ge 2x+1$, and it suffices to maximize for n up to this bound. Below is a table of the values of DC which covers our two constructions. For all of these values, the maximum in (5) is achieved when n = 10, except for $DC_{3,9}(8)$ which is maximized at n = 11.

x	0	1	2	3	4	5	6	7	8
$DC_{2,9}(x)$	1/2	_	5/8	5/7	5/6	1	_	_	_
$DC_{3,9}(x)$	1/3	_	5/12	10/21	5/9	2/3	5/6	10/9	55/27

Plugging DC into (4), we obtain a uniform bound on $|\mu(G)|$ for all connected G with more than n_0 vertices.

Lemma 3.3. Fix q > 0, a list $\{H\}$ of unlabeled graphs on at most n_0 vertices and $\{c_H\}$ a list of coefficients. Then for any connected G on $n > n_0$ labeled vertices, we have

$$\sum_{H} |c_H| \cdot \mathbb{P}[G_q \cong H] \le \max_{G' \subseteq K_{n_0}} \left\lfloor \frac{1}{q^{k(G')-1}} \sum_{H} |c_H| \cdot \mathbb{P}[G'_q \cong H] \cdot DC_{q,n_0}(v(H)) \right\rfloor.$$
(6)

Now we extend this to a bound for all G. Let G be a graph with at least two connected components, and let $u, v \in V(G)$ be a pair vertices in distinct connected components. Our first guess would be to replace G by the graph obtained by identifying u and v, which we denote $G/(u \sim v)$, and hope that $\sum_{H} |c_{H}| \cdot \mathbb{P}[G_{\varphi} \cong H]$ increases. This is nearly the case.

Lemma 3.4. Fix q > 0, a list $\{H\}$ of unlabeled graphs on at most n_0 vertices and $\{c_H\}$ a list of coefficients. Then for any G on n labeled vertices with at least two connected components, u and v a pair of disconnected vertices in G (i.e., there is no path in G from u to v), and $G' = G/(u \sim v)$, we have

$$\sum_{H} |c_{H}| \cdot \mathbb{P}[G_{q} \cong H] \leq \sum_{H} c'_{H} \cdot \mathbb{P}[G'_{q} \cong H],$$

where c'_H is the maximum value of $|c_{H'}|$ over all graphs H' which may be transformed to H by contracting at most one pair of disconnected vertices.

Proof. Let φ be a random q-coloring of V(G) distributed uniformly on colorings satisfying $\varphi(u) = \varphi(v)$. Since u and v are in distinct connected components, it is easy to see that G_q and G_{φ} have the same distribution. Any fixed φ_0 satisfying $\varphi_0(u) = \varphi_0(v)$ descends to a well-defined coloring φ_1 on $G' = G/(u \sim v)$. Furthermore, we may get from G_{φ_0} to G'_{φ_1} by identifying at most one pair of disconnected vertices. Consequently, the following inequality holds term by term, equating terms under this map $\varphi_0 \mapsto \varphi_1$.

$$\sum_{H} |c_{H}| \cdot \mathbb{P}[G_{q} \cong H] = \frac{1}{q^{n-1}} \sum_{\substack{\varphi_{0} \in [q]^{V(G)} \\ \varphi_{0}(u) = \varphi_{0}(v)}} |c_{G_{\varphi_{0}}}|$$
$$\leq \frac{1}{q^{n-1}} \sum_{\substack{\varphi_{1} \in [q]^{V(G')} \\ \varphi_{1} \in [q]^{V(G')}}} c'_{G'_{\varphi_{1}}}$$
$$= \sum_{H} c'_{H} \cdot \mathbb{P}[G'_{q} \cong H].$$

Iteratively applying Lemma 3.4 until G is connected, and then applying Lemma 3.3, we obtain the following result, which allows us to upper bound μ by a finite computation.

Proposition 3.5. Fix q > 0, a list $\{H\}$ of unlabeled graphs on at most n_0 vertices and $\{c_H\}$ a list of coefficients. Then for any G on $n > n_0$ labeled vertices, we have

$$\sum_{H} |c_{H}| \cdot \mathbb{P}[G_{q} \cong H] \le \max_{G' \subseteq K_{n_{0}}} \left[\frac{1}{q^{k(G')-1}} \sum_{H} \widetilde{c_{H}} \cdot \mathbb{P}[G'_{q} \cong H] \cdot DC_{q,n_{0}}(v(H)) \right],$$
(7)

where $\widetilde{c_H}$ is the maximum value of $|c_{H'}|$ over all graphs H' which may be transformed to H by repeatedly identifying pairs of disconnected vertices.

4. Selecting a feasible dual solution

We now prove Proposition 2.5 by computer verification. Computations were scripted in Python 3.6.5 using integer arithmetic. We use the list of all graphs up to nine vertices, taken from [12]. Our code is attached as ancillary files to the arXiv version of this article.

4.1. The case t = 3. The choices of c_H are listed in Table 1. It turns out that in order to satisfy the conclusions of Proposition 2.5, the values of c_H for $H \subseteq K_3$ are uniquely determined. The other coefficients were chosen with some flexibility.

H	Ø	.^•	\land	•—• •—•	V.	\mathcal{N}	.		\sim
c_H	7	-5	-1	1.7	3	0.3	-0.2	-3.7	-0.75

TABLE 1. Coefficients for t = 3.

With this choice of c_H , we verify Proposition 2.5 for t = 3.

Lemma 4.1. With c_H chosen as in Table 1, and $\mu(G) = (-1)^{e(G)} \sum_H c_H \mathbb{P}[G_q \cong H]$, one has

- (1) $\mu(0) = 7$.
- (2) $|\mu(G)| \leq 1$ whenever G has at most 3 edges.
- (3) $|\mu(G)| \leq 0.9875$ whenever G has more than 3 edges.

Proof. For the empty graph, G_q is always empty, and so $\mu(0) = c_{\emptyset} = 7$. Properties (2) and (3) are verified by a finite computation, iterating through all 2-colorings of all graphs with up to 9 vertices. Property (3) is verified for graphs with more than 9 vertices by applying Proposition 3.5 with $n_0 = 9$, again by iterating through all 2-colorings of all graphs with up to 9 vertices. \Box

4.2. The case t = 4. Our choices for the coefficients c_H are listed in Table 2. For t = 4 one needs to choose a substantially larger number of coefficients to be nonzero. Despite this, there is still considerable flexibility. To make a reasonably concise presentation, we choose our list of coefficients $\{c_H\}$ to be constant on certain equivalence classes of H, so that we may simply write down a single choice for each class. The classes are defined as follows:

Definition 4.2. A block of a graph H is a maximal connected subgraph with at least one edge and no cut vertex. The collection of blocks of H partitions E(H). We say two graphs H and H'are equivalent and write $H \sim H'$ if the collection of blocks of H and the collection of blocks of H'are equal as multisets of unlabeled graphs.

Example 4.3. The graphs \checkmark and $\stackrel{\leftarrow \bullet}{\leftarrow}$ are equiavlent, with blocks \checkmark , \checkmark .

Example 4.4. The graph $\stackrel{\wedge}{\leftrightarrow}$ has blocks $\stackrel{\wedge}{\leftrightarrow}$.

We will choose c_H to be equal for all graphs within each equivalence class. This additional restriction is motivated in two ways. First, all graphs H' involved in the computation of \widetilde{c}_H (see Proposition 3.5) satisfy $H' \sim H$, and so with this choice one has $\widetilde{c}_H = |c_H|$ for all H. Second, it turns out that for general t, if $H \sim H'$ and both are contained in a K_t , in order to satisfy the conclusions of Proposition 2.5 one must choose $c_H = c_{H'}$ regardless.

Blocks of H	c_H	Blocks of H	c_H
Ø	63		-20.3
	-30	\bowtie	24.75
• , • [•]	12		-2.274
Δ	-63	⊑>, ✓	2
Δ , \checkmark	6		-11.528
	-39		-22.928
, , , , , , , , , , , , , , , , , , ,	-1.926	\sim	-12.8
	12	□>, ✓	-1.2
△, ✓, ✓	5.478	×, /	-4.362
, ,	4.293		1.138
	-16.5		

TABLE 2. Coefficients for t = 4

Lemma 4.5. With c_H chosen as in Table 1, and $\mu(G) = (-1)^{e(G)} \sum_H c_H \mathbb{P}[G_q \cong H]$, one has

- (1) $\mu(0) = 63$,
- (2) $|\mu(G)| \leq 1$ whenever G has at most 6 edges, and
- (3) $|\mu(G)| \leq 0.999$ whenever G has more than 6 edges.

Proof. For the empty graph, G_q is always empty, and so $\mu(0) = c_{\emptyset} = 63$. Properties (2) and (3) are verified by a finite computation, iterating through all 3-colorings of all graphs with up to 9 vertices. On graphs with at most 9 vertices and more than 6 edges, the maximum value of $|\mu(G)|$ is $1 - 188406/(1000 \cdot 3^9)$, which is attained by 24. Property (3) is verified for graphs with more than 9 vertices by applying Proposition 3.5 with $n_0 = 9$, by iterating through all 3-colorings of all graphs. \Box graphs with up to 9 vertices, which gives a bound of 1 - 20/19683.

5. Additional proofs

Here we collect the proofs deducing Theorem 1.5 and Corollary 1.9 from Theorem 1.8. The following fact is well known (see [8, Theorem 2.38] or [5, Lemma 2.6]).

Lemma 5.1. Let $\mathcal{F} \subseteq 2^{[n]}$ be a nonempty monotone increasing family (that is, for all $X \in \mathcal{F}$ and $Y \supseteq X$, $Y \in \mathcal{F}$). For $0 let <math>\mu_p(\mathcal{F}) = \sum_{X \in \mathcal{F}} p^{|X|} (1-p)^{n-|X|}$. Then $\log_p \mu_p(\mathcal{F})$ is non-increasing in p.

Corollary 5.2. Let $0 and let <math>(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-intersecting pair of subsets of [N]. Then $\mu_p(\mathcal{F}_1)\mu_p(\mathcal{F}_2) \le p^2$.

Proof. It suffices to prove this bound when \mathcal{F}_1 and \mathcal{F}_2 are replaced by their upwards closures (the upwards closure of \mathcal{F} is the set $\{H : \exists G \in \mathcal{F} : G \subseteq H\}$). By Lemma 5.1 it suffices to consider the p = 1/2 case. In this setting, consider complementary pairs of sets (G, \overline{G}) . For $(\mathcal{F}_1, \mathcal{F}_2)$ to be an intersecting pair of families, we must have $\mathbb{1}_{G \in \mathcal{F}_1} + \mathbb{1}_{\overline{G} \in \mathcal{F}_2} \leq 1$. Taking the expectation over all G, by linearity of expectation this becomes $\mu_{1/2}(\mathcal{F}_1) + \mu_{1/2}(\mathcal{F}_2) \leq 1$, and so by AM-GM $\mu_{1/2}(\mathcal{F}_1)\mu_{1/2}(\mathcal{F}_2) \leq 1/4$, which completes the proof.

We continue with a lemma that allows us to strengthen the conclusions of Theorem 1.8 from K_t -conjunctions to K_t -unvirates under the stronger hypothesis that $\mathcal{F}_1, \mathcal{F}_2$ are cross-*t*-chromaticintersecting. Recall that $\mu_p(\mathcal{F}) := \sum_{G \in \mathcal{F}} p^{e(G)}(1-p)^{e(\overline{G})}$. For a set of edges $E_0 \subseteq {\binom{[n]}{2}}$, we say a family of graphs \mathcal{F} is an E_0 -conjunction if there is some graph H with $E(H) \subseteq E_0$ so that $\mathcal{F} = \{G : E_0 \subseteq E(G \triangle H)\}$. If $|E_0| = k$ we say \mathcal{F} is a *k*-conjunction. **Lemma 5.3.** Let $t \in \mathbb{N}$ and let $0 . There exists a constant <math>\epsilon_{p,t} > 0$ so that the following holds. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-intersecting pair of families of graphs on n labeled vertices, and let $\mathcal{T} \subseteq 2^{\binom{[n]}{2}} \times 2^{\binom{[n]}{2}}$ be a k-conjunction for $k \leq 2\binom{t}{2}$. If \mathcal{T} is not the Cartesian product of identical K_t -unvirates, then $\mu_p((\mathcal{F}_1 \times \mathcal{F}_2) \cap \mathcal{T}) \leq (1 - \epsilon_{p,t})\mu_p(\mathcal{T})$.

Proof. Any conjunction \mathcal{T} admits a product decomposition $\mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2$, where $\mathcal{T}_i : 2^{\binom{[n]}{2}} \to \{0, 1\}$. We can write $\mathcal{T}_i = \{G : E_i \subseteq E(G \triangle H_i)\}$ for some graphs H_i and sets of edges E_i where $|E_1| + |E_2| \leq 2\binom{t}{2}$. If we are not in the case where $E_1 = E_2 \cong K_t$ and $H_1 = H_2 = \emptyset$, then for any $G_1 \in \mathcal{T}_1$, $G_2 \in \mathcal{T}_2$, we see $(G_1 \cap G_2) \cap (E_1 \cap E_2)$ has at most $\binom{t}{2}$ edges and is not a K_t , and therefore must be (t-1)-colorable, and so G_1 and G_2 must share an edge outside $E_1 \cap E_2$.

Now we construct subsets of each conjunction that "disagree with the other conjunction as much as possible." Consider $E_2 \setminus E_1$. Graphs in \mathcal{T}_2 must disagree with H_2 on this set, but graphs in \mathcal{T}_1 are unconstrained on this set of edges. Consider the subset of \mathcal{T}_1 given by graphs that agree with H_2 on this set: $\mathcal{T}'_1 := \{G \in \mathcal{T}_1 : (G \triangle H_2) \cap (E_2 \setminus E_1) = \emptyset\}$. Define \mathcal{T}'_2 similarly. Then since $|E_2 \setminus E_1| \le 2\binom{t}{2}$, we see $\mu_p(\mathcal{T}'_i) \ge p^{2\binom{t}{2}}\mu_p(\mathcal{T})$, and each pair $G_1 \in \mathcal{F}_1 \cap \mathcal{T}'_1, G_2 \in \mathcal{F}_2 \cap \mathcal{T}'_2$ intersects on no edges in $E_1 \triangle E_2$. We know that G_1 and G_2 must share an edge outside $E_1 \cap E_2$, which must therefore lie outside $E_1 \cup E_2$. Letting $\overline{\mathcal{F}_i} = \{E(G) \setminus (E_1 \cup E_2) : G \in \mathcal{F}_i \cap \mathcal{T}'_i\}$ (removing duplicates), we have that $(\overline{\mathcal{F}_1}, \overline{\mathcal{F}_2})$ is a cross-intersecting pair of families. Restricting our domain to $\binom{[n]}{2} \setminus (E_1 \cup E_2)$, by Corollary 5.2 we have $\mu_p(\overline{\mathcal{F}_1})\mu_p(\overline{\mathcal{F}_2}) < p^2$. Pulling this back, we see $\mu_p(\mathcal{F} \cap (\mathcal{T}'_1 \times \mathcal{T}'_2)) \le p^2\mu_p(\mathcal{T}'_1 \times \mathcal{T}'_2)$.

To conclude, write $\mu_p(\mathcal{F} \cap \mathcal{T}) = \mu_p((\mathcal{F} \cap \mathcal{T}) \setminus (T'_1 \times T'_2)) + \mu_p(\mathcal{F} \cap (T'_1 \times T'_2))$. The first term is trivially bounded by $\mu_p(\mathcal{T} \setminus (T'_1 \times T'_2))$, and we just saw that the second term is at most $p^2 \mu_p(T'_1 \times T'_2)$. Adding in our lower bound $\mu_p(\mathcal{T}_i) \geq p^{-2\binom{t}{2}}$, we conclude $\mu_p(\mathcal{F} \cap \mathcal{T}) \leq (1 - (1 - p^2)p^{4\binom{t}{2}})\mu_p(\mathcal{T})$, which completes the proof.

We can apply this to immediately deduce Theorem 1.5 from Theorem 1.8.

Lemma 5.4. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-intersecting pair of families of graphs on n labeled vertices. Then equality holds in Theorem 1.8 if and only if $\mathcal{F}_1 = \mathcal{F}_2$ is a K_t -umvirate, and for sufficiently small ϵ the K_t -conjunction \mathcal{T} from (stability) must be a K_t -umvirate.

Proof. The upper bound and uniqueness follow immediately by noting that the only K_t -conjunction that is K_t -intersecting is a K_t -unvirate. For stability, if Theorem 1.8 outputs some \mathcal{T} that is not a K_t -unvirate, we can apply Lemma 5.3 to $\mathcal{T} \times \mathcal{T}$ with p = 1/2 and deduce $|\mathcal{F}_1 \cap \mathcal{T}||\mathcal{F}_2 \cap \mathcal{T}| \leq (1 - c_t)|\mathcal{T}|^2$. But then we cannot have satisfied the hypothesis of Theorem 1.8 (stability) for sufficiently small ϵ .

Proof of Theorem 1.5. Lemma 5.4 completes the proof of Theorem 1.5 in the more general cross-t-chromatic-intersecting setting, for sufficiently small ϵ . Choosing a new C_t large enough makes the conclusion of (stability) trivially satisfied for large ϵ , which completes the deduction.

Now that we have handled the p = 1/2 case for cross-t-chromatic-intersecting families, we next apply the results of [5] to obtain results for all p < 1/2 to yield Corollary 1.9.

Corollary 5.5. Let $t \in \{3,4\}$, $0 , and let <math>(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-intersecting pair of families of graphs on n labeled vertices. Then $\mu_p(\mathcal{F}_1)\mu_p(\mathcal{F}_2) \leq p^{2\binom{t}{2}}$, with equality if and only if $\mathcal{F}_1 = \mathcal{F}_2$ is a K_t -unvirate.

Proof. The upwards closures of \mathcal{F}_1 and \mathcal{F}_2 also form a cross-*t*-chromatic-intersecting pair of families, so for the original pair to be maximal both families be upwards-closed. The result then follows from the p = 1/2 version above and Lemma 5.1.

In order to deduce stability we apply the following powerful result from [5].

Theorem 5.6 ([5, Theorem 3.1]). Let $t \in \mathbb{N}$ and 0 . Then exist constants <math>C > 4 and c > 0 so that the following holds. Let $f : \mathbb{F}_2^N \to \{0,1\}$ be increasing with $\mathbb{E}[f] \leq 2^{-t}$ and

$$\mu_p(\mathcal{F}) \ge \begin{cases} p^t \left(1 - c \left(\frac{1}{2} - p \right) \right) & p \ge 1/C\\ C p^{t+1} & p < 1/C. \end{cases}$$

For all $\epsilon > 0$, if

$$\mu_p(f) \ge p^t \left(1 - \epsilon^{\log_p(1-p)}\right) + p^{t-1}(1-p)\epsilon_p$$

then there is some t-unvirate \mathcal{G} such that $\mu_p(\mathcal{F} \setminus \mathcal{G}) \leq (1-p)p^{t-1}\epsilon$.

Corollary 5.7. Let $t \in \mathbb{N}$ and let $0 . Then there exists a constant <math>C_{p,t} > 0$ so that the following holds. Let $(\mathcal{F}_1, \mathcal{F}_2)$ be a cross-t-chromatic-intersecting pair of families of graphs on n labeled vertices. Then if $\mu_p(\mathcal{F}) \ge (1 - \epsilon)p^{\binom{t}{2}}$, then there exists some K_t -unvirate \mathcal{T} so that $\mu_p(\mathcal{F}_i \setminus \mathcal{T}) \le C_{p,t} \epsilon^{\log_{(1-p)}(p)}$ for $i \in \{1, 2\}$.

Proof. By Lemma 5.4, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ has size $\mu_{1/2}(\mathcal{F}) \leq 2^{-2\binom{t}{2}}$. Then by Theorem 5.6 if $\mu_p(\mathcal{F}) \geq (1-\epsilon)p^{2\binom{t}{2}}$, there is some $2\binom{t}{2}$ -umvirate \mathcal{T} so that $\mu_p(\mathcal{F} \setminus \mathcal{T}) \leq C'_{p,t} \epsilon^{\log_{(1-p)}(p)}$. If \mathcal{T} is not the Cartesian product of identical K_t -umvirates, apply Lemma 5.3. Then $\mu_p(\mathcal{F}_1)\mu_p(\mathcal{F}_2) = \mu_p(\mathcal{F} \cap \mathcal{T}) + \mu_p(\mathcal{F} \setminus \mathcal{T}) \leq (1-\epsilon_{p,t})\mu_p(\mathcal{T}) + C'_{p,t} \epsilon^{\log_{(1-p)}(p)} < (1-\epsilon)p^{2\binom{t}{2}}$ for ϵ sufficiently small. Choosing $C_{p,t}$ large enough makes the conclusion trivial for large ϵ , which completes the proof.

6. Further work

Given these results, we are optimistic that this framework is strong enough to obtain optimal bounds on K_t -intersecting families for each fixed t (given enough computational power). Despite this, proving a result for all t will require some new insight. We see two potential directions to proving a result in full generality. The first is to bypass the c_H setup entirely by writing down a μ which satisfies the conclusions of Proposition 2.5. We see no obvious candidates for such μ , and the constructions of [4] and this paper provide no clear pattern for generalization. The second option would be to write down a general form for the coefficients c_H , and use some technique akin to Proposition 3.5 to reduce checking Proposition 2.5 to a bounded computation that can be done by hand. This also seems difficult, as again there is no clear pattern for generalization among the feasible solutions for t = 3, 4.

The *tetrahedron* is the complete 3-uniform hypergraph on 4 vertices. In correspondence, Ellis, Filmus, and Friedgut suggested that the following version of Problem 1.1 would be interesting.

Problem 6.1. What is the maximal size of a tetrahedron-intersecting family of 3-uniform hypergraphs on n labeled vertices?

References

- Aaron Berger, Ross Berkowitz, Pat Devlin, Michael Doppelt, Sonali Durham, Tessa Murthy, and Harish Vemuri, Connected-intersecting families of graphs, arXiv:1901.01616.
- [2] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer, Some intersection theorems for ordered sets and graphs, J. Combin. Theory Ser. A 43 (1986), 23–37.
- [3] Philippe Delsarte, An algebraic approach to the association schemes of coding theory, Philips Res. Rep. Suppl. (1973), vi+97.
- [4] David Ellis, Yuval Filmus, and Ehud Friedgut, Triangle-intersecting families of graphs, J. Eur. Math. Soc. (JEMS) 14 (2012), 841–885.
- [5] David Ellis, Nathan Keller, and Noam Lifshitz, Stability versions of Erdős-Ko-Rado type theorems via isoperimetry, J. Eur. Math. Soc. (JEMS) 21 (2019), 3857–3902.

BERGER AND ZHAO

- [6] Paul Erdős, Chao Ko, and Richard Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford Ser. (2) 12 (1961), 313–320.
- [7] Ehud Friedgut, On the measure of intersecting families, uniqueness and stability, Combinatorica 28 (2008), 503–528.
- [8] Geoffrey Grimmett, Percolation, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 321, Springer-Verlag, Berlin, 1999.
- [9] Nathan Keller and Noam Lifshitz, A note on large H-intersecting families, SIAM J. Discrete Math. 33 (2019), 398–401.
- [10] Guy Kindler and Shmuel Safra, Noise-resistant boolean functions are juntas, unpublished (2004), https://www.cs.huji.ac.il/~gkindler/.
- [11] János Körner, Silvia Messuti, and Gábor Simonyi, Families of graph-different Hamilton paths, SIAM J. Discrete Math. 26 (2012), 321–329.
- [12] Brendan McKay, Combinatorial data, https://users.cecs.anu.edu.au/~bdm/data/.

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA *Email address*: {bergera,yufeiz}@mit.edu