

## CONGRUENCE CLASSES OF SIMPLEX STRUCTURES IN FINITE FIELD VECTOR SPACES

TIMOTHY CHEEK, JOSEPH COOPER, PICO GILMAN, ALEX IOSEVICH, KAREEM JABER,  
EYVINDUR PALSSON, VISMAY SHARAN, JENNA SHUFFELTON,  
AND MARIE-HÉLÈNE TOMÉ

**ABSTRACT.** We study a generalization of the Erdős-Falconer distance problem over finite fields. For a graph  $G$ , two embeddings  $p, p' : V(G) \rightarrow \mathbb{F}_q^d$  of a graph  $G$  are congruent if for all edges  $(v_i, v_j)$  of  $G$  we have that  $\|p(v_i) - p(v_j)\| = \|p'(v_i) - p'(v_j)\|$ . What is the infimum of  $s$  such that for any subset  $E \subset \mathbb{F}_q^d$  with  $|E| \gtrsim q^s$ ,  $E$  contains a positive proportion of congruence classes of  $G$  in  $\mathbb{F}_q^d$ ? Bennett et al. and McDonald used group action methods to prove results in the case of  $k$ -simplices. The work of Iosevich, Jardine and McDonald as well as that of Bright et al. have proved results in the case of trees and trees of simplices, utilizing the inductive nature of these graphs.

Recently Aksoy, Iosevich, and McDonald combined these two approaches to obtain nontrivial bounds on the “bowtie” graph, two triangles joined at a vertex. Their proof relies on an application of the Hadamard three-lines theorem to pass to a different graph. We develop novel geometric techniques called branch shifting and simplex unbalancing to reduce our analysis of trees of simplices to a much smaller class of simplex structures. This allows us to establish a framework that handles a wide class of graphs exhibiting a combination of rigid and loose behavior. In  $\mathbb{F}_q^2$ , this approach gives new nontrivial bounds on chains and trees of simplices. In  $\mathbb{F}_q^d$ , we improve on the results of Bright et al. in many cases and generalize their work to a wider class of simplex trees. We discuss partial progress on how this framework can be extended to more general simplex structures, such as cycles of simplices and structures of simplices glued together along an edge or a face.

*“The shortest path between two truths in the real domain passes through the complex domain.”*

- Jacques Hadamard

### CONTENTS

1. Introduction	42
1.1. Statement of main results	45
2. Preliminaries	49
2.1. Analytic preliminaries	49

---

*Date:* Submitted May 3, 2025. Accepted October 10, 2025.

*2020 Mathematics Subject Classification.* 52C10, 42B35.

*Key words and phrases.* Erdős distance problem, simplices in subsets of finite fields, group actions, Fourier transform.

2.2.	Geometric preliminaries	50
2.3.	Simplex trees	52
3.	Functional sums and useful bounds	53
3.1.	Proof of Lemma 3.3	56
3.2.	Proof of Lemma 3.4	58
4.	Proof of main theorems	61
4.1.	The setup	61
4.2.	Two geometric operations	66
4.3.	Proof of Proposition 4.4 and Theorem 1.4	72
5.	Cycles of simplices	75
5.1.	The setup and Hadamard three-lines reduction	76
5.2.	Technical obstructions	80
6.	Future work	83
6.1.	Edge gluing	83
6.2.	Thresholds	85
7.	Acknowledgements	85
	References	86

## 1. INTRODUCTION

The Erdős distance problem is a classical problem in geometric combinatorics. Originally introduced by Erdős in 1946, it asks for the smallest number of distinct distances determined by a subset of  $\mathbb{R}^d$  under the standard Euclidean metric. The original conjecture in [Erd46] states that for a subset  $P \subset \mathbb{R}^d$ ,  $d \geq 2$ , with size  $|P| = N$ , the least number of distances determined by  $P$  is on the order of  $N^{\frac{2}{d}-\epsilon}$  for  $\epsilon > 0$ . The case of  $d = 2$  was resolved by Guth and Katz in 2011 within a  $\sqrt{\log N}$  factor of the best known lower bound ([GK15]). The case where  $d \geq 3$  is still open with best known bounds from Solymosi and Vu in 2008 ([SV08]). The Falconer distance problem is similar in nature to the original Erdős distance problem and asks for the minimum Hausdorff dimension of a set  $E \subset \mathbb{R}^d$  needed to achieve a positive Lebesgue measure of distances. In [Fal85], Falconer conjectured that if the Hausdorff dimension of  $E$  is strictly greater than  $\frac{d}{2}$ , the set of distances  $\{\|x - y\|, x, y \in E\}$  with the standard Euclidean metric has positive (one-dimensional) Lebesgue measure. This conjecture is still open in all dimensions. For  $d = 2$ , the best threshold is  $\frac{5}{4}$  from Guth et al. in 2018 ([Gut+20]), and for  $d \geq 3$  the best known bound is  $\frac{d}{2} + \frac{1}{4} - \frac{1}{8d+4}$  were found by Du et al. in 2023 ([Du+23]).

In [IR07], Iosevich and Rudnev introduced the Erdős-Falconer distance problem over finite fields, an arithmetic analogue to these two problems. In  $\mathbb{F}_q^d$ , for  $q$  a prime power, we define the distance between two points  $x, y \in \mathbb{F}_q^d$  as

$$\|x - y\| = (x_1 - y_1)^2 + \cdots + (x_d - y_d)^2. \quad (1.1)$$

The Erdős-Falconer distance problem over  $\mathbb{F}_q^d$  asks for the minimum size of a subset  $E$  such that pairs of points in  $E$  determine all possible distances in  $\mathbb{F}_q$ , or more generally, a positive proportion of distances in  $\mathbb{F}_q$ <sup>1</sup>. More precisely, it asks for the infimum of all  $s \in \mathbb{R}$  such that  $|E| \gtrsim q^s$  implies that  $E$  contains all distances, or a positive proportion of distances, in  $\mathbb{F}_q$ , where  $q$  is treated as an asymptotic parameter.

**Remark.** Here, as in most other papers on the subject, we write  $X \lesssim Y$  to mean there is a constant  $C$  with  $X \leq CY$  independent of  $q$ , and similarly, we write  $X \gtrsim Y$  if  $X \geq CY$  for a constant  $C$  independent of  $q$ . We say  $X \approx Y$ , or that  $X$  is on the order of  $Y$ , if  $X \lesssim Y \lesssim X$ .

This problem has been actively researched over the past two decades. In [IR07], it was proved that for  $E \subset \mathbb{F}_q^d$  and  $s = \frac{d+1}{2}$ , if  $E \gtrsim q^s$  then  $E$  contains all distances in  $\mathbb{F}_q$ . In [Har+11], the authors showed that this  $s$  is tight in odd dimensions even for the weaker problem of considering a positive proportion of distances. In  $d = 2$ , this exponent has been improved when considering a positive proportion of distances, first to  $\frac{4}{3}$  in [Cha+12], and recently to  $\frac{5}{4}$  in [Mur+22], which is the best known result as of this writing.

An interesting analogue of the original Erdős-Falconer distance problem in  $\mathbb{F}_q^d$  generalizes the notion of the distance set determined by a set of points  $E$  in  $\mathbb{F}_q^d$ .

**Definition 1.1.** *Given a graph  $G$ , we say that two embeddings  $p, p' : V(G) \rightarrow \mathbb{F}_q^d$  of a graph  $G$  are congruent if for every  $(v_i, v_j)$  in the edge set of  $G$ , we have that  $\|p(v_i) - p(v_j)\| = \|p'(v_i) - p'(v_j)\|$ .*

For a general graph  $G$ , what is the infimum of  $s$  such that  $|E| \gtrsim q^s$  implies that  $E$  contains an embedding of  $G$  in each congruence class, or a positive proportion of congruence classes of  $G$  in  $\mathbb{F}_q^d$ ? The original Erdős-Falconer distance problem over  $\mathbb{F}_q^d$  relates to the case where  $G$  is the complete graph on two vertices.

A natural graph to consider in this generalization is the complete graph on  $k + 1$  vertices, whose embeddings correspond to  $k$ -simplices in  $\mathbb{F}_q^d$ . To this end, we often refer to both the complete graph on  $k + 1$  vertices and embeddings of this graph in  $\mathbb{F}_q^d$  as simplices of dimension  $k$ . In [Ben+17], Bennett et al. proved that for  $n \leq d$ ,  $E \subset \mathbb{F}_q^d$ , and  $s = \frac{dn+1}{n+1}$ , then  $E$  contains a positive proportion of congruence classes of  $n$ -simplices in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ . To achieve this, the authors adapted group action methods from the continuous case of the Falconer distance problem, see [Gre+15]. While the distance function in (1.1) isn't a norm, it is still invariant under transformations in  $O_d(\mathbb{F}_q)$ , the group of orthogonal transformations over  $\mathbb{F}_q^d$  with respect to the standard inner product. These group action methods over finite fields have found success in proving results for rigid graphs. For  $d = 2$ , [Ben+17] improved the  $s = \frac{dn+1}{n+1}$  exponent for 2-simplices due to a stronger bound on the finite field analogue

<sup>1</sup>Although this is not a metric, it is standard in the literature to refer to it as a distance due to historical connections to the Erdős and Falconer distance problems.

of the Mattila integral in this setting. In 2019, McDonald [McD20] used group action techniques to extend the results of [Ben+17] in both the  $\mathbb{F}_q^d$  and  $\mathbb{F}_q^2$  settings to  $n$ -simplices for  $n \geq d$ .

Results have been proven for loose graphs as well, such as paths, cycles, and trees, where a group action approach is less clear. In this setting, success was found in using techniques that take advantage of the inductive nature of these graphs. In [Ben+16], the authors proved that for  $G$  a path of arbitrary length,  $E \subset \mathbb{F}_q^d$  and  $s = \frac{d+1}{2}$ , if  $|E| \gtrsim q^s$ , then  $E$  contains at least one embedding of  $G$  in every congruence class, and so  $E$  contains all congruence classes of  $G$  in  $\mathbb{F}_q^d$ . Similar results have been proved in [IJM21] when  $G$  is a tree for  $s = \frac{d+1}{2}$ . These results have been generalized in [Bri+23], where for simplices of sufficiently small dimension relative to the ambient space, the authors proved that for  $E \subset \mathbb{F}_q^d$ , and  $s = m + \frac{d-1}{2}$ , if  $|E| \gtrsim q^s$  then  $E$  contains all congruence classes of embeddings of a certain class of trees of  $m$ -simplices.

In [IP19], Iosevich and Parshall proved a much more general result: for any connected graph  $G$  with maximum vertex degree  $t$ ,  $E \subset \mathbb{F}_q^d$ , and  $s = t + \frac{d-1}{2}$ , if  $|E| \gtrsim q^s$  then  $E$  contains at least one embedding of  $G$  in every congruence class. While this result gives nontrivial  $s$  for any  $G$  in a sufficiently high dimension space, the specialized results of [Ben+17] and [McD20] yield better results on high-dimensional simplices, and [Bri+23] yields better results on the class of simplex trees they handle for all dimensions of the space. This suggests that the techniques that are used to work with a graph  $G$  should vary with the properties that  $G$  exhibits.

Very recently, in [AIM24], Aksoy, Iosevich, and McDonald took a step toward combining the group action techniques for simplices and the inductive techniques for paths and trees to handle the bowtie graph  $B$  shown in Figure 1, a graph that exhibits a combination of rigid and loose behavior. To do this, the authors first reduced the problem to bounding a sum corresponding to the number of pairs of congruent embeddings of  $B$  in  $E$ , following a similar setup as in [Ben+17]. Then, by the Hadamard three-lines theorem from complex analysis, they convert the problem from bounding the sum corresponding to  $B$  (a 2-simplex attached to a 2-simplex) to a sum corresponding to the kite graph  $K$  (a 3-simplex attached to a 1-simplex), as shown in Figure 1.

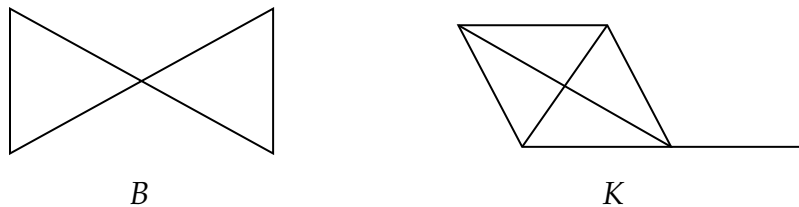


FIGURE 1. The “bowtie” graph,  $B$ , and the “kite” graph,  $K$ .

Morally, as paths are much more abundant than high-dimensional simplices, if  $E$  is large enough to contain a positive proportion of congruence classes of 3-simplices,

then  $E$  should also contain a positive proportion of congruence classes of  $K$ . To prove this rigorously, the authors split up the sum corresponding to  $K$  to isolate the contributions from the path and those from the 3-simplex. Then, the path terms can be bounded using the inductive techniques from [Ben+16], and the 3-simplex terms could be bounded using the group action techniques from [Ben+17].

In this paper, we generalize the techniques of [AIM24] to obtain nontrivial bounds on  $s$  for a much wider and more complex class of graphs exhibiting a combination of rigid and loose structure. In particular, we focus our attention on a class of graphs that we call simplex trees.

**Definition 1.2.** For a simplex  $S$ , let  $V(S)$  denote the set of vertices of  $S$ . Let  $\mathcal{T}$  be a single simplex or let  $\mathcal{T} = \{S_i : i \in I\}$  be a finite set of simplices. We denote by  $N_{\mathcal{T}}(S_i)$  the set of all simplices in  $\mathcal{T}$  sharing a vertex with  $S_i$ . We drop the subscript  $\mathcal{T}$  when it is clear from context. We say  $\mathcal{T}$  is a **(connected) simplex tree** if  $|\mathcal{T}| = 1$  or if  $|\mathcal{T}| > 1$  and  $\mathcal{T}$  satisfies the following properties:

- (1) For any  $i, j \in I$  distinct, we have that  $|V(S_i) \cap V(S_j)| \leq 1$ .
- (2) For each  $a, b \in I$ , there exists a path of simplices in  $\mathcal{T}$  from  $S_a$  to  $S_b$ , i.e., there exist distinct integers  $i_1, \dots, i_n \in I$  such that  $S_{i_1} = S_a$ ,  $S_{i_n} = S_b$  and  $S_{i_{j+1}} \in N(S_{i_j})$  for each  $1 \leq j \leq n - 1$ .
- (3) There is no closed loop of simplices, i.e. there do not exist distinct integers  $i_1, \dots, i_n \in I$  such that for each  $1 \leq j \leq n$ , we have that  $S_{i_{j+1}} \in N(S_{i_j})$  and  $S_{i_n} \in N(S_{i_1})$ . Note that for a connected simplex tree,  $S_i \in N(S_j)$  if and only if  $|V(S_i) \cap V(S_j)| = 1$ .

Our definition may be viewed as a natural generalization of an undirected tree, where each simplex may be viewed as a vertex and each vertex shared by two simplices may be viewed as an edge between them.

**Remark.** Our definition of a simplex tree is more general than the definition given in [Bri+23]. Their class of simplex trees is formed by taking a standard graph-theoretic tree, and to each edge, attaching a  $k$ -simplex sharing this edge. This means that for any simplex in a tree considered by [Bri+23], only two of its vertices may be shared by other simplices. Under our definition of simplex trees, however, a simplex can share any number of its vertices with other simplices. Secondly, while the definition in [Bri+23] requires all simplices in the tree to have the same dimension, in our definition, simplices can have different dimensions throughout the tree.

**1.1. Statement of main results.** Our main theorem deals with how big  $E$  needs to be to ensure it contains a positive proportion of congruence classes of  $\mathcal{T}$ .

**Theorem 1.3.** For a simplex tree  $\mathcal{T}$  and an integer  $k$  such that  $1 \leq k < \frac{d+1}{2}$ , define

$$N_k = k + \sum_{\substack{S \in \mathcal{T} \\ \dim(S) > k}} (\dim(S) - k). \quad (1.2)$$

For an  $E \subset \mathbb{F}_q^d$ , suppose that for  $s = \max\left(\frac{dN_k+1}{N_k+1}, k + \frac{d-1}{2}\right)$ ,  $|E| \gtrsim q^s$ . Then  $E$  contains a positive proportion of congruence classes of embeddings of  $\mathcal{T}$  in  $\mathbb{F}_q^d$ .

Note that for a given  $\mathcal{T}$  and  $k$ , Theorem 1.3 results in some  $s < d$ , and therefore this result is nontrivial for all simplex trees  $\mathcal{T}$ . Naively, for any graph  $G$  and a space  $\mathbb{F}_q^d$ , one can find a nontrivial  $s$  (i.e.  $s < d$ ) such that for any  $E \subset \mathbb{F}_q^d$  and  $|E| \gtrsim q^s$ ,  $E$  contains a positive proportion of congruence classes of  $G$  in  $\mathbb{F}_q^d$  by first finding an  $s$  for the complete graph on the number of vertices of  $G$ , and then applying the corresponding results of [Ben+17] and [McD20]. However, Theorem 1.3 and all subsequent theorems in the introduction yield better bounds than this method.

The idea of the proof of Theorem 1.3 is as follows. We first generalize the setup of [Ben+17] to reduce the problem to bounding a sum corresponding to the number of pairs of congruent embeddings of  $\mathcal{T}$  in  $E$ . We develop two techniques called *branch shifting* and *simplex unbalancing* that algebraically modify this sum in ways that correspond to geometric modifications of  $\mathcal{T}$ . This allows us to reduce to the problem to bounding a sum corresponding to the modified  $\mathcal{T}$ . We apply these techniques repeatedly until our modified  $\mathcal{T}$  consists of one large simplex of dimension  $N_k$  with trees of  $k$ -simplices attached to a single vertex. From here, we use Fourier analytic techniques to isolate the contributions of each of these two different structures from the sum, handling the large simplex with group action techniques in the spirit of [Ben+17] and [McD20] and handling the smaller trees of simplices with the inductive techniques from [IJM21] and [Bri+23]. Once  $E$  is large enough to contain a positive proportion of congruence classes of  $N_k$ -simplices and all congruence classes of trees of  $k$ -simplices, which occur at  $s = \frac{dN_k+1}{N_k+1}$  and  $s = k + \frac{d-1}{2}$ , respectively, we show that  $E$  contains a positive proportion of congruence classes of  $\mathcal{T}$ . Therefore, the geometry of our reduced structure dictates our threshold for  $s$ .

The value of  $k$  in Theorem 1.3 can be viewed as a parameter that determines how much the resulting  $s$  relies on inductive techniques rather than group action techniques. When  $k$  is small, we modify  $\mathcal{T}$  to have a very large central simplex and simplex trees of very small dimension attached to a vertex. Then, the strength of our result relies much more heavily on group action techniques rather than inductive techniques. On the other hand, when  $k$  is large, we modify  $\mathcal{T}$  to have a smaller central simplex, while the simplex trees sticking out of it are of higher dimension. In this scenario, the group action techniques yield a smaller  $s$ , at the cost of inductive techniques yielding a larger  $s$  to guarantee that  $E$  contains all congruence classes of trees of  $k$ -simplices. The flexibility we gain by adaptively scaling between group action techniques and inductive techniques yields a vast generalization of the methods of [AIM24]. See Figure 2 for an illustration of how scaling in  $k$  affects the modification of  $\mathcal{T}$ . Note that the edges inside of the simplices are omitted for clarity.

Note that when the dimension of  $\mathbb{F}_q^d$  is much larger than the dimension of simplices in a simplex tree  $\mathcal{T}$ , it is always optimal to pick  $k$  as the dimension of the

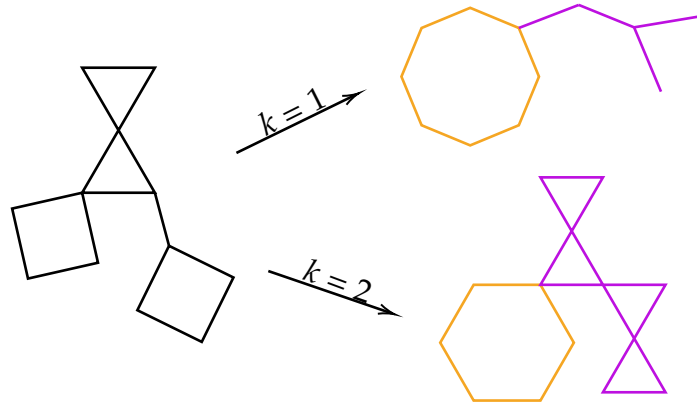


FIGURE 2. Two modifications of a simplex tree corresponding to different  $k$ .

largest simplex in  $\mathcal{T}$ . However, when the dimension of  $\mathbb{F}_q^d$  is comparable to the dimension of the simplices, the choice of  $k$  is more nuanced. For example, Theorem 1.3 gives the best results for a chain of three triangles in  $\mathbb{F}_q^4$  for  $k = 1$ , and gives the best results for a 6-simplex attached to a 3-simplex in  $\mathbb{F}_q^6$  for  $k = 2$ . If the dimension of the simplices in  $\mathcal{T}$  is bounded, it becomes optimal to pick a larger  $k$  as the number of simplices in  $\mathcal{T}$  increases.

In the case where the simplices in our simplex tree are sufficiently small, we get a result using purely inductive techniques.

**Theorem 1.4.** *Consider a simplex tree  $\mathcal{T}$  consisting only of simplices of dimension at most  $n < \frac{d+1}{2}$ . Fix  $E \subset \mathbb{F}_q^d$  and put  $s = n + \frac{d-1}{2}$ . If  $|E| \gtrsim q^s$  then  $E$  contains a positive proportion of congruence classes of  $\mathcal{T}$  in  $\mathbb{F}_q^d$ .*

This result follows by using branch shifting to convert our simplex tree  $\mathcal{T}$  to the form of a tree of  $n$ -simplices in the sense of [Bri+23]. From here, using Theorem 1.12 of [Bri+23] gives the desired result.

When working with simplex trees of only  $n$ -simplices for  $n < \frac{d+1}{2}$ , Theorem 1.4 generalizes the results of [Bri+23]. We achieve the same exponent  $s = n + \frac{d-1}{2}$  in Theorem 1.12 of [Bri+23] but for a wider class of simplex trees than they consider. In many cases where a simplex tree contains simplices of dimension close to  $\frac{d+1}{2}$ , Theorem 1.3 improves on the results of [Bri+23]. For instance, for a chain of three triangles in  $\mathbb{F}_q^4$ , choosing  $k = 1$  in Theorem 1.3 gives  $s = \frac{17}{5}$ , while Theorem 1.9 of [Bri+23] gives  $s = \frac{7}{2}$ . It is worth noting that for  $n < \frac{d+1}{2}$ , [Bri+23] and Theorem 1.4 give values of  $s$  independent of the size of the tree, but Theorem 1.3 weakens as the size of the tree grows. This is because upon Hadamard three-lines reduction, the results of group action techniques are sensitive to the size of the large simplex, while the results of inductive techniques are independent of the size of the simplex tree. As  $n$  grows, Theorem 1.3 becomes more and more reliant on group action techniques, even when  $k$  is chosen to be maximal.

For all simplex trees, we also improve the results of Iosevich and Parshall in [IP19]. When  $\mathcal{T}$  has a simplex of dimension at least  $\frac{d+1}{2}$ , the methods in [IP19] give trivial results, and when  $\mathcal{T}$  has all simplices of dimension at most  $n$  for  $n < \frac{d+1}{2}$ , applying Theorem 1.4 improves on the result of [IP19]. Furthermore, our results are independent of the maximum vertex degree of  $\mathcal{T}$ , as Theorems 1.3 and 1.4 only depend on the set of simplices in  $\mathcal{T}$  rather than their configuration. For example, these theorems yield the same bounds for a chain of  $n$ -simplices of length  $\ell$  and a collection of  $\ell$   $n$ -simplices sharing a common vertex. This gives further indication, as [Bri+23] discusses, that maximum vertex degree is not always the right notion of complexity.

Due to a stronger bound on the finite field analogue of the Mattila integral in  $\mathbb{F}_q^2$ , in this setting, we obtain a slightly better result.

**Theorem 1.5.** *For a simplex tree  $\mathcal{T}$ , define*

$$N = 1 + \sum_{\substack{S \in \mathcal{T}, \\ \dim(S) > 1}} (\dim(S) - 1).$$

*If  $E \subset \mathbb{F}_q^2$  with  $q \equiv 3 \pmod{4}$  is such that*

$$|E| \gtrsim q^{\frac{4N}{2N+1}},$$

*then  $E$  contains a positive proportion of congruence classes of embeddings of  $\mathcal{T}$  in  $\mathbb{F}_q^2$ .*

This result is novel in the sense that [IP19] and [Bri+23] both give trivial results for simplex trees in  $\mathbb{F}_q^2$ . The restriction of  $q \equiv 3 \pmod{4}$  is to rule out the existence of a square root of  $-1$  and the resulting radius zero spheres in  $\mathbb{F}_q^2$ . These cause technical obstructions in obtaining the improved bound on the finite field analogue of the Mattila integral, necessary when applying group action techniques to our large central simplex (as in the discussion after the statement of Theorem 1.3 above). An argument in [Ben+17] was able to work around the  $q \equiv 1 \pmod{4}$  obstructions for embeddings of 1 and 2-simplices in  $\mathbb{F}_q^2$  via explicit analysis of the “bad cases” in this setting. Therefore, with a more careful analysis of the  $q \equiv 1 \pmod{4}$  setting in the case of overconstrained simplices in  $\mathbb{F}_q^2$ , it is believable that Theorem 1.5 should hold for  $q \equiv 1 \pmod{4}$  as well.

It is interesting to note that we obtain the same results as in Theorems 1.3 and 1.5 when we delete a number of edges from some complete graphs in our simplex tree. More precisely, in Section 2, we note that for  $k \geq d$ , a congruence class of an embedding of a complete graph on  $k+1$  vertices (i.e. a  $k$ -simplex) is defined by  $d(k+1) - \binom{d+1}{2}$  pairwise distances rather than all  $\binom{k+1}{2}$  distances. Therefore, given a simplex structure, we may remove up to  $\binom{k+1}{2} - d(k+1) + \binom{d+1}{2}$  edges of any simplex of dimension at least  $d$  and obtain the same bounds as in Theorems 1.3 and 1.5.

In Section 2, we discuss some preliminaries from Fourier analysis and the techniques used in [Ben+17] and [AIM24]. In Section 3, we generalize the methods

used in [AIM24] to isolate terms corresponding to high-dimensional simplices and low-dimensional simplex trees in the sum corresponding to a simplified simplex tree  $\mathcal{T}$ . In Section 4, we present the proof of Theorems 1.3 and 1.5. Given a simplex tree  $\mathcal{T}$ , we recursively build a sum corresponding to  $\mathcal{T}$  and reduce the problem to bounding this sum. Then, we define and prove branch shifting and simplex unbalancing. Finally, we inductively apply these two geometric operations to reduce to a case where we may apply the methods in Section 3.

In Section 5, we discuss partial progress in applying the Hadamard three-lines framework to cycles of simplices. As in Section 4, we are able to reduce to a class of cycles with only one large simplex. However, in bounding the sum corresponding to this simple class of simplex cycles, technical issues arise in trying to separate the terms corresponding to the large simplex and to the small simplices.

In Section 6, we briefly discuss future work on structures of simplices glued together at an edge or a face, as well as a probabilistic approach to the Erdős-Falconer distance problem that may circumvent the technical obstructions in Section 5.

## 2. PRELIMINARIES

**2.1. Analytic preliminaries.** We first state some of the Fourier analytic machinery we need.

**Definition 2.1.** *Suppose  $f: \mathbb{F}_q^d \rightarrow \mathbb{C}$  and suppose  $\chi$  is a nontrivial additive character on  $\mathbb{F}_q$ . Then we define the  $m$ -th Fourier coefficient of  $f$  as*

$$\widehat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-m \cdot x) f(x). \quad (2.1)$$

We also have the usual inverse finite Fourier transform, defined by

$$f(x) = \sum_{m \in \mathbb{F}_q^d} \chi(m \cdot x) \widehat{f}(m). \quad (2.2)$$

**Lemma 2.2** (Parseval). *For  $f$  as above,*

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|^2 = q^{-d} \sum_{x \in \mathbb{F}_q^d} |f(x)|^2. \quad (2.3)$$

For further reading on the finite Fourier transform see, for example, Chapter 7 of [SS03].

We now review some tools we use in bounding sums later on in this paper.

**Lemma 2.3** (Log-convexity of  $L^p$ -norms). *For  $p_0, p_1 \in \mathbb{R}^+$ , define*

$$p_r = \left( \frac{1-r}{p_0} + \frac{r}{p_1} \right)^{-1}. \quad (2.4)$$

*Then, for  $f \in L^{p_0} \cap L^{p_1}$ , we have*

$$\|f\|_{p_r} \leq \|f\|_{p_0}^{1-r} \|f\|_{p_1}^r. \quad (2.5)$$

**Theorem 2.4** (Hadamard three-lines theorem). *Suppose  $f(z) = f(x+iy)$  is bounded, continuous on the strip  $\{z : \operatorname{Re}(z) \in [a, b]\}$ , and holomorphic in the interior of the strip. Put*

$$M(x) = \sup_{y \in \mathbb{R}} |f(x+iy)|. \quad (2.6)$$

*Then  $\log(M(x))$  is convex on  $[a, b]$ . Therefore, for any  $t \in [0, 1]$ , we have*

$$M(ta + (1-t)b) \leq M(a)^t M(b)^{1-t}. \quad (2.7)$$

Throughout this paper, we often apply the Hadamard three-lines Theorem in the setting below, so we state it here for convenience.

**Corollary 2.5.** *Consider a function  $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$  that can be extended to a function from  $\mathbb{C}^2$  to  $\mathbb{C}$ . Suppose that the complex function  $\Psi(z) = \psi(a_1 + z, a_2 - z)$  satisfies the criteria of Lemma 2.4 in the strip  $\{z : \operatorname{Re}(z) \in [-a_1, a_2]\}$ , and for each  $x \in [-a_1, a_2]$ ,  $|\Psi(x+iy)|$  is maximized at  $y = 0$ . Then, we have that*

$$\psi(a_1, a_2) \leq \psi(a_1 + a_2, 0)^{s_1} \psi(0, a_1 + a_2)^{s_2}, \quad (2.8)$$

for  $s_1 = \frac{a_2}{a_1+a_2}$  and  $s_2 = \frac{a_1}{a_1+a_2}$ .

Finally, we have a finite field variant of Taylor's Theorem that is a special case of Lemma 2.1 of [Ben+17].

**Lemma 2.6.** *For any function  $\phi : \mathbb{F}_q^d \rightarrow \mathbb{R}^{\geq 0}$  and any  $n \geq 2$ , we have that*

$$\sum_{x \in \mathbb{F}_q^d} \phi^n(x) \lesssim q^{-d(n-1)} \|\phi\|_1^n + \|\phi\|_\infty^{n-2} \sum_{x \in \mathbb{F}_q^d} \left( \phi(x) - \frac{\|\phi\|_1}{q^d} \right)^2. \quad (2.9)$$

**2.2. Geometric preliminaries.** We define the geometric counting function  $\lambda_\theta$  and highlight some important properties, as in [Ben+17].

**Definition 2.7.** *Given a subset  $E \subset \mathbb{F}_q^d$  and a  $\theta \in O_d(\mathbb{F}_q)$ , the group of orthogonal transformations over  $\mathbb{F}_q^d$ , define the function  $\lambda_\theta : \mathbb{F}_q^d \rightarrow \mathbb{R}$  by*

$$\lambda_\theta(w) = |\{(u, u') \in E^2 : u - \theta u' = w\}|. \quad (2.10)$$

*We also have  $\|\lambda_\theta(w)\|_1 = |E|^2$ ,  $\|\lambda_\theta(w)\|_\infty = |E|$ , and  $\widehat{\lambda}_\theta(0) = \frac{|E|^2}{q^d}$ .*

For all  $d \geq 2$ , we have the following bounds on sums over  $\lambda_\theta$ .

**Proposition 2.8** (Lemmas 4 and 5 of [McD20]). *For  $\lambda_\theta$  as defined in  $\mathbb{F}_q^d$ , we have that for  $|E| \gtrsim q^{\frac{dn-d+1}{n}}$ ,*

$$\sum_{\theta \in O_d(\mathbb{F}_q)} \sum_{w \in \mathbb{F}_q^d} \lambda_\theta^n(w) \lesssim |E|^{2n} q^{-dn + \binom{d+1}{2}}. \quad (2.11)$$

*We also have, independent of the size of  $E$ , that*

$$\sum_{\theta \in O_d(\mathbb{F}_q)} \sum_{w \in \mathbb{F}_q^d} (\lambda_\theta(w) - \widehat{\lambda}_\theta(0))^2 \lesssim q^{\binom{d}{2}+1} |E|^2. \quad (2.12)$$

In  $\mathbb{F}_q^2$ , we have the following result.

**Lemma 2.9.** *For any function  $h : \mathbb{F}_q^2 \rightarrow \mathbb{R}_{\geq 0}$ , for  $q \equiv 3 \pmod{4}$*

$$\|\widehat{hS}_t\|_4 \lesssim q^{-\frac{3}{2}} \|hS_t\|_2. \quad (2.13)$$

See Theorem 4.4 of [Cha+12] for the original proof, or the proof of Lemma 2.4 in [AIM24] which more closely matches our notation. Lemma 2.9 allows us to get better bound on sums of  $\lambda_\theta$  as presented below.

**Proposition 2.10.** *For  $q \equiv 3 \pmod{4}$ ,  $\lambda_\theta$  defined as in (2.10), and  $|E| \gtrsim q^{\frac{4n-4}{2n-1}}$ , we have*

$$\sum_{\theta \in O_2(\mathbb{F}_q)} \sum_{w \in \mathbb{F}_q^2} \lambda_\theta^n(w) \lesssim \frac{|E|^{2n}}{q^{2n-3}}. \quad (2.14)$$

Furthermore, independent of the size of  $E$ , we have

$$\sum_{\theta \in O_2(\mathbb{F}_q)} \sum_{w \in \mathbb{F}_q^2} (\lambda_\theta - \widehat{\lambda}_\theta(0))^2 \lesssim |E|^{\frac{5}{2}} q. \quad (2.15)$$

*Proof.* The second part of the above Proposition is proven in [Ben+17], but more explicitly proven in Lemma 2.4 of [AIM24]. To prove the first part of the above Proposition, apply Lemma 2.6 to the sum over  $w$ , and apply Lemma 2.4 of [AIM24].  $\square$

We now review some definitions and notation about congruence classes.

**Definition 2.11.** *For a graph  $G$  with  $e(G)$  edges, the **congruence class** of  $G$  corresponding to the values  $t_1, \dots, t_{e(G)} \in \mathbb{F}_q$  is the set of embeddings  $\mathcal{E}$  in  $G$  such that for each embedding  $h \in \mathcal{E}$  and for all  $i$ , the  $i$ th edge  $(v_{i,1}, v_{i,2})$  of  $G$  satisfies  $\|h(v_{i,1}) - h(v_{i,2})\| = t_i$ . This forms an equivalence relation on the set of embeddings of  $G$ .*

When  $G$  is a simplex, we can rephrase this definition of congruence classes in terms of group actions. For a  $k$ -simplex  $S = (x_0, \dots, x_k) \in (\mathbb{F}_q^d)^{k+1}$ , the **rigid-motion congruence class** of  $G$  is the set of  $k$ -simplices  $S = (y_0, \dots, y_k) \in (\mathbb{F}_q^d)^{k+1}$  such that there exists a  $\theta \in O_d(\mathbb{F}_q)$  and  $w \in \mathbb{F}_q^d$  such that  $x_i = \theta y_i + w$  for all  $i$ . We now show that these two definitions coincide.

A  $k$ -simplex (more precisely, an embedding of a  $k$ -simplex)  $(x_0, \dots, x_k) \in (\mathbb{F}_q^d)^{k+1}$  is called **nondegenerate** if the vectors  $x_1 - x_0, \dots, x_n - x_0$  are linearly independent or they span  $\mathbb{F}_q^d$ . It is well known that for  $k \leq d$ , two nondegenerate  $k$ -simplices  $(x_0, \dots, x_k)$  and  $(y_0, \dots, y_k)$  in  $\mathbb{F}_q^d$  are congruent in the sense of Definition 2.11 if and only if there exists a  $\theta \in O_d(\mathbb{F}_q)$  and  $w \in \mathbb{F}_q^d$  for which  $x_i = \theta y_i + w$  for all  $i$  (see Lemma 2.11 of [Cha+12]). In the case that  $k \geq d$ , we can form this nondegenerate  $k$ -simplex as the union of nondegenerate  $d$ -simplices sharing a common  $d - 1$ -simplex, i.e.,

$$(x_0, \dots, x_k) = \bigcup_{i \geq d} (x_0, \dots, x_{d-1}, x_i) \quad (2.16)$$

Then for two congruent  $k$ -simplices  $(x_0, \dots, x_k)$  and  $(y_0, \dots, y_k)$  in the sense of Definition 2.11, for  $d \leq i \leq k$  there exists  $\theta_i \in O_d(\mathbb{F}_q), w_i \in \mathbb{F}_q^d$  sending  $(x_0, \dots, x_{d-1}, x_i)$  to  $(y_0, \dots, y_{d-1}, y_i)$ . These  $\theta_i$  must be chosen from the two  $\theta$  sending  $(x_0, \dots, x_{d-1})$  to  $(y_0, \dots, y_{d-1})$  (there are two since this  $\theta$  is unique up to reflection). Therefore, for  $k \geq d$ , every congruence class of  $k$ -simplices contains most  $2^{d-k+1}$  rigid-motion congruence classes of  $k$ -simplices defined by rotations and translations, and every rigid-motion congruence class lies in a single congruence class. We see that rigid-motion congruence classes of simplices are a finer equivalence relation than standard congruence classes, but only by a constant factor. As our results only care about  $E$  containing a positive proportion of congruence classes of embeddings of some graph, we may use these definitions interchangeably for congruence classes of nondegenerate simplices.

**Remark.** In Theorem 7.5 of [Ben+17] for  $k \leq d$  and Theorem 1 of [McD20] for  $k \geq d$ , it is proved that the number of rigid-motion congruence classes of degenerate  $k$ -simplices is negligible compared to the number of rigid-motion congruence classes of all  $k$ -simplices. Therefore, for the rest of the paper, we may assume that all embeddings of simplices are nondegenerate.

We recall the following fact about embeddings of simplices in [McD20].

**Lemma 2.12.** *The number of congruence classes of  $n$ -simplices in  $\mathbb{F}_q^d$  is on the order of*

$$\begin{cases} q^{\binom{n+1}{2}} & n \leq d \\ q^{d(n+1) - \binom{d+1}{2}} & n \geq d. \end{cases} \quad (2.17)$$

We note that these two values agree for  $n = d - 1$  and  $n = d$ .

We establish the following notation.

**Definition 2.13.** *For a simplex  $S$ , let  $c_d(S)$  be the integer such that the number of congruence classes of  $S$  in  $\mathbb{F}_q^d$  is on the order of  $q^{c_d(S)}$ . We write  $c(S)$  when the dimension is clear.*

Finally, we recall a well-known fact about the size of the orthogonal group over  $\mathbb{F}_q$ .

**Lemma 2.14.** *Let  $O_d(\mathbb{F}_q)$  denote the orthogonal group over  $\mathbb{F}_q$ . Then we have that*

$$|O_d(\mathbb{F}_q)| \approx q^{\binom{d}{2}}. \quad (2.18)$$

**2.3. Simplex trees.** We develop notation for simplex trees in order to work with them more easily throughout the paper.

**Definition 2.15.** *We define a **rooted simplex tree**  $(\mathcal{T}, S_0)$  to be a connected simplex tree with a designated root simplex  $S_0$ .*

**Definition 2.16.** For a rooted simplex tree  $(\mathcal{T}, S_0)$ , we define the **depth of a simplex**  $S$  in  $\mathcal{T}$  to be the minimal length  $n$  of a path of simplices in  $\mathcal{T}$  from  $S_0$  to  $S$ . Specifically, for  $i_0, \dots, i_m \in I$  such that  $S_0 = S_{i_0}$  and  $S = S_{i_m}$ , we define

$$n := \min \left\{ m \in \mathbb{N} : \exists S_{i_1}, \dots, S_{i_{m-1}} \in \mathcal{T}, S_{i_{j+1}} \in \mathcal{N}(S_{i_j}) \text{ for all } 0 \leq j \leq m-1 \right\}. \quad (2.19)$$

**Definition 2.17.** For a rooted simplex tree  $(\mathcal{T}, S_0)$ , and a simplex  $S \in \mathcal{T}$ , we call a vertex  $v$  of  $S$  a **parent vertex of  $S$**  if it is shared by a simplex of less depth. Note that each non-root simplex has exactly one parent vertex, while the root has no parent vertex. If  $v$  is shared by a simplex of higher depth, it is called a **child vertex of  $S$** . If  $v$  is not shared by any simplex other than  $S$ , it is called a **free vertex**.

A **free rooted simplex tree**  $(\mathcal{T}, S_0, v_0)$  is a rooted simplex tree with designated free vertex  $v_0$  of  $S_0$ .

**Definition 2.18.** For a simplex  $S$  in a rooted tree  $(\mathcal{T}, S_0)$ , denote  $\mathcal{P}_{\mathcal{T}}(S)$  ( $\mathcal{P}$  for “pruned”) as the free rooted simplex tree consisting of  $S$  and all simplices connected to  $S$  through a path of simplices in  $\mathcal{T}$  of greater depth than  $S$ . We define  $\mathcal{P}_{\mathcal{T}}(S)$  to be rooted at  $S$  with the designated vertex being the parent vertex of  $S$  in  $\mathcal{T}$ . We drop the subscript  $\mathcal{T}$  when it is clear from context.

**Definition 2.19.** For each  $v \in V(S)$  and simplices  $S_{i_1}, \dots, S_{i_n} \in \mathcal{N}(S)$  of greater depth sharing the vertex  $v$  with  $S$ , we define the  $j$ -th **branch** of  $S$  at  $v$  to be the free rooted simplex tree  $\mathcal{P}(S_{i_j})$  with root  $S_{i_j}$  and designated free vertex  $v$ . Let  $\mathcal{B}_{\mathcal{T}}(S, v)$  be the set of branches of  $S$  at  $v$ , where the  $\mathcal{T}$  subscript is dropped when it is clear from context. If a simplex  $S$  has no branches, we call the simplex a **leaf**. If the simplex  $S$  is dimension greater than  $n$  and has branches consisting only of simplices of dimension at most  $n$ , we call the simplex an  **$n$ -leaf**.

It is clear that any simplex tree with a simplex of dimension greater than  $n$  contains an  $n$ -leaf.

Note that choosing a congruence class of each simplex  $S \in \mathcal{T}$  determines a unique congruence class of  $\mathcal{T}$ , since it determines the distance between every pair of vertices corresponding to an edge in an embedding of  $\mathcal{T}$ . Furthermore, every congruence class of  $\mathcal{T}$  uniquely determines the congruence class of each  $S \in \mathcal{T}$ . This immediately gives us following lemma:

**Lemma 2.20.** Given a simplex tree  $\mathcal{T}$  we have that the number of congruence classes of  $\mathcal{T}$  in  $\mathbb{F}_q^d$  is equal to  $q^{\sum_{S \in \mathcal{T}} c(S)}$ . We denote the quantity  $\sum_{S \in \mathcal{T}} c(S)$  as  $c(\mathcal{T})$ .

**Remark.** By the discussion and remark under 2.11, we may assume that in any embedding of  $\mathcal{T}$ , all embeddings of simplices in  $\mathcal{T}$  are nondegenerate.

### 3. FUNCTIONAL SUMS AND USEFUL BOUNDS

In this section, we prove several bounds on sums and products of geometric counting functions that will become relevant in the upcoming sections. Some of the

bounds in this section rely on the results of [Bri+23], who work with a weaker definition of simplex trees than how we have defined them. We recall their definition here:

**Definition 3.1** (Definition 1.11 of [Bri+23]). *We define an  $k$ -weak simplex tree  $T_k$  as the graph obtained from a standard graph-theoretic tree, where each edge is replaced with a  $k$ -simplex containing this edge. We assume  $k \leq \frac{d+1}{2}$ .*

For a  $k$ -weak simplex tree  $T_k$  with a distinguished root vertex  $r$ , we enumerate each of the  $e(T_k) = \ell \binom{k+1}{2}$  edges in  $T_k$ , where  $\ell$  is the number of simplices in  $T_k$ . For  $t_1, \dots, t_{e(T_k)} \in \mathbb{F}_q$  such that  $t_i \neq 0$ , we identify the congruence class of embeddings of  $T_k$  in  $\mathbb{F}_q^d$  by assigning the distance  $t_i$  to the  $i$ th edge. For an  $E \subset \mathbb{F}_q^d$ , we define the function  $f_{T_k, t_1, \dots, t_{e(T_k)}}(x)$  as the number of embeddings  $h : V(T_k) \rightarrow E$  of  $T_k$  in the congruence class defined by  $t_1, \dots, t_{e(T_k)} \in \mathbb{F}_q$ , with the condition that  $h(r) = x$ . In our analysis of  $f_{T_k, t_1, \dots, t_{e(T_k)}}(x)$ , our results will be independent of the  $t_i$  (as long as they are nonzero), so we denote this function simply as  $f_{T_k}(x)$ .

For a  $\theta \in O_d(\mathbb{F}_q)$ , it is useful to define the function  $\Gamma_{\theta, T_k}(w) : \mathbb{F}_q^d \rightarrow \mathbb{R}$  as

$$\Gamma_{\theta, T_k}(w) = \sum_{\substack{x, x' \in \mathbb{F}_q^d \\ x - \theta x' = w}} f_{T_k}(x) f_{T_k}(x'). \quad (3.1)$$

This function counts the number of pairs of congruent embeddings  $h_1, h_2 : V(T_k) \rightarrow E$  such that the transformation given by a rotation by  $\theta$  and a translation by  $w$  sends  $h_1(r)$  to  $h_2(r)$ . We prove some facts about  $f_{T_k}$  and  $\Gamma_{\theta, T_k}$  which are of use to us.

**Lemma 3.2.** *Suppose  $|E| \gtrsim q^{k + \frac{d-1}{2}}$  and let  $T_k$  be a  $k$ -weak simplex tree with  $\ell$  simplices. Then*

$$\begin{aligned} (1) \quad \|f_{T_k}\|_1 &\approx \frac{|E|^{\ell k+1}}{q^{\ell \binom{k+1}{2}}}. \\ (2) \quad \|f_{T_k}\|_2 &\approx \frac{|E|^{\ell k + \frac{1}{2}}}{q^{\ell \binom{k+1}{2}}}. \\ (3) \quad \|\Gamma_{\theta, T_k}\|_1 &\approx \frac{|E|^{2\ell k+2}}{q^{2\ell \binom{k+1}{2}}} \text{ and } \widehat{\Gamma}_{\theta, T_k}(0) \approx \frac{|E|^{2\ell k+2}}{q^{d+2\ell \binom{k+1}{2}}}. \end{aligned}$$

*Proof.* By definition,  $f_{T_k}(x)$  counts the number of embeddings of  $T_k$  rooted at  $x$  in a fixed congruence class. Hence,  $\|f_{T_k}\|_1$  is the total number of trees in this congruence class. From Theorem 1.12 of [Bri+23], this is on the order of  $|E|^{\ell k+1} q^{-\ell \binom{k+1}{2}}$  when  $|E| \gtrsim q^{k + \frac{d-1}{2}}$ .

To find  $\|f_{T_k}\|_2$ , note that  $f_{T_k}^2(x)$  counts the number of pairs of embeddings of  $T_k$  rooted at  $x$  in a fixed congruence class. Equivalently,  $f_{T_k}^2(x)$  counts the number of embeddings of a larger  $k$ -weak simplex tree  $T'_k$  rooted at  $x$ , where  $T'_k$  is obtained by adjoining two copies of  $T_k$  at  $r$ , and denoting  $r$  as the root vertex of  $T'_k$ .  $T_k$  has  $2\ell$

many  $k$ -simplices, so by Theorem 1.12 of [Bri+23], we have that for  $|E| \gtrsim q^{k+\frac{d-1}{2}}$ ,

$$\sum_{x \in \mathbb{F}_q^d} f_{T_k}^2(x) \approx \frac{|E|^{2\ell k+1}}{q^{2\ell \binom{k+1}{2}}}. \quad (3.2)$$

Therefore, taking square roots, we get that

$$\|f_{T_k}\|_2 \approx \frac{|E|^{\ell k + \frac{1}{2}}}{q^{\ell \binom{k+1}{2}}}. \quad (3.3)$$

By the definition of the Fourier transform we have that

$$\widehat{\Gamma}_{\theta, T_k}(m) = q^{-d} \sum_{w \in \mathbb{F}_q^d} \chi(-m \cdot w) \Gamma_{\theta, T_k}(w) \quad (3.4)$$

$$= q^{-d} \sum_{x, x' \in \mathbb{F}_q^d} \chi(-m \cdot (x - \theta x')) f_{T_k}(x) f_{T_k}(x') = q^d \widehat{f_{T_k}}(m) \overline{\widehat{f_{T_k}}(\theta^{-1} m)}. \quad (3.5)$$

We also have that for  $|E| \gtrsim q^{k+\frac{d-1}{2}}$ ,

$$\widehat{f_{T_k}}(0) = q^{-d} \|f_{T_k}\|_1 \approx q^{-d} \left( \frac{|E|^{\ell k+1}}{q^{\ell \binom{k+1}{2}}} \right) = \frac{|E|^{\ell k+1}}{q^{d+\ell \binom{k+1}{2}}}. \quad (3.6)$$

Therefore, we get the following results.

$$\widehat{\Gamma}_{\theta, T_k}(0) = q^d |\widehat{f_{T_k}}(0)|^2 \approx \frac{|E|^{2\ell k+2}}{q^{d+2\ell \binom{k+1}{2}}} \quad (3.7)$$

$$\|\Gamma_{\theta, T_k}\|_1 = q^d \widehat{\Gamma}_{\theta, T_k}(0) \approx \frac{|E|^{2\ell k+2}}{q^{2\ell \binom{k+1}{2}}}. \quad (3.8)$$

□

In the coming sections, we need bounds for sums of the form

$$\sum_{\theta \in O_d(\mathbb{F}_q)} \sum_{w \in \mathbb{F}_q^d} \lambda_{\theta}^n(w) \Gamma_{\theta, T_k}(w). \quad (3.9)$$

This bound looks slightly different in  $\mathbb{F}_q^2$  than in higher dimensions, due to our ability to bound the finite field analogue of the Mattila integral better in two dimensions.

**Lemma 3.3.** *Let  $T$  be a 1-weak simplex tree with  $\ell$  simplices (i.e. a standard graph-theoretic tree with  $\ell$  edges) and let  $E \subset \mathbb{F}_q^2$  for  $q \equiv 3 \pmod{4}$ . If  $|E| \gtrsim q^{\frac{4n}{2n+1}}$  for some  $n \geq 1$ , then we have that*

$$\sum_{\theta, w} \lambda_{\theta}^n(w) \Gamma_{\theta, T}(w) \lesssim \frac{|E|^{2\ell+2n+2}}{q^{2\ell+2n-1}}. \quad (3.10)$$

**Lemma 3.4.** *Let  $T$  be a  $k$ -weak simplex tree with  $\ell$  simplices, for  $1 \leq k < \frac{d+1}{2}$ , and let  $E \subset \mathbb{F}_q^d$ . If  $|E| \gtrsim q^{\max(\frac{dn+1}{n+1}, k + \frac{d-1}{2})}$  for some  $n \geq 1$ , we have that*

$$\sum_{\theta, w} \lambda_{\theta}^n(w) \Gamma_{\theta, T_k}(w) \lesssim \frac{|E|^{2\ell k + 2n + 2}}{q^{2\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}}. \quad (3.11)$$

The proofs of these two Lemmas are contained in the subsequent subsections.

**3.1. Proof of Lemma 3.3.** We first prove a useful bound, which is a generalization of Lemma 3.10 of [AIM24]. Recall that in the setting of Lemma 3.3 we have  $d = 2$ , so our 1-weak simplex tree  $T_k$  is just a standard graph-theoretic tree  $T$ .

**Lemma 3.5.** *For  $q \equiv 3 \pmod{4}$ ,  $E \subset \mathbb{F}_q^2$  and  $|E| \gtrsim q^{3/2}$ ,*

$$\sum_{\theta, w} \left( \Gamma_{\theta, T}(w) - \widehat{\Gamma}_{\theta, T}(0) \right)^2 \lesssim \frac{|E|^{4\ell + \frac{5}{2}}}{q^{4\ell - 1}}. \quad (3.12)$$

*Proof.* We first apply Parseval's theorem.

$$\sum_{\theta, w} \left( \Gamma_{\theta, T}(w) - \widehat{\Gamma}_{\theta, T}(0) \right)^2 = q^2 \sum_{\theta} \sum_{m \neq 0} \left| \widehat{\Gamma}_{\theta, T}(m) \right|^2. \quad (3.13)$$

We note that

$$\widehat{\Gamma}_{\theta, T}(m) = q^{-2} \sum_w \chi(-m \cdot w) \Gamma_{\theta, T}(w) = q^{-2} \sum_{u, u' \in \mathbb{F}_q^2} \chi(-m \cdot (u - \theta u')) f_T(u) f_T(u') \quad (3.14)$$

$$= q^2 \widehat{f}_T(m) \overline{\widehat{f}_T(\theta^{-1} m)}. \quad (3.15)$$

Plugging the expression (3.15) for  $\widehat{\Gamma}_{\theta, T}(m)$  into (3.13) gives us

$$\sum_{\theta, w} \left( \Gamma_{\theta, T}(w) - \widehat{\Gamma}_{\theta, T}(0) \right)^2 = q^6 \sum_{\theta} \sum_{m \neq 0} |\widehat{f}_T(m)|^2 |\widehat{f}_T(\theta^{-1} m)|^2. \quad (3.16)$$

We sum first in  $\theta$ . Defining  $\|m\| = t$  and  $h(m) = \overline{\widehat{f}_T(m)}$ , we get that

$$\sum_{\theta} |\widehat{f}_T(\theta^{-1} m)|^2 \approx \sum_{\|\ell\|=t} |\widehat{f}_T(\ell)|^2 = \sum_{\ell} \widehat{f}_T(\ell) S_t(\ell) h(\ell) \quad (3.17)$$

$$= q^{-2} \sum_{\ell} \sum_w \chi(-\ell \cdot w) f_T(w) S_t(\ell) h(\ell) \quad (3.18)$$

$$= \sum_{\ell} f_T(\ell) \widehat{h S}_t(\ell) = \|f_T\|_{4/3} \|\widehat{h S}_t\|_4, \quad (3.19)$$

where the last equality follows from an application of Hölder's inequality. We recall that by Lemma 2.9,  $\|\widehat{h S}_t\|_4 \lesssim q^{-3/2} \|h S_t\|_2$  and we use the log-convexity of  $L^p$  norms to bound  $\|f_T\|_{4/3}$ .

$$\|f_T\|_{4/3} \leq (\|f_T\|_1)^{1/2}(\|f_T\|_2)^{1/2}. \quad (3.20)$$

These give

$$2\|hS_t\|_2^2 = \sum_{\theta} |\widehat{f_T}(\theta^{-1}m)|^2 \lesssim q^{-3/2}\|hS_t\|_2(\|f_T\|_1)^{1/2}(\|f_T\|_2)^{1/2}, \quad (3.21)$$

from which we get that

$$\sum_{\theta} |\widehat{f_T}(\theta^{-1}m)|^2 \lesssim q^{-3}\|f_T\|_1\|f_T\|_2. \quad (3.22)$$

We calculate the remaining sum over  $m$  in (3.16).

$$\sum_{m \in \mathbb{F}_q^2} |\widehat{f_T}(m)|^2 = q^{-2} \sum_{x \in \mathbb{F}_q^2} f(x)^2 = q^{-2}\|f_T\|_2^2. \quad (3.23)$$

Putting this all together, we conclude that

$$\sum_{\theta, w} \left( \Gamma_{\theta, T}(w) - \widehat{\Gamma}_{\theta, T}(0) \right)^2 \lesssim q^6 \cdot q^{-2}\|f_T\|_2^2 \cdot q^{-3}\|f_T\|_1\|f_T\|_2 = q\|f_T\|_1\|f_T\|_2^3. \quad (3.24)$$

For  $|E| \gtrsim q^{3/2}$ , we may plug in our bounds for  $\|f_T\|_1$  and  $\|f_T\|_2$  from Lemma 3.2, completing the proof.  $\square$

We are now ready to prove Lemma 3.3 by induction on  $n$ . Let  $A_n = \sum_{\theta, w} \lambda_{\theta}^n(w) \Gamma_{\theta, T}(w)$ . We first find a bound on  $A_1$ . By Cauchy-Schwarz, we have that

$$\sum_{\theta, w} \lambda_{\theta}(w) \Gamma_{\theta, T}(w) \leq \left( \sum_{\theta, w} \lambda_{\theta}^2(w) \right)^{1/2} \left( \sum_{\theta, w} \Gamma_{\theta, T}^2(w) \right)^{1/2}. \quad (3.25)$$

By Proposition 2.10, we have that  $\sum_{\theta, w} \lambda_{\theta}^2(w) \lesssim |E|^4 q^{-1}$  whenever  $|E| \gtrsim q^{4/3}$ . To bound  $\sum_{\theta, w} \Gamma_{\theta, T}^2(w)$ , we apply Lemma 2.6 and plug in our bound from Lemma 3.5: for  $|E| \gtrsim q^{3/2}$ , we obtain

$$\sum_{\theta, w} \Gamma_{\theta, T}^2(w) \lesssim \sum_{\theta} \frac{|E|^{4\ell+4}}{q^{4\ell+2}} + \sum_{\theta, w} \left( \Gamma_{\theta, T}(w) - \widehat{\Gamma}_{\theta, T}(0) \right)^2 \quad (3.26)$$

$$\lesssim \frac{|E|^{4\ell+4}}{q^{4\ell+1}} + \frac{|E|^{4\ell+\frac{5}{2}}}{q^{4\ell-1}}. \quad (3.27)$$

We see that the first term in (3.27) dominates the second term at  $|E| \gtrsim q^{4/3}$ . Putting everything together, we have for  $|E| \gtrsim q^{4/3}$ ,

$$\sum_{\theta, w} \lambda_{\theta}(w) \Gamma_{\theta, T}(w) \lesssim \frac{|E|^{2\ell+4}}{q^{2\ell+1}}. \quad (3.28)$$

This completes the base case. Now assume that  $n \geq 2$ , and assume the inductive hypothesis that for  $|E| \gtrsim q^{4(n-1)/(2(n-1)+1)}$ ,

$$A_{n-1} \lesssim \frac{|E|^{2\ell+2(n-1)+2}}{q^{2\ell+2(n-1)-1}} \quad (3.29)$$

Since  $\lambda_\theta^n(w)\Gamma_{\theta,T}(w)$  is positive by definition and  $\widehat{\lambda}_\theta(0) = |E|^2q^{-2}$ , we may write

$$\begin{aligned} \sum_{\theta,w} \lambda_\theta^n(w)\Gamma_{\theta,T}(w) &\leq \left| \sum_{\theta,w} \lambda_\theta^{n-1}(w) \left( \lambda_\theta(w) - \frac{|E|^2}{q^2} \right) (\Gamma_{\theta,T}(w) - \widehat{\Gamma}_{\theta,T}(0)) \right| \\ &\quad + \sum_{\theta,w} \lambda_\theta^n \widehat{\Gamma}_{\theta,T}(0) + \frac{|E|^2}{q^2} A_{n-1} \end{aligned} \quad (3.30)$$

$$=: I + II + III. \quad (3.31)$$

We bound  $I$  by pulling absolute values inside the sum, taking the supremum bound on  $\lambda_\theta^{n-1}(w)$ , and applying Cauchy-Schwarz to the remaining terms, as follows:

$$I \leq |E|^{n-1} \left( \sum_{\theta,w} (\lambda_\theta(w) - \widehat{\lambda}_\theta(0))^2 \right)^{1/2} \left( \sum_{\theta,w} (\Gamma_{\theta,T}(w) - \widehat{\Gamma}_{\theta,T}(0))^2 \right)^{1/2} \quad (3.32)$$

$$\lesssim \frac{|E|^{2\ell+\frac{3}{2}+n}}{q^{2\ell-1}}. \quad (3.33)$$

The bound above follows from plugging in Proposition 2.10 and Lemma 3.5. For  $|E| \gtrsim q^{4n/(2n+1)}$ , we have that

$$\frac{|E|^{2\ell+n+\frac{3}{2}}}{q^{2\ell-1}} \lesssim \frac{|E|^{2\ell+2n+2}}{q^{2\ell+2n-1}}. \quad (3.34)$$

From Proposition 2.10 and Lemma 3.2, we get that for  $|E| \gtrsim q^{4(n-1)/(2(n-1)+1)}$ ,

$$II = \frac{|E|^{2\ell+2}}{q^{2\ell+2}} \sum_{\theta,w} \lambda_\theta^n(w) \lesssim \frac{|E|^{2\ell+2n+2}}{q^{2\ell+2n-1}}. \quad (3.35)$$

Finally, we note that for  $|E| \gtrsim q^{4(n-1)/(2(n-1)+1)}$ ,

$$III = \frac{|E|^2}{q^2} A_{n-1} \lesssim \frac{|E|^{2\ell+2n+2}}{q^{2\ell+2n-1}}. \quad (3.36)$$

Combining these bounds completes the proof.

**3.2. Proof of Lemma 3.4.** Before proceeding to the proof of Lemma 3.4, we prove a  $d$ -dimensional analogue of the bound in Lemma 3.5. This is essentially a generalization of Lemma 4 in [McD20].

**Lemma 3.6.** For  $E \subset \mathbb{F}_q^d$  and  $|E| \gtrsim q^{k+(d-1)/2}$

$$\sum_{\theta,w} (\Gamma_{\theta,T_k}(w) - \widehat{\Gamma}_{\theta,T_k}(0))^2 \lesssim |E|^{4\ell k+2} q^{d+\binom{d-1}{2}-4\ell\binom{k+1}{2}}.$$

*Proof.* We see that

$$\sum_{\theta, w} (\Gamma_{\theta, T_k}(w) - \widehat{\Gamma}_{\theta, T_k}(0))^2 = q^d \sum_{\theta, m \neq 0} |\Gamma_{\theta, T_k}(m)|^2 \quad (3.37)$$

$$= q^{3d} \sum_{\theta, m \neq 0} |\widehat{f}_{T_k}(m)|^2 |\widehat{f}_{T_k}(\theta^{-1}m)|^2 \quad (3.38)$$

$$= q^{3d} \sum_{m \neq 0} |\widehat{f}_{T_k}(m)|^2 \sum_{\theta} |\widehat{f}_{T_k}(\theta^{-1}m)|^2 \quad (3.39)$$

$$= q^{3d} \sum_t \sum_{\substack{m \neq 0 \\ \|m\|=t}} |\widehat{f}_{T_k}(m)|^2 \sum_{\substack{n \neq 0 \\ \|n\|=t}} |\widehat{f}_{T_k}(n)|^2 \sum_{\theta m=n} 1. \quad (3.40)$$

Note that for  $\theta, \phi \in O_d(\mathbb{F}_q)$ , we have that  $\theta m = n$  and  $\phi m = n$  if and only if  $\phi^{-1}\theta$  fixes  $m$ . Therefore, the sum over  $\theta$  occurs the size of the stabilizer of  $m$  many times, which is just  $|O_{d-1}(\mathbb{F}_q)|$  since  $m, n \neq 0$ . Therefore, we may write (3.40) as

$$q^{3d} |O_{d-1}(\mathbb{F}_q)| \sum_t \left( \sum_{\substack{m \neq 0 \\ \|m\|=t}} |\widehat{f}_{T_k}(m)|^2 \right)^2. \quad (3.41)$$

We write the square of sums as the product of two sums. We dominate one of these sums by dropping all conditions on  $m$  by positivity:

$$\sum_{\substack{m \neq 0 \\ \|m\|=t}} |\widehat{f}_{T_k}(m)|^2 \leq \sum_m |\widehat{f}_{T_k}(m)|^2 = q^{-d} \|f_{T_k}\|_2^2, \quad (3.42)$$

where the last equality follows from Parseval. We dominate the rest of (3.41) by removing the non-zero condition:

$$\sum_t \sum_{\substack{m \neq 0 \\ \|m\|=t}} |\widehat{f}_{T_k}(m)|^2 \leq \sum_m |\widehat{f}_{T_k}(m)|^2 = q^{-d} \|f_{T_k}\|_2^2. \quad (3.43)$$

All in all, we get that

$$\sum_{\theta, w} (\Gamma_{\theta, T_k}(w) - \widehat{\Gamma}_{\theta, T_k}(0))^2 \leq q^d |O_{d-1}(\mathbb{F}_q)| \|f_{T_k}\|_2^4. \quad (3.44)$$

Plugging in  $\|f_{T_k}\|_2$  from Lemma 3.2 and the size of  $|O_{d-1}(\mathbb{F}_q)|$  from Proposition 2.14 gives the desired result.  $\square$

*Proof of Lemma 3.4.* As in the proof of Lemma 3.3, we proceed by induction on  $n$ . Denote by  $A_n$  the sum  $\sum_{\theta, w} \lambda_{\theta}^n(w) \Gamma_{\theta, T_k}(w)$ . We first obtain a bound for  $A_1$ . By Cauchy Schwarz, we have that

$$\sum_{\theta, w} \lambda_{\theta}(w) \Gamma_{\theta, T_k}(w) \leq \left( \sum_{\theta, w} \lambda_{\theta}^2(w) \right)^{1/2} \left( \sum_{\theta, w} \Gamma_{\theta, T_k}^2(w) \right)^{1/2}. \quad (3.45)$$

Applying Lemma 2.6 to the  $\Gamma$  sum and plugging in Lemma 3.6 for  $|E| \gtrsim q^{k+(d-1)/2}$  gives us that

$$\sum_{\theta, w} \Gamma_{\theta, T_k}^2(w) \lesssim q^{\binom{d}{2}} \frac{|E|^{4\ell k+4}}{q^{4\ell \binom{k+1}{2}+d}} + \sum_{\theta, w} (\Gamma_{\theta, T_k}(w) - \widehat{\Gamma}_{\theta, T_k}(0))^2 \quad (3.46)$$

$$\lesssim q^{\binom{d}{2}} \frac{|E|^{4\ell k+4}}{q^{4\ell \binom{k+1}{2}+d}} + |E|^{4\ell k+2} q^{d+\binom{d-1}{2}-4\ell \binom{k+1}{2}}. \quad (3.47)$$

By a straightforward computation, we see that the first term dominates in (3.47) for  $|E| \gtrsim q^{(d+1)/2}$ . After plugging in the appropriate bounds from Proposition 2.8, we conclude that for  $|E| \gtrsim q^{\max((d+1)/2, k+(d-1)/2)}$ ,

$$A_1 \lesssim \left( |E|^4 q^{\binom{d+1}{2}-2d} \right)^{1/2} \left( \frac{|E|^{4\ell k+4}}{q^{4\ell \binom{k+1}{2}+d}} q^{\binom{d}{2}} \right)^{1/2} = |E|^{2\ell k+4} q^{\frac{d^2-3d-4\ell \binom{k+1}{2}}{2}} \quad (3.48)$$

$$= \frac{|E|^{2\ell k+4}}{q^{2\ell \binom{k+1}{2}+2d-\binom{d+1}{2}}}. \quad (3.49)$$

This concludes the base case. Now, for  $n \geq 2$  and  $|E| \gtrsim q^{\max(\frac{d(n-1)+1}{n}, k+\frac{d-1}{2})}$ , assume that

$$A_{n-1} \lesssim \frac{|E|^{2\ell k+2(n-1)+2}}{q^{2\ell \binom{k+1}{2}+d((n-1)+1)-\binom{d+1}{2}}} \quad (3.50)$$

We break up  $A_n$  similarly as before.

$$\begin{aligned} A_n &= \sum_{\theta, w} \lambda_{\theta}^n(w) \Gamma_{\theta, T_k}(w) \leq \left| \sum_{\theta, w} \lambda_{\theta}^{n-1}(w) \left( \lambda_{\theta}(w) - \frac{|E|^2}{q^d} \right) (\Gamma_{\theta, T_k}(w) - \widehat{\Gamma}_{\theta, T_k}(0)) \right| \\ &\quad + \sum_{w, \theta} \lambda_{\theta}^n(w) \widehat{\Gamma}_{\theta, T_k}(0) + \frac{|E|^2}{q^d} A_{n-1} \end{aligned} \quad (3.51)$$

$$=: I + II + III. \quad (3.52)$$

We bound each term individually. To bound  $I$ , we pull absolute values inside the sum, take the supremum of  $\lambda_{\theta}^{n-1}$ , and apply Cauchy Schwarz to the remaining terms to obtain

$$I \leq |E|^{n-1} \left( \sum_{\theta, w} \left( \lambda_{\theta}(w) - \frac{|E|^2}{q^d} \right)^2 \right)^{1/2} \left( \sum_{\theta, w} (\Gamma_{\theta, T_k}(w) - \widehat{\Gamma}_{\theta, T_k}(0))^2 \right)^{1/2}. \quad (3.53)$$

By Proposition 2.8, we have that  $\sum_{\theta, w} \left( \lambda_{\theta}(w) - \frac{|E|^2}{q^d} \right)^2 \lesssim |E|^2 q^{\binom{d}{2}+1}$ . Bounding the  $\Gamma$  sum with Lemma 3.6 for  $|E| \gtrsim q^{k+(d-1)/2}$ , taking square roots, and simplifying gives us for  $|E| \gtrsim q^{k+(d-1)/2}$ ,

$$I \leq |E|^{2\ell k+n+1} q^{\frac{d^2-d+2}{2}-2\ell \binom{k+1}{2}}. \quad (3.54)$$

By a straightforward computation, we see that for  $|E| \gtrsim q^{\frac{dn+1}{n+1}}$ ,

$$I \lesssim \frac{|E|^{2\ell k+2n+2}}{q^{2\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}}. \quad (3.55)$$

Next, we bound  $II$ . From Proposition 2.8, when  $|E| \gtrsim q^{\frac{dn-d+1}{n}}$ , we have  $\lambda_{\theta}^n(x) \leq |E|^{2n} q^{-dn + \binom{d+1}{2}}$ . Plugging this in to our expression for  $II$  and using our bounds from 3.2 gives us that

$$II \lesssim \frac{|E|^{2\ell k+2n+2}}{q^{2\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}}. \quad (3.56)$$

Finally, the bound for  $III$  quickly follows from the inductive hypothesis for  $|E| \gtrsim q^{\max(\frac{d(n-1)+1}{n}, k + \frac{d-1}{2})}$ :

$$III \lesssim \frac{|E|^2}{q^d} \frac{|E|^{2\ell k+2(n-1)+2}}{q^{2\ell \binom{k+1}{2} + d(n) - \binom{d+1}{2}}} = \frac{|E|^{2\ell k+2n+2}}{q^{2\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}}. \quad (3.57)$$

□

#### 4. PROOF OF MAIN THEOREMS

In this section, we will prove Theorem 1.3, 1.4, and 1.5. Recall from the discussion and remarks under 2.11 and 2.20 that for an embedding  $h$  of a simplex tree  $\mathcal{T}$ , we may assume that the embeddings of all simplices  $S \in \mathcal{T}$  (obtained by restricting the domain of  $h$  to vertices of  $S$ ) are nondegenerate.

**4.1. The setup.** For a given simplex tree  $\mathcal{T}$ , we would like to find an  $s$  such that  $|E| \gtrsim q^s$  implies that  $E$  contains a positive proportion of congruence classes of  $\mathcal{T}$  in  $\mathbb{F}_q^d$ . Denote  $\mathbb{D}$  as the set of congruence classes of  $\mathcal{T}$ . For  $\delta \in \mathbb{D}$  ( $\delta$  a congruence class of  $\mathcal{T}$ ), we define  $\nu_{\mathcal{T}}(\delta)$  as the number of embeddings of  $\mathcal{T}$  in  $E$  that are in the congruence class  $\delta$ . We denote by  $\Delta(E)$  the size of the support of  $\nu_{\mathcal{T}}$ , or in other words, the number of congruence classes of  $\mathcal{T}$  in  $E$ .

By an application of Cauchy-Schwarz, we have that

$$\Delta(E) \geq \frac{(\sum_{\delta \in \mathbb{D}} \nu_{\mathcal{T}}(\delta))^2}{\sum_{\delta \in \mathbb{D}} \nu_{\mathcal{T}}^2(\delta)}. \quad (4.1)$$

The sum in the numerator runs over all possible  $|V(\mathcal{T})|$ -tuples in  $E$  (where  $V(\mathcal{T})$  denotes the vertex set of  $\mathcal{T}$ ). Hence we have that

$$\Delta(E) \gtrsim \frac{|E|^{2|V(\mathcal{T})|}}{\sum_{\delta \in \mathbb{D}} \nu_{\mathcal{T}}^2(\delta)}. \quad (4.2)$$

To obtain a lower bound on  $\Delta(E)$ , it suffices to get an upper bound on  $\sum_{\delta \in \mathbb{D}} \nu_{\mathcal{T}}^2(\delta)$ . More precisely, to prove Theorems 1.3 and 1.5, we need to show that for the  $s$  in the

theorem statements,  $\Delta(E)$  is on the order of  $q^{c(\mathcal{T})}$  whenever  $|E| \gtrsim q^s$ . By the setup above, we see that

$$\sum_{\delta \in \mathbb{D}} v_{\mathcal{T}}^2(\delta) \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}} \implies \Delta(E) \gtrsim q^{c(\mathcal{T})}. \quad (4.3)$$

Our first goal is to find an expression of a more workable form for  $\sum_{\delta \in \mathbb{D}} v_{\mathcal{T}}^2(\delta)$ , the number of pairs of congruent embeddings of  $\mathcal{T}$ . We build this expression recursively, by first counting the number of pairs of congruent embeddings of simplices, then sub-simplex trees, and finally the full simplex tree  $\mathcal{T}$ .

For  $u, u' \in E$  and an  $n$ -simplex  $S$ , we begin by finding an expression for the number of pairs of congruent embeddings of  $S$  in  $E$  based at  $u$  and  $u'$ , i.e. congruent embeddings  $h_1, h_2 : V(S) \rightarrow E$  of  $n$ -simplices into  $E$  such that  $h_1(v_0) = u$  and  $h_2(v_0) = u'$ . The following exposition follows [Ben+17] closely, but our notation differs and these ideas will be developed further later, so we present it here.

We define  $\lambda_{\theta}(w)$  as in Definition 2.7. This is the number of pairs  $x, y \in E$  such that the ‘rigid motion’  $\rho(\theta, w)$ , consisting of a rotation by  $\theta \in O_d(\mathbb{F}_q)$  followed by a translation by  $w$ , sends  $y$  to  $x$ . By the discussion under 2.11, two nondegenerate  $n$ -simplices  $S_x = (x_0, \dots, x_n)$  and  $S_y = (y_0, \dots, y_n)$  with vertices in  $\mathbb{F}_q^d$  are congruent if and only if there exists a rigid motion  $\rho(\theta, w)$  such that  $x_i = \rho(\theta, w)y_i$  for all  $i$ . Therefore, we see that  $\lambda_{\theta}^{n+1}(w)$  counts the number of pairs of congruent  $n$ -simplices sent to each other by the transformation  $\rho(\theta, w)$  (as we need to find  $n + 1$  pairs of points such that  $\rho(\theta, w)$  sends the first element of each pair to the second). For the rest of this section, when the context is clear, a simplex  $S$  living in  $\mathbb{F}_q^d$  is actually an embedding of a simplex into  $\mathbb{F}_q^d$ .

For  $u, u' \in \mathbb{F}_q^d$ , we also frequently make use of the term  $\lambda_{\theta}(u - \theta u')$ , which counts pairs  $x, x'$  such that the transformation given by a rotation by  $\theta$  and translating the image of  $u'$  to  $u$ , also transforms  $x'$  to  $x$ . Therefore,  $\lambda_{\theta}^n(u - \theta u')$  counts the number of pairs of embeddings of  $n$ -simplices, one of which contains  $u$  as a vertex and the other of which contains  $u'$ , that are sent to each other by first rotating by  $\theta$  and then translating the image of  $u'$  to  $u$  (the exponent is  $n$  and not  $n + 1$  as  $u, u'$  are already fixed).

For  $\theta, \phi \in O_d(\mathbb{F}_q)$  and  $n$ -simplices  $S_x = (x_0, \dots, x_n)$  and  $S_y = (y_0, \dots, y_n)$  the transformations  $\rho(\theta, x_0 - \theta y_0)$  and  $\rho(\phi, x_0 - \phi y_0)$  both send  $y_i$  to  $x_i$  for all  $i$  if and only if  $\theta$  and  $\phi$  differ by an element in the stabilizer of the  $n$ -simplex  $S_{y-y_0} = (0, y_1 - y_0, \dots, y_n - y_0)$ . Since for congruent  $n$ -simplices  $S_x$  and  $S_y$ , the corresponding pinned simplices  $S_{x-x_0}$  and  $S_{y-y_0}$  have conjugate stabilizer groups, we may define  $\text{Stab}(S_x)$  to be the common size of the stabilizer  $S_{y-y_0}$  of simplices  $S_y$  congruent to  $S_x$ . We see that for  $u, u' \in E \subset \mathbb{F}_q^d$ , the sum

$$\sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_{\theta}^n(u - \theta u') \quad (4.4)$$

counts each pair of congruent embeddings of  $n$ -simplices  $(S_x, S_y)$  in  $E$  exactly  $\text{Stab}(S_x)$  many times. Let  $\text{Stab}(n)$  denote the minimum size of the stabilizer of a  $n$ -simplex

in  $\mathbb{F}_q^d$ . We have the following fact from [Ben+17] for  $n \leq d$ , which we easily extend to  $n \geq d$ :

**Proposition 4.1.** *The minimum size of the stabilizer of a dimension  $n$  simplex is*

$$\text{Stab}(n) \approx \begin{cases} |O_{d-n}(\mathbb{F}_q)| \approx q^{\binom{d-n}{2}} & n < d \\ 1 & n \geq d. \end{cases}$$

The case for  $n \geq d$  follows trivially from the fact that a simplex  $S_x = (x_0, \dots, x_n)$  of dimension  $d$  or larger, the vectors  $x_1 - x_0, \dots, x_n - x_0$  may span all of  $\mathbb{F}_q^d$  hence the stabilizer is minimally 1.

We note that since we can assume that all simplices are nondegenerate, the stabilizers of our simplices are always of minimal size. Therefore, the expression

$$D_S(u, u') = \frac{1}{\text{Stab}(\dim(S))} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_\theta^{\dim(S)}(u - \theta u') \quad (4.5)$$

counts each pair of congruent embeddings of  $S$  in  $E$  based at  $u$  and  $u'$  at on the order of one time, since without the stabilizer term out front we count each pair of congruent embeddings on the order of  $\text{Stab}(\dim(S))$  times.

We now use the above methods to obtain an expression for the number of pairs of congruent embeddings  $h_1, h_2$  of a free rooted simplex tree  $(\mathcal{T}, S, v_0)$  based at  $u$  and  $u'$  (i.e.  $h_1(v_0) = u$  and  $h_2(v_0) = u'$ ).

**Proposition 4.2.** *Suppose we have a free rooted simplex tree  $(\mathcal{T}, S, v_0)$ , for  $S \in \mathcal{T}$  and  $v_0 \in S$  a free vertex. Define  $D_{(\mathcal{T}, S, v_0)} : E^2 \rightarrow \mathbb{R}$  recursively as*

$$D_{(\mathcal{T}, S, v_0)}(u, u') = \frac{1}{\text{Stab}(\dim(S))} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{\substack{v \in V(S) \\ v \neq v_0}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \prod_{B \in \mathcal{B}(S, v)} D_B(x, x').$$

Then for each  $u, u' \in E$ ,  $D_{(\mathcal{T}, S, v_0)}(u, u')$  is on the order of the number of pairs of congruent embeddings of  $(\mathcal{T}, S, v_0)$  in  $E$  based at  $u$  and  $u'$ .

*Proof.* If  $(\mathcal{T}, S, v_0)$  is a simplex, then  $|\mathcal{B}(S, v)| = 0$ . Defining the empty product over  $\mathcal{B}(S, v) = \emptyset$  to be 1, we see that  $D_{(\mathcal{T}, S, v_0)}(u, u')$  reduces to  $D_S(u, u')$ , and by the discussion above,  $D_S(u, u')$  is on the order of the number of pairs of congruent embeddings of  $(\mathcal{T}, S, v_0)$  based at  $u, u'$ .

We now proceed by induction on the number of simplices in  $(\mathcal{T}, S, v_0)$ . We can obtain an expression on the order of the number of pairs of congruent embeddings of  $(\mathcal{T}, S, v_0)$  in  $E$  based at  $u$  and  $u'$  as follows:

- (1) Fix a  $\theta$ , and consider all embeddings of simplices  $h_1, h_2 : V(S) \rightarrow E$  based at  $u, u'$  such that the rotation  $\theta$  followed by the translation sending  $\theta u'$  to  $u$  also sends  $h_2(v)$  to  $h_1(v)$  for all  $v \in V(S) \setminus v_0$ .
- (2) For each such fixed pair of embeddings of  $S$ , count the number of ways this pair  $h_1, h_2$  can be extended to a pair of congruent embeddings of  $(\mathcal{T}, S, v_0)$  in  $E$ .

(3) Sum over all  $\theta$  and divide the final sum by  $\text{Stab}(\dim(S))$ .

First, we explain why

$$\prod_{\substack{v \in V(S) \\ v \neq v_0}} \sum_{\substack{x, x' \in E \\ x - \theta x' = u - \theta u'}} \prod_{B \in \mathcal{B}(S, v)} D_B(x, x') \quad (4.6)$$

corresponds to steps 1 and 2. For notational convenience, let  $\dim(S) = n$ , so that  $S$  has  $n+1$  vertices. For a fixed  $\theta$ , we consider all congruent embeddings  $h_1, h_2 : V(S) \rightarrow E$  such that the transformation  $\rho(\theta, u - \theta u')$  sends  $h_2(v)$  to  $h_1(v)$  for all  $v \in V(S)$ . Each pair of these congruent embeddings  $h_1, h_2$  corresponds to  $n$  pairs of points  $(a_i, b_i) \in \mathbb{F}_q^2$  (one pair for each  $v \neq v_0$ , as  $u$  and  $u'$  are already fixed) such that  $\rho(\theta, u - \theta u')b_i = a_i$ . For a fixed vertex  $v \in V(S) \setminus v_0$ , the sum over  $x, x' \in E$  ranges over all possible candidates for  $h_1(v), h_2(v)$  with respect to a fixed transformation. For each such  $x, x'$  and a fixed branch  $B$  of  $(\mathcal{T}, S, v_0)$  at the vertex  $v$ , by the inductive hypothesis,  $D_B(x, x')$  is on the order of the number of pairs of congruent embeddings of  $B$  in  $E$  based at  $x$  and  $x'$ . The product over all branches  $B$  at the vertex  $v$  is on the order of the number of pairs of congruent embeddings based at  $x$  and  $x'$  of the structure obtained by joining all branches in  $\mathcal{B}(S, v)$  at  $v$ . Taking the product over all vertices  $v \in V(S)$  distinct from  $v_0$  yields an expression on the order of the number of pairs of congruent embeddings of  $(\mathcal{T}, S, v_0)$  extending  $h_1, h_2$ .

We now argue step 3. Fixing a pair of based congruent embeddings  $h_1, h_2 : V(S) \rightarrow E$ , of  $S$ , for each  $\phi \in \text{Stab}(h_2(S))$  and each  $\theta$  such that  $\rho(\theta, u - \theta u')$  sends  $h_2(S)$  to  $h_1(S)$ , we also have that  $\rho(\theta\phi, u - \theta\phi u')$  sends  $h_2(S)$  to  $h_1(S)$ . Therefore, once we sum over all  $\theta$ , we overcount each pair of embeddings of  $S$  on the order of  $\text{Stab}(h_2(S)) = \text{Stab}(\dim(S))$  many times (as  $h_2(S)$  is nondegenerate). Therefore, Steps 1 and 2 overcount each pair of congruent embeddings of  $\mathcal{T}$  by the same factor, and we divide by  $\text{Stab}(\dim(S))$  to complete the proof of the Proposition.  $\square$

Finally, we construct an expression for the number of pairs of congruent embeddings of a general rooted simplex tree  $(\mathcal{T}, S_0)$ .

**Proposition 4.3.** *Consider a rooted simplex tree  $(\mathcal{T}, S_0)$ , and define*

$$R_{(\mathcal{T}, S_0)} = \frac{1}{\text{Stab}(\dim(S_0))} \sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{v \in V(S_0)} \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \prod_{B \in \mathcal{B}(S_0, v)} D_B(x, x').$$

Then we have that

$$\sum_{\delta \in \mathbb{D}} v_{\mathcal{T}}^2(\delta) \approx R_{(\mathcal{T}, S_0)}.$$

*Proof.* Note that  $\sum_{\delta \in \mathbb{D}} v_{\mathcal{T}}^2(\delta)$  counts the number of pairs of congruent embeddings of  $(\mathcal{T}, S_0)$  in  $E$ . We can obtain an expression on the order of this quantity as follows:

- (1) Fix a  $\theta, w$ , and consider all embeddings of simplices  $h_1, h_2 : V(S_0) \rightarrow E$  such that  $h_1(S_0) = \rho(\theta, w)h_2(S_0)$ .

- (2) For each fixed pair of congruent embeddings of the root  $S_0$  corresponding to  $\rho(\theta, w)$ , count the number of ways this pair of embeddings  $h_1, h_2$  can be extended to a pair of congruent embeddings of  $(\mathcal{T}, S_0)$  in  $E$ .
- (3) Sum over all transformations  $\rho(\theta, w)$ , and divide our final sum by  $\text{Stab}(\dim(S_0))$ .

The quantity

$$\prod_{v \in V(S_0)} \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \prod_{B \in \mathcal{B}(S_0, v)} D_B(x, x') \quad (4.7)$$

corresponds to steps 1 and 2. For notational convenience, let  $\dim(S_0) = n$ . For a fixed  $\theta$ , we consider all congruent embeddings  $h_1, h_2 : V(S) \rightarrow E$  such that the transformation  $\rho(\theta, w)$  sends  $h_2(v)$  to  $h_1(v)$  for all  $v \in V(S)$ . Each pair of these embeddings  $h_1, h_2 : V(S) \rightarrow E$  corresponds to  $n + 1$  pairs of points  $(a_i, b_i)$  such that  $\rho(\theta, w)b_i = a_i$  for all  $i$ . For a  $v \in V(S_0)$ , the sum over  $x, x' \in E$  ranges over all possible candidates for  $h_1(v), h_2(v)$  with respect to a fixed  $\rho(\theta, w)$ . We weight each candidate as in Proposition 4.2, and taking the product over all  $v \in V(S_0)$  yields a quantity on the order of the number of pairs of congruent embeddings of  $(\mathcal{T}, S_0)$  extending  $h_1, h_2$ .

Now we argue Step 3 by a similar argument as in Proposition 4.2. Fixing a pair of congruent embeddings  $h_1, h_2 : V(S_0) \rightarrow E$ , for each  $\phi \in \text{Stab}(h_2(S_0))$  and each pair  $\theta, w$  such that  $\rho(\theta, w)$  sends  $h_2(S_0)$  to  $h_1(S_0)$ , we also have that  $\rho(\theta\phi, w)$  sends  $h_2(S_0)$  to  $h_1(S_0)$ . Therefore, the sum over all  $\theta$  and  $w$  counts  $h_1, h_2$  exactly  $\text{Stab}(h_1(S_0)) = \text{Stab}(\dim(S_0))$  many times. Therefore, after summing over all  $\theta$  and dividing by  $\text{Stab}(\dim(S_0))$ , we see that each pair of congruent embeddings of  $\mathcal{T}$  is counted on the order of one time. □

Therefore, in order to prove both Theorem 1.3 and 1.5, it is equivalent to show the following:

**Proposition 4.4.** *For a rooted simplex tree  $(\mathcal{T}, S_0)$ , and for any  $1 \leq k < \frac{d+1}{2}$ , define*

$$N_k = k + \sum_{\substack{S \in \mathcal{T} \\ \dim(S) > k}} (\dim(S) - k). \quad (4.8)$$

*Then for  $s = \max\left(\frac{dN_k+1}{N_k+1}, k + \frac{d-1}{2}\right)$ , we have that whenever  $|E| \gtrsim q^s$ ,*

$$R_{(\mathcal{T}, S_0)} \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.9)$$

*If we additionally impose the condition that  $d = 2$  and that  $q \equiv 3 \pmod{4}$ , for  $s = \frac{4N_1}{2N_1+1}$  we have that whenever  $|E| \gtrsim q^s$  the above holds as well.*

From here, our goal will be to apply the Hadamard three-lines Theorem to bound  $R_{(\mathcal{T}, S_0)}$  effectively. This geometrically corresponds to reducing the proof of Proposition 4.4 for  $\mathcal{T}$  to proving Proposition 4.4 for some collection of nicer simplex trees.

**4.2. Two geometric operations.** In this subsection, we describe and prove the two operations we use to simplify our sum: branch shifting and simplex unbalancing. Before we do so, we introduce some notation to make the statement and proof of these operations more clear.

**Definition 4.5.** For a rooted simplex tree  $(\mathcal{T}, S_0)$ , denote  $p(\mathcal{T}, S_0)$  as the infimum over  $s \in \mathbb{R}$  such that  $R_{(\mathcal{T}, S_0)} \lesssim E^{2|V(\mathcal{T})|} q^{-c(\mathcal{T})}$  (equivalently,  $\sum_{\delta \in \mathbb{D}} v_{\mathcal{T}}^2(\delta) \lesssim E^{2|V(\mathcal{T})|} q^{-c(\mathcal{T})}$ ) whenever  $|E| \gtrsim q^s$ .

**Lemma 4.6** (Branch Shifting). Given a rooted simplex tree  $(\mathcal{T}, S_0)$ , consider two simplices  $S_1$  and  $S_2$ , that are not necessarily distinct from each other or the root vertex. Further, suppose each of  $S_1, S_2$  contain respective, necessarily distinct, child vertices  $v_1$  and  $v_2$ . Consider the rooted tree  $(\mathcal{T}_1, S_0)$  formed from  $(\mathcal{T}, S_0)$  by deleting all branches of  $S_2$  at  $v_2$  and duplicating all branches of  $S_1$  at  $v_1$ . Similarly, consider the rooted tree  $(\mathcal{T}_2, S_0)$  formed from  $(\mathcal{T}, S_0)$  by deleting all branches of  $S_1$  at  $v_1$  and duplicating all branches of  $S_2$  at  $v_2$ . Then  $p(\mathcal{T}, S_0) \leq \max(p(\mathcal{T}_1, S_0), p(\mathcal{T}_2, S_0))$

Note that the quantity  $N_k$  in Proposition 4.4 is the same for  $\mathcal{T}, \mathcal{T}_1$ , and  $\mathcal{T}_2$ . Therefore, the above Lemma states that in order to prove Proposition 4.4 for  $\mathcal{T}$ , it suffices to prove Proposition 4.4 for  $\mathcal{T}_1$  and  $\mathcal{T}_2$ . An example of branch shifting shown in Figure 3, with  $S_0$  being the black 3-simplex,  $S_1, S_2$  both the red triangle,  $v_1$  being the parent node of the green branches, and  $v_2$  being the parent node of the blue branch.

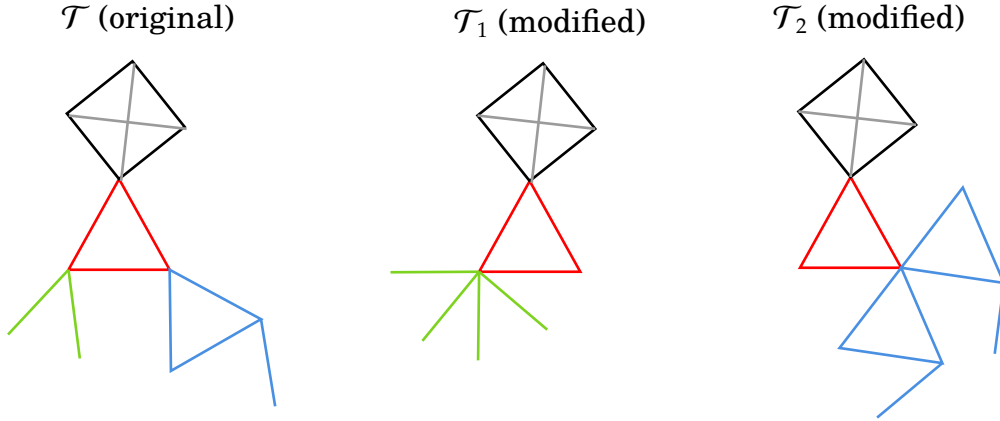


FIGURE 3. An example of branch shifting.

*Proof.* First, suppose that  $S_1$  and  $S_2$  are distinct from each other and from the root simplex, with parent vertices  $v_{0,1}$  and  $v_{0,2}$  respectively. Then for  $\ell \in \{1, 2\}$ , the expression  $R_{(\mathcal{T}, S_0)}$  contains the term

$$D_{\mathcal{P}(S_\ell)}(u, u') = \frac{1}{\text{Stab}(\dim(S_\ell))} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{\substack{v \in V(S_\ell) \\ v \neq v_{0,\ell}}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \prod_{B \in \mathcal{B}(S_\ell, v)} D_B(x, x'). \quad (4.10)$$

Now consider the modified expression

$$D_{\mathcal{P}(S_\ell)}(u, u', a_\ell) = \frac{1}{\text{Stab}(\dim(S_\ell))} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{\substack{v \in V(S_\ell) \\ v \neq v_{0,\ell}}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \left( \prod_{B \in \mathcal{B}(S_\ell, v)} D_B(x, x') \right)^{G_\ell(v)} \quad (4.11)$$

where  $G_\ell : V(S_\ell) \rightarrow \mathbb{R}$  is defined as

$$G_\ell(v) = \begin{cases} 1 & v \neq v_\ell \\ a_\ell & v = v_\ell. \end{cases} \quad (4.12)$$

We now consider the function  $\psi(a_1, a_2)$  defined as the expression for  $R_{(\mathcal{T}, S_0)}$  where we substitute all instances of  $D_{\mathcal{P}(S_\ell)}(u, u')$  with  $D_{\mathcal{P}(S_\ell)}(u, u', a_\ell)$ . Therefore, we see that

$$R_{(\mathcal{T}, S_0)} = \psi(1, 1). \quad (4.13)$$

We plan to invoke the Hadamard three-lines theorem to “unbalance”  $\psi(1, 1)$ , by bounding it by the geometric average of  $\psi(2, 0)$  and  $\psi(0, 2)$ . As the expression for  $\psi(2, 0)$  corresponds to duplicating the contributions of the branches of  $S_1$  at  $v_1$  and removing the contributions of the branches of  $S_2$  at  $v_2$ , and analogously for  $\psi(0, 2)$ , these two terms represent simplex trees of a “lower complexity” than that of  $\psi(1, 1)$ . For an example, consider the case in Figure 3, with  $S_1 = S_2$  as the red triangle, the green simplices as branches off of child vertex  $v_1$ , and the blue simplices as branches off of child vertex  $v_2$ . Then  $\mathcal{T}_1$  corresponds to the expression  $\psi(2, 0)$  and  $\mathcal{T}_2$  corresponds to the expression  $\psi(0, 2)$ .

It is easy to check that  $\psi$  satisfies the requirements for the special case of Hadamard three-lines in Corollary 2.5 (by an iterated use of the triangle inequality), so applying this gives us that

$$\psi(1, 1) \leq \psi(2, 0)^{\frac{1}{2}} \psi(0, 2)^{\frac{1}{2}} \quad (4.14)$$

Note that the terms  $\psi(2, 0)$  and  $\psi(0, 2)$  are equal to  $R_{(\mathcal{T}_1, S_0)}$  and  $R_{(\mathcal{T}_2, S_0)}$  respectively, where  $\mathcal{T}_1$  and  $\mathcal{T}_2$  are as defined in the Lemma statement. We also easily see that  $2c(\mathcal{T}) = c(\mathcal{T}_1) + c(\mathcal{T}_2)$  and that  $2|V(\mathcal{T})| = |V(\mathcal{T}_1)| + |V(\mathcal{T}_2)|$ . Since for any  $s > \max(p(\mathcal{T}_1), p(\mathcal{T}_2))$ , we have that  $|E| \gtrsim q^s$  implies

$$\psi(2, 0)\psi(0, 2) \lesssim \frac{|E|^{2|V(\mathcal{T}_1)|+2|V(\mathcal{T}_2)|}}{q^{c(\mathcal{T}_1)+c(\mathcal{T}_2)}}, \quad (4.15)$$

we conclude that

$$\psi(1, 1) \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.16)$$

In the case where  $S_1$  and  $S_2$  are the same simplex and are still distinct from the root node, we denote both of them as  $S$  for simplicity. We denote the parent vertex of  $S$  as  $v_0$ , and the two distinguished child vertices as  $v_1$  and  $v_2$ . Then the expression  $R_{(\mathcal{T}, S_0)}$  contains the term

$$D_{\mathcal{P}(S)}(u, u') = \frac{1}{\text{Stab}(\dim(S))} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{\substack{v \in V(S) \\ v \neq v_0}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \prod_{B \in \mathcal{B}(S, v)} D_B(x, x'). \quad (4.17)$$

Now consider the modified expression

$$\begin{aligned} & D_{\mathcal{P}(S)}(u, u', a_1, a_2) \\ &= \frac{1}{\text{Stab}(\dim(S))} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{\substack{v \in V(S) \\ v \neq v_0}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \left( \prod_{B \in \mathcal{B}(S, v)} D_B(x, x') \right)^{G_1(v) + G_2(v) - 1} \end{aligned} \quad (4.18)$$

where  $G_1$  and  $G_2$  are defined as in (4.12) above. We define  $\psi(a_1, a_2)$  as the expression for  $R_{(\mathcal{T}, S_0)}$  where we substitute each instance of  $D_{\mathcal{P}(S)}(u, u')$  with  $D_{\mathcal{P}(S)}(u, u', a_1, a_2)$ . We note that  $R_{(\mathcal{T}, S_0)} = \psi(1, 1)$ . Applying Hadamard three-lines and arguing as in the previous case completes the proof of the Lemma in this case.

The cases where one or more of  $S_1, S_2$  is the root of  $\mathcal{T}$  follow in an extremely similar fashion. Suppose that  $S_1$  is the root of  $\mathcal{T}$ , then  $R_{(\mathcal{T}, S_0)}$  is of the form

$$R_{(\mathcal{T}, S_0)} = \sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{v \in V(S_1)} \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \prod_{B \in \mathcal{B}(S_1, v)} D_B(x, x'). \quad (4.19)$$

Define  $\psi(a_1, a_2)$  as the modified expression

$$\sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{v \in V(S_1)} \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \left( \prod_{B \in \mathcal{B}(S_1, v)} D_B(x, x') \right)^{G_1(v)} \quad (4.20)$$

where we substitute each instance of  $D_{\mathcal{P}(S_2)}(u, u')$  with  $D_{\mathcal{P}(S_2)}(u, u', a_\ell)$ . From here, we apply Hadamard three-lines to  $\psi(1, 1)$  and argue as in the previous cases. Finally, in the case where  $S_1 = S_2 = S_0$ , we define  $\psi(a_1, a_2)$  as the modified expression for  $R_{(\mathcal{T}, S_0)}$  below.

$$\sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \prod_{v \in V(S_1)} \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \left( \prod_{B \in \mathcal{B}(S_1, v)} D_B(x, x') \right)^{G_1(v) + G_2(v) - 1} \quad (4.21)$$

Then applying Hadamard three-lines to  $\psi(1, 1)$  finishes this case.  $\square$

We now proceed to our second geometric operation:

**Lemma 4.7** (Simplex Unbalancing). *Given a rooted simplex tree  $(\mathcal{T}, S_0)$ , consider two distinct simplices  $S_1$  and  $S_2$  (one of which may be the root), each of dimension at least 2 and each with at least one free vertex. Choose positive integers  $k_1, k_2$  such that  $k_i$  does not surpass the number of free vertices of  $S_i$ . Consider the rooted tree  $(\mathcal{T}_1, S_0)$*

formed by deleting  $k_2$  free vertices of  $S_2$  and adding  $k_2$  free vertices to  $S_1$ . Similarly, consider the rooted tree  $(\mathcal{T}_2, S_0)$  formed by deleting  $k_1$  free vertices of  $S_1$  and adding  $k_1$  free vertices to  $S_2$ . Then  $p(\mathcal{T}) \leq \max(p(\mathcal{T}_1), p(\mathcal{T}_2))$ .

Similar to branch shifting, the operation of simplex unbalancing preserves the quantity  $N_k$  as defined in 4.4. Therefore, this lemma allows us to rephrase proving Proposition 4.4 for  $\mathcal{T}$  into proving Proposition 4.4 for  $\mathcal{T}_1$  and  $\mathcal{T}_2$ . An example of simplex unbalancing is shown in Figure 4, where we take  $S_0$  to be the black triangle,  $S_1$  to be the green triangle, and  $S_2$  to be the blue triangle.

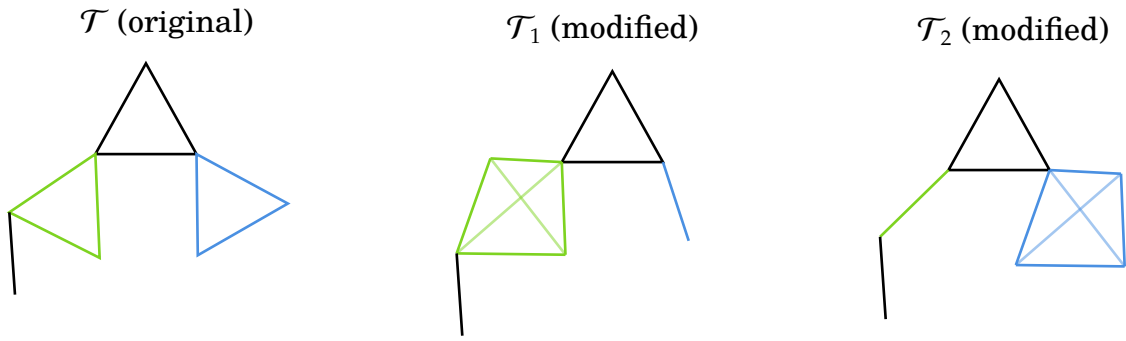


FIGURE 4. An example of simplex unbalancing.

To prove simplex unbalancing, we need the following Lemma:

**Lemma 4.8.** *Given a rooted simplex tree  $(\mathcal{T}, S_0)$  in  $\mathbb{F}_q^d$ , consider two distinct simplices  $S_1$  and  $S_2$ , each of dimension at least 2 and with at least one free vertex. Consider the rooted tree  $(\mathcal{T}', S_0)$  formed by deleting one free vertex of  $S_1$  and adding one free vertex to  $S_2$ . Then we have that*

$$\left( \frac{\text{Stab}(\dim(S_1))}{\text{Stab}(\dim(S'_1))} \cdot \frac{\text{Stab}(\dim(S_2))}{\text{Stab}(\dim(S'_2))} \right) q^{c(\mathcal{T}')} = q^{c(\mathcal{T})}. \quad (4.22)$$

*Proof.* By Lemma 2.20, we see that  $c(\mathcal{T}) - c(\mathcal{T}') = c(S_1) - c(S'_1) + c(S_2) - c(S'_2)$ . Let  $\dim(S_1) = n$  and  $\dim(S_2) = m$ , so that  $\dim(S'_1) = n - 1$  and  $\dim(S'_2) = m + 1$ . We split into cases based on the dimension of simplices  $S_1$  and  $S_2$ .

- If  $n, m \geq d$ , then by Lemma 2.12, we see that  $c(S)$  changes linearly by  $d$  as the dimension of both simplices change by 1. Therefore,  $c(S_1) - c(S'_1) + c(S_2) - c(S'_2) = d + (-d) = 0$ . The sizes of the stabilizers are trivial in this case, so we have the equality in (4.22).
- If  $n, m < d$ , then by Lemma 2.12 we have that

$$\begin{aligned} & c(S_1) - c(S'_1) + c(S_2) - c(S'_2) \\ &= \binom{n+1}{2} - \binom{n}{2} + \binom{m+1}{2} - \binom{m+2}{2} = n - m - 1. \end{aligned} \quad (4.23)$$

Similarly, by Proposition 4.1, we see that

$$\frac{\text{Stab}(n)}{\text{Stab}(n-1)} \cdot \frac{\text{Stab}(m)}{\text{Stab}(m+1)} = q^{\binom{d-n}{2} - \binom{d-n+1}{2} + \binom{d-m}{2} - \binom{d-m-1}{2}} = q^{n-m-1}. \quad (4.24)$$

So again, we get the equality in (4.22).

- If  $n < d, m \geq d$ , then again by Lemma 2.12 we have that

$$\begin{aligned} & c(S_1) - c(S'_1) + c(S_2) - c(S'_2) \\ &= \binom{n}{2} - \binom{n-1}{2} + d(m+1) - d(m+2) = n - d. \end{aligned} \quad (4.25)$$

The stabilizers  $\text{Stab}(m)$  and  $\text{Stab}(m+1)$  are trivial, and we have by Proposition 4.1 that

$$\frac{\text{Stab}(n)}{\text{Stab}(n-1)} = q^{\binom{d-n}{2} - \binom{d-n+1}{2}} = q^{n-d}. \quad (4.26)$$

Thus, we have the equality on (4.22). The case where  $n \geq d, m < d$  follows similarly.  $\square$

*Proof of Simplex Unbalancing.* For now, assume that  $S_1$  and  $S_2$  are not the root of  $\mathcal{T}$ , and denote the parent vertices of  $S_1$  and  $S_2$  as  $v_{0,1}$  and  $v_{0,2}$  respectively. Denote the set of free vertices of  $S_1$  as  $F_1$ , and the set of free vertices of  $S_2$  as  $F_2$ . For  $\ell \in \{1, 2\}$  the expression  $R_{(\mathcal{T}, S_0)}$  contains a term of the form

$$D_{\mathcal{P}(S_\ell)}(u, u') = \frac{1}{\text{Stab}(\dim(S_\ell))} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_\theta^{|F_\ell|} (u - \theta u') \prod_{\substack{v \in V(S_\ell) \\ v \neq v_{0,\ell} \\ v \notin F_\ell}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \prod_{B \in \mathcal{B}(S_\ell, v)} D_B(x, x'). \quad (4.27)$$

Now consider the modified expression

$$D_{\mathcal{P}(S_\ell)}(u, u', a_\ell) = \frac{1}{\text{Stab}(\dim(S_\ell))} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_\theta^{a_\ell + (|F_\ell| - k_\ell)} (u - \theta u') \prod_{\substack{v \in V(S_\ell) \\ v \neq v_{0,\ell} \\ v \notin F_\ell}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = u - \theta u'}} \prod_{B \in \mathcal{B}(S_\ell, v)} D_B(x, x'). \quad (4.28)$$

Consider the function  $\psi(a_1, a_2)$ , defined as the expression for  $R_{(\mathcal{T}, S_0)}$  where we substitute each instance of  $D_{\mathcal{P}(S_\ell)}(u, u')$  with  $D_{\mathcal{P}(S_\ell)}(u, u', a_\ell)$ . Therefore, we see that

$$R_{(\mathcal{T}, S_0)} = \psi(k_1, k_2). \quad (4.29)$$

It is clear that  $\psi$  satisfies the conditions needed for the special case of Hadamard three-lines as stated in Corollary 2.5 (again by iterated triangle inequality), so applying this gives us that

$$\psi(k_1, k_2) \leq \psi(k_1 + k_2, 0)^{s_1} \psi(0, k_1 + k_2)^{s_2} \quad (4.30)$$

for some positive  $s_i$  such that  $s_1 + s_2 = 1$ . Note that the only difference between  $\psi(k_1 + k_2, 0)$  and  $R_{(\mathcal{T}_1, S_0)}$  is the stabilizer constant in front of  $\psi(k_1 + k_2, 0)$ 's modified  $D_{\mathcal{P}(S_1)}$  and  $D_{\mathcal{P}(S_2)}$  terms. This is because in the process of applying Hadamard three-lines above, we changed the dimension of  $S_1$  and  $S_2$  by changing the power on the  $\lambda_\theta$  terms, but didn't change the stabilizer terms accordingly. Therefore, we have that

$$\psi(k_1 + k_2, 0) = \left( \frac{\text{Stab}(\dim(S_1) + k_2)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(\dim(S_2) - k_2)}{\text{Stab}(\dim(S_2))} \right) R_{(\mathcal{T}_1, S_0)}. \quad (4.31)$$

By a repeated application of Lemma 4.8, we have that

$$\left( \frac{\text{Stab}(\dim(S_1) + k_2)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(\dim(S_2) - k_2)}{\text{Stab}(\dim(S_2))} \right) \frac{1}{q^{c(\mathcal{T}_1)}} = \frac{1}{q^{c(\mathcal{T})}}. \quad (4.32)$$

After noting that  $|V(\mathcal{T}_1)| = |V(\mathcal{T})|$ , we see that for  $s > p(\mathcal{T}_1, S_0)$ ,

$$\psi(k_1 + k_2, 0) \lesssim \left( \frac{\text{Stab}(\dim(S_1) + k_2)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(\dim(S_2) - k_2)}{\text{Stab}(\dim(S_2))} \right) \left( \frac{|E|^{2|V(\mathcal{T}_1)|}}{q^{c(\mathcal{T}_1)}} \right) \quad (4.33)$$

$$\lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.34)$$

By an analogous argument, we also have that for  $s > p(\mathcal{T}_2, S_0)$ ,

$$\psi(0, k_1 + k_2) \lesssim \left( \frac{\text{Stab}(\dim(S_1) - k_1)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(\dim(S_2) + k_1)}{\text{Stab}(\dim(S_2))} \right) \left( \frac{|E|^{2|V(\mathcal{T}_2)|}}{q^{c(\mathcal{T}_2)}} \right) \quad (4.35)$$

$$\lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.36)$$

Therefore, whenever  $s > \max(p(\mathcal{T}_1), p(\mathcal{T}_2))$ , we may conclude that  $R_{(\mathcal{T}, S_0)} = \psi(k_1, k_2) \lesssim E^{2|V(\mathcal{T})|} q^{-c(\mathcal{T})}$ .

In the case that  $S_1$  or  $S_2$  is the root of  $(\mathcal{T}, S_0)$ , the proof follows similarly. Without loss of generality, let  $S_1$  be the root of  $\mathcal{T}$ . Again, let  $F_1$  and  $F_2$  be the set of free vertices of  $S_1$  and  $S_2$  respectively. The term  $R_{(\mathcal{T}, S_0)}$  is of the form

$$R_{(\mathcal{T}, S_0)} = \sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_\theta^{|F_1|}(w) \prod_{\substack{v \in V(S_1) \\ v \notin F_1}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = w}} \prod_{B \in \mathcal{B}(S_1, v)} D_B(x, x') \quad (4.37)$$

Define  $\psi(a_1, a_2)$  as the modified expression

$$\sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_\theta^{a_1 + |F_1| - k_1}(w) \prod_{\substack{v \in V(S_1) \\ v \notin F_1}} \sum_{\substack{x, x' \in E \\ x - \theta x' \\ = w}} \prod_{B \in \mathcal{B}(S_1, v)} D_B(x, x'), \quad (4.38)$$

where we substitute each instance of  $D_{\mathcal{P}(S_2)}(u, u')$  with  $D_{\mathcal{P}(S_2)}(u, u', a_\ell)$  as we did in the previous case. Applying Hadamard three-lines and arguing as in the previous case completes the proof of the Lemma.  $\square$

**4.3. Proof of Proposition 4.4 and Theorem 1.4.** We have the following simple fact that we use throughout this subsection.

**Lemma 4.9.** *Consider a graph  $G$  and a subgraph  $H$  of  $G$ . If for some  $s \in \mathbb{R}$ ,  $E \subset \mathbb{F}_q^d$  contains a positive proportion of congruence classes of embeddings of  $G$  in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ , then  $E$  contains a positive proportion of congruence classes of embeddings of  $H$  in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ .*

*Proof.* Note that a congruence class of  $G$  determines a congruence class of  $H$ . As long as the number of congruence classes of  $H$  in  $\mathbb{F}_q^d$  is unbounded as  $q$  grows, it is clear that  $E$  cannot contain a positive proportion of congruence classes of embeddings of  $G$  without containing a positive proportion of congruence classes of embeddings of  $H$ . □

*Proof of Theorem 1.4.* We may assume that all simplices in  $\mathcal{T}$  are exactly dimension  $k$  by Lemma 4.9. We first choose a root  $S_0$  of  $\mathcal{T}$ . Then, to any simplex  $S \in \mathcal{T}$  such that  $S$  has more than one child vertex, we apply branch shifting to two child vertices of  $S$  to reduce to a collection of modified  $\mathcal{T}_i$  where  $S$  has one less child vertex. Repeating this process for all  $S \in \mathcal{T}$  allows us to reduce to the case where  $\mathcal{T}$  is a simplex tree consisting only of  $n$ -simplices with at most one child vertex. Simplex trees of this form are  $k$ -weak simplex trees, and we apply Theorem 1.12 of [Bri+23] to conclude that for  $s = k + \frac{d-1}{2}$ ,  $|E| \gtrsim q^s$  implies that  $E$  contains all congruence classes of  $\mathcal{T}$  in  $\mathbb{F}_q^d$ . □

We first prove a special case of Proposition 4.4 for a simple class of simplex trees  $\mathcal{T}$  before continuing to the proof of 4.4 in full generality.

**Lemma 4.10.** *For an integer  $k$  such that  $1 \leq k < \frac{d+1}{2}$ , let  $\mathcal{T}$  be a simplex tree consisting of exactly one simplex  $S$  of dimension greater than  $k$ . For  $s = \max\left(\frac{d \dim(S)+1}{\dim(S)+1}, k + \frac{d-1}{2}\right)$ , we have that for  $E \subset \mathbb{F}_q^d$ , whenever  $|E| \gtrsim q^s$ ,*

$$\sum_{\delta \in \text{ID}} v_{\mathcal{T}}^2(\delta) \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.39)$$

*If we additionally impose the condition that  $d = 2$ ,  $k = 1$  and that  $q \equiv 3 \pmod{4}$ , for  $s = \frac{4 \dim(S)}{2 \dim(S)+1}$  we have that whenever  $|E| \gtrsim q^s$  the above holds as well.*

*Proof.* We view  $S$  as the root of  $\mathcal{T}$ , and by Lemma 4.9, assume that all simplices of dimension at most  $k$  are exactly dimension  $k$ . If  $S$  has no child vertices, we directly apply the bounds found in [Ben+17] and [McD20] and no further work is needed. If  $S$  has at least one child vertex, we may reduce to the case where  $S$  has exactly one child vertex. To do this, we apply branch shifting to a pair of child vertices of  $S$  to reduce to the case where  $S$  has one less child vertex, and repeat this process. Denote the sole child vertex of  $S$  as  $v$ .

Next, we modify the branches of  $S$ . To any  $k$ -simplex with multiple child vertices, we repeatedly apply branch shifting to reduce to the case that each  $k$ -simplex only has a single child vertex. After this modification, each branch of  $S$  is of the form of a  $k$ -weak simplex tree as defined in Section 3. Then we may write the sum  $R_{(\mathcal{T}, S)}$  as follows:

$$R_{(\mathcal{T}, S)} = \frac{1}{\text{Stab}(\dim(S))} \sum_{w \in \mathbb{F}_q^d} \sum_{\theta \in O_d(\mathbb{F}_q)} \lambda_{\theta}^{\dim(S)}(w) \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \prod_{B \in \mathcal{B}(S, v)} D_B(x, x'). \quad (4.40)$$

We note that  $\prod_{B \in \mathcal{B}(S, v)} D_B(x, x')$  counts the number of congruent embeddings (based at  $x$  and  $x'$ ) of the structure formed by the union of all the branches of  $S$  at  $v$ . We note that the union of all of the branches of  $S$  at  $v$  is a  $k$ -weak simplex tree  $T_k$  rooted at  $v$ . Therefore,  $\prod_{B \in \mathcal{B}(S, v)} D_B(x, x')$  counts the number of congruent pairs of embeddings of  $T_k$  into  $E$  based at  $x$  and  $x'$ .

Defining  $e(T_k) = \ell \binom{k+1}{2}$  as the number of edges of  $T_k$  as in Section 3, let  $t_1, \dots, t_{e(T_k)} \in \mathbb{F}_q$  determine a congruence class of embeddings of  $T_k$ . Defining  $f_{T_k, t_1, \dots, t_{e(T_k)}}$  as in Section 3, we may write

$$\begin{aligned} R_{(\mathcal{T}, S)} &= \frac{1}{\text{Stab}(\dim(S))} \sum_{\theta, w} \lambda_{\theta}^{\dim(S)}(w) \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} \prod_{B \in \mathcal{B}(S, v)} D_B(x, x') \\ &= \frac{1}{\text{Stab}(\dim(S))} \sum_{t_1, \dots, t_{e(T_k)}} \sum_{\theta, w} \lambda_{\theta}^{\dim(S)}(w) \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} f_{T_k, t_1, \dots, t_{e(T_k)}}(x) f_{T_k, t_1, \dots, t_{e(T_k)}}(x'). \end{aligned} \quad (4.41)$$

We also may assume that  $t_1, \dots, t_{e(T_k)}$  range over nonzero elements in  $\mathbb{F}_q$  rather than all of  $\mathbb{F}_q$ . This is because we only care about containing a positive proportion of congruence classes of  $\mathcal{T}$  in  $\mathbb{F}_q^d$ , and the number of congruence classes of  $T_k$  determined by a set of  $e(T_k)$  distances with at least one distance zero is negligible compared to the number of congruence classes determined by a set of  $e(T_k)$  nonzero distances. Fixing a congruence class  $t_1, \dots, t_{e(T_k)}$  of  $T_k$ , we note that the inside of the outermost sum of (4.41) is exactly of the form  $\sum_{\theta, w} \lambda_{\theta}^{\dim(S)}(w) \Gamma_{\theta, T_k}(w)$ , where  $\Gamma_{\theta, T_k}(w)$  is defined as in Section 3. From this point onward, we split into cases for  $\mathbb{F}_q^2$  and  $\mathbb{F}_q^d$ . For notational convenience, we let  $n$  denote the dimension of  $S$  for the rest of this Lemma.

In  $\mathbb{F}_q^2$ , from an immediate application of Lemma 3.3 we have that for  $|E| \gtrsim q^{\frac{4n}{2n+1}}$ ,

$$\begin{aligned} &\frac{1}{\text{Stab}(n)} \sum_{t_1, \dots, t_{e(T_k)}} \sum_{\theta, w} \lambda_{\theta}^n(w) \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} f_{T_k, t_1, \dots, t_{e(T_k)}}(x) f_{T_k, t_1, \dots, t_{e(T_k)}}(x') \\ &\lesssim \frac{1}{\text{Stab}(n)} q^{\ell \binom{k+1}{2}} \left( \frac{|E|^{2\ell k + 2n + 2}}{q^{2\ell \binom{k+1}{2} + 2n - 1}} \right) = \frac{1}{\text{Stab}(n)} \left( \frac{|E|^{2\ell k + 2n + 2}}{q^{\ell \binom{k+1}{2} + 2n - 1}} \right). \end{aligned} \quad (4.42)$$

It is easy to show that  $\mathcal{T}$  has  $\ell k + n + 1$  vertices. We also see that by Lemma 2.20,  $c(\mathcal{T}) = c(S) + \ell \binom{k+1}{2}$ . In  $d = 2$ , we see that  $\text{Stab}(n) \approx 1$  and  $c(S) = 2n - 1$  by Lemma 2.12. Therefore, we have shown that for  $|E| \gtrsim q^{\frac{4n}{2n+1}}$ ,

$$\sum_{d \in \mathbb{D}} v_{\mathcal{T}}^2(d) \approx R_{(\mathcal{T}, S)} \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.43)$$

In  $\mathbb{F}_q^d$ ,  $d > 2$ , we apply Lemma 3.4 to (4.41), so for  $|E| \gtrsim q^{\max(\frac{dn+1}{n+1}, k + \frac{d-1}{2})}$ ,

$$\frac{1}{\text{Stab}(n)} \sum_{t_1, \dots, t_{\ell} \in (T_k)} \sum_{\theta, w} \lambda_{\theta}^n(w) \sum_{\substack{x, x' \in E \\ x - \theta x' = w}} f_{T_k, t_1, \dots, t_{\ell}}(x) f_{T_k, t_1, \dots, t_{\ell}}(x') \quad (4.44)$$

$$\lesssim \frac{1}{\text{Stab}(n)} q^{\ell \binom{k+1}{2}} \left( \frac{|E|^{2\ell k + 2n + 2}}{q^{2\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}} \right) \quad (4.45)$$

$$= \frac{1}{\text{Stab}(n)} \left( \frac{|E|^{2\ell k + 2n + 2}}{q^{\ell \binom{k+1}{2} + d(n+1) - \binom{d+1}{2}}} \right). \quad (4.46)$$

Again we see that  $\mathcal{T}$  has  $\ell k + n + 1$  vertices, and that  $c(\mathcal{T}) = c(S) + \ell \binom{k+1}{2}$ . When  $\dim(S) = n \geq d - 1$ , we have that  $\text{Stab}(n) \approx 1$  and that  $c(S) = d(n + 1) - \binom{d+1}{2}$ . When  $\dim(S) = n < d - 1$ , we have that  $\text{Stab}(n)$  is  $q^{\binom{d-n}{2}}$  by Proposition 4.1. As  $\binom{d-n}{2} + d(n + 1) - \binom{d+1}{2} = \binom{n+1}{2}$ , we see that for  $|E| \gtrsim q^{\max(\frac{dn+1}{n+1}, k + \frac{d-1}{2})}$ ,

$$\sum_{d \in \mathbb{D}} v_{\mathcal{T}}^2(d) \approx R_{(\mathcal{T}, S)} \lesssim \frac{|E|^{2|V(\mathcal{T})|}}{q^{c(\mathcal{T})}}. \quad (4.47)$$

□

*Proof of Proposition 4.4 for general simplex trees.* Fix an integer  $k$  in the statement of Proposition 4.4 such that  $1 \leq k < \frac{d+1}{2}$ . If there are less than two simplices of dimension greater than  $k$  in  $\mathcal{T}$ , then we are done by Lemma 4.10 and Theorem 1.4. Now suppose there are at least two simplices in  $\mathcal{T}$  of dimension greater than  $k$ . Then there must be at least two  $k$ -leafs in  $\mathcal{T}$ . Fix two distinct  $k$ -leafs  $S_1$  and  $S_2$  of dimensions  $n, m > k$  respectively.

- Step 1: We reduce to the case where one of  $S_1, S_2$  has at most one child vertex and the other has no child vertices. To do this, we iteratively apply branch shifting. Consider the case where  $S_1$  has more than one child vertex. Here, we apply branch shifting to a pair of child vertices of  $S_1$  to reduce to the case where  $S_1$  has one less child vertex, and repeat this process until  $S_1$  has a single child vertex. Therefore, we may assume that  $S_1$  has either zero or one child vertices. Similarly, we may assume that  $S_2$  has either zero or one child vertices. If either  $S_1$  or  $S_2$  has no child vertices, we are done. Otherwise, apply branch shifting to the child vertex of  $S_1$  and the child vertex of  $S_2$ .

- **Step 2:** Denote  $F_1$  and  $F_2$  as the set of free vertices of  $S_1$  and  $S_2$ . If  $S_1$  has a child vertex, set  $k_1 = |F_1| - k + 1$ , otherwise, set  $k_1 = |F_1| - k$ . Similarly, if  $S_2$  has a child vertex, set  $k_2 = |F_2| - k + 1$ , otherwise, set  $k_2 = |F_2| - k$ . Then apply simplex unbalancing to  $S_1$  and  $S_2$  with the appropriate  $k_1, k_2$ . Note that in all of these cases, this process lets us reduce to the case where one of  $S_1, S_2$  is an  $k$ -simplex, and the other is a  $(\dim(S_1) + \dim(S_2) - k)$ -dimensional simplex.

After an application of these two steps, we have reduced the goal of proving Proposition 4.4 for  $\mathcal{T}$  to proving Proposition 4.4 for a collection of  $\mathcal{T}_i$ , all with one less simplex of dimension greater than  $k$  than  $\mathcal{T}$ , since the quantity  $N_k$  is invariant under branch shifting and simplex unbalancing. Note that for each of these  $\mathcal{T}_i$ , we have that

$$\sum_{\substack{S \in \mathcal{T}_i \\ \dim(S) > k}} \dim(S) = \left( \sum_{\substack{S \in \mathcal{T} \\ \dim(S) > k}} \dim(S) \right) - k. \quad (4.48)$$

Therefore, if we start with a simplex tree  $\mathcal{T}$ , and repeatedly apply the above two steps to reduce proving Proposition 4.4 for  $\mathcal{T}$  to proving Proposition 4.4 for a collection of  $\mathcal{T}_j$  each containing exactly one simplex  $S_j$  of dimension greater than  $k$ , we see that for each  $\mathcal{T}_j$ ,

$$\dim(S_j) = k + \sum_{\substack{S \in \mathcal{T} \\ \dim(S) > k}} (\dim(S) - k) \quad (4.49)$$

Applying Lemma 4.10 to each  $T_j$  completes the proof of Proposition 4.4 in full generality.  $\square$

## 5. CYCLES OF SIMPLICES

In this section, we discuss partial results related to cycles of simplices.

**Definition 5.1.** *Let  $C$  be a finite set of simplices  $S_1, \dots, S_k$  where  $S_i$  shares a vertex with  $S_{i+1}$  for all  $1 \leq i \leq k - 1$ , and  $S_k$  shares a vertex with  $S_1$ . Then we call  $C$  a cycle of simplices.*

Our goal is to find an  $s$  such that  $\mathbb{F}_q^d$  contains a positive proportion of congruence classes of  $C$  in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ . As argued for simplex trees in Lemma 2.20, we have the following:

**Lemma 5.2.** *The number of congruence classes of  $C$  in  $\mathbb{F}_q^d$  is on the order of  $q^{\sum_{S \in C} c(S)}$ . We denote  $\sum_{S \in C} c(S)$  as  $c(C)$ .*

Following the framework developed in Section 4, to find an  $s$  corresponding to a given simplex cycle  $C$ , we first reduce this goal finding an  $s$  for a simple class of simplex cycles consisting of one large simplex and all other simplices of dimension one. This is more precisely described in the proposition below.

**Proposition 5.3.** *Consider a cycle of simplices  $C$  with  $k$  simplices, and define.*

$$N = 1 + \sum_{S \in C, \dim(S) > 1} (\dim(S) - 1)$$

*Let  $C'$  be a cycle of simplices containing  $k - 1$  simplices of dimension 1 and one simplex of dimension  $N$ . Then if for some  $s \in \mathbb{R}$ ,  $E$  contains a positive proportion of congruence classes of  $C'$  in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ , then the same is true for  $C$  and the same  $s$  as well.*

The proof of the above Proposition requires a modification of the simplex unbalancing technique from earlier. However after we reduce to this class of simple cycles, we find that various technical obstructions occur in trying to generalize the methods of Section 3 to this setting. Morally, however, we expect the  $s$  for  $C'$  to be slightly larger or equal to that of simplices of dimension  $N$ , as cycles of arbitrary length are much more abundant than simplices of dimension  $N$  in  $\mathbb{F}_q^d$ .

**5.1. The setup and Hadamard three-lines reduction.** Consider a cycle of simplices  $C$  given by a collection of simplices  $S_1, \dots, S_k$ , where  $S_i$  shares a vertex with  $S_{i+1}$  for all  $1 \leq i \leq k - 1$  and  $S_k$  shares a vertex with  $S_1$ . Again, we are interested in finding an  $s$  such that  $\mathbb{F}_q^d$  contains a positive proportion of congruence classes of  $C$  in  $\mathbb{F}_q^d$  whenever  $|E| \gtrsim q^s$ . Using the same Cauchy-Schwarz setup as in Section 4.1, we show that this problem is equivalent to finding an  $s$  such that whenever  $|E| \gtrsim q^s$ ,

$$\sum_{\delta \in \mathbb{D}} v_C^2(\delta) \lesssim \frac{|E|^{2|V(C)|}}{q^{c(C)}} \quad (5.1)$$

As usual,  $\mathbb{D}$  denotes the space of congruence classes of  $C$ ,  $v_C(\delta)$  denotes the number of embeddings in the congruence class  $\delta$  contained in  $E$ , and  $V(C)$  denotes the vertex set of  $C$ . Analogous to simplex trees, a vertex in  $C$  is called a free vertex if it is shared by only one simplex. As per Section 4.1, the goal is now be to write  $\sum_{\delta \in \mathbb{D}} v_C^2(\delta)$  in a way that is more amenable to applying Hadamard three-lines. We have two different ways of doing this, depending on whether the two simplices  $S_1, S_2 \in C$  we distinguish share a vertex or not.

First, consider nonadjacent simplices  $S_1, S_2$  of  $C$  with dimension  $m$  and  $n$  respectively. Denote  $\mathcal{H}_1$  and  $\mathcal{H}_2$  as the two chains of simplices connecting  $S_1$  to  $S_2$ , denote  $v_{1,i}$  as the vertex shared by  $S_1$  and  $\mathcal{H}_i$ , and denote  $v_{2,i}$  similarly as the vertex shared by  $S_2$  and  $\mathcal{H}_i$ . Let  $\delta_1$  and  $\delta_2$  be respective congruence class of these two chains. Let the function  $f_{\delta_1}(x_1, x_2)$  count the number of embeddings of  $\mathcal{H}_1$  into  $E$  in a given congruence class  $\delta_1$ , such that the embedding sends  $v_{1,1}$  to  $x_1$  and  $v_{1,2}$  to  $x_2$ . Similarly, let the function  $g_{\delta_2}(x_1, x_2)$  count the number of embeddings of  $\mathcal{H}_2$  into  $E$  in a given congruence class  $\delta_2$ , such that the embedding sends  $v_{2,1}$  to  $x_1$  and  $v_{2,2}$  to  $x_2$ . Then

we define

$$D_{\text{nonadj}}(\mathcal{C}, S_1, S_2) = \frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\delta_1, \delta_2} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{w_1, w_2 \in \mathbb{F}_q^d} \lambda_\theta^{m-1}(w_1) \lambda_\phi^{n-1}(w_2) \sum_{\substack{u_1, u'_1, v_1, v'_1 \in E \\ u_1 - \theta u'_1 = \\ v_1 - \theta v'_1 = w_1}} \sum_{\substack{u_2, u'_2, v_2, v'_2 \in E \\ u_2 - \phi u'_2 = \\ v_2 - \phi v'_2 = w_2}} f_{\delta_1}(u_1, u_2) f_{\delta_1}(u'_1, u'_2) g_{\delta_2}(v_1, v_2) g_{\delta_2}(v'_1, v'_2). \quad (5.2)$$

**Proposition 5.4.** *For nonadjacent simplices  $S_1, S_2 \in \mathcal{C}$  with dimensions  $m$  and  $n$  respectively, we have that*

$$\sum_{\delta \in \mathcal{D}} v_{\mathcal{C}}^2(\delta) \approx D_{\text{nonadj}}(\mathcal{C}, S_1, S_2). \quad (5.3)$$

*Proof.* We first count the number of pairs of congruent embeddings  $h, h' : V(S_1 \cup S_2) \rightarrow E$  of the collection of two disjoint simplices  $S_1, S_2$ . Let the dimensions of  $S_1$  and  $S_2$  be  $m$  and  $n$  respectively. By the methods of Section 4.1, we see that this is on the order of

$$\frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{w_1, w_2 \in \mathbb{F}_q^d} \lambda_\theta^{m+1}(w_1) \lambda_\theta^{n+1}(w_2) \quad (5.4)$$

$$= \frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{w_1, w_2 \in \mathbb{F}_q^d} \lambda_\theta^{m-1}(w_1) \lambda_\phi^{n-1}(w_2) \sum_{\substack{u_1, u'_1, v_1, v'_1 \in E \\ u_1 - \theta u'_1 = \\ v_1 - \theta v'_1 = w_1}} \sum_{\substack{u_2, u'_2, v_2, v'_2 \in E \\ u_2 - \phi u'_2 = \\ v_2 - \phi v'_2 = w_2}} 1. \quad (5.5)$$

In the equality above, we just unpack the definition of  $\lambda_\theta$  and  $\lambda_\phi$ . For every congruent pair of embeddings  $h, h'$  the expression above counts,  $u_1$  and  $v_1$  are the vertices  $h(v_{1,1})$  and  $h(v_{1,2})$  of  $h(S_1)$ , and  $u'_1$  and  $v'_1$  are the vertices  $h'(v_{1,1})$  and  $h'(v_{1,2})$  of  $h'(S_1)$ . Similarly  $u_2$  and  $v_2$  are the the vertices  $h(v_{2,1})$  and  $h(v_{2,2})$  of  $h(S_2)$ , and  $u'_2$  and  $v'_2$  are the vertices  $h'(v_{2,1})$  and  $h'(v_{2,2})$  of  $h'(S_2)$ . See Figure 5 for an example where  $\dim(S_1) = 2$  and  $\dim(S_2) = 3$ .

We now weight the count of each pair of congruent embeddings  $h, h' : V(S_1 \cup S_2) \rightarrow E$  by how many pairs of congruent embeddings  $\tilde{h}, \tilde{h}' : V(\mathcal{C}) \rightarrow E$  it extends to. This is exactly

$$\sum_{\delta_1, \delta_2} f_{\delta_1}(u_1, u_2) f_{\delta_1}(u'_1, u'_2) g_{\delta_2}(v_1, v_2) g_{\delta_2}(v'_1, v'_2), \quad (5.6)$$

completing the proof of the proposition.  $\square$

Now consider adjacent simplices  $S_1, S_2$  of  $\mathcal{C}$  with dimensions  $m$  and  $n$  respectively. Denote  $\mathcal{H}$  as the chain of simplices connecting  $S_1$  to  $S_2$ , denote  $v_1$  as the vertex shared by  $\mathcal{H}$  and  $S_1$ , and denote  $v_2$  similarly as the vertex shared by  $\mathcal{H}$  and  $S_2$ . Let  $\delta'$  be a congruence class of  $\mathcal{H}$ . Let the function  $F_{\delta'}(x_1, x_2)$  count the number of

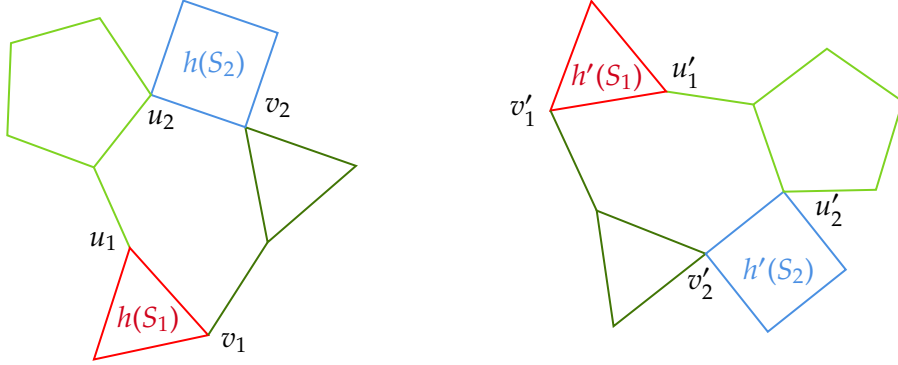


FIGURE 5. A potential extension  $\tilde{h}$  and  $\tilde{h}'$  of two embeddings  $h, h' : V(S_1 \cup S_2) \rightarrow E$ .

embeddings of  $\mathcal{H}$  into  $E$  in the given congruence class  $\delta'$ , such that the embedding sends  $v_1$  to  $x_1$  and  $v_2$  to  $x_2$ . Then we define

$$D_{\text{adj}}(C, S_1, S_2) = \frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\delta'} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{x_1, x_2 \in \mathbb{F}_q^d} \lambda_{\theta}^{m-1}(x - \theta x') \lambda_{\phi}^{n-1}(x - \phi x') \sum_{\substack{u_1, u_1' \in E \\ u_1 - \theta u_1' \\ = x - \theta x'}} \sum_{\substack{u_2, u_2' \in E \\ u_2 - \phi u_2' \\ = x - \phi x'}} F_{\delta'}(u_1, u_2) F_{\delta'}(u_1', u_2'). \quad (5.7)$$

**Proposition 5.5.** *For adjacent simplices  $S_1, S_2 \in C$  with dimensions  $m$  and  $n$  respectively, we have that*

$$\sum_{\delta \in \mathbb{D}} v_C^2(\delta) \approx D_{\text{adj}}(C, S_1, S_2). \quad (5.8)$$

*Proof.* We first count the number of pairs  $h, h' : V(S_1 \cup S_2) \rightarrow E$  of congruent embeddings of the collection of two simplices  $S_1, S_2$  sharing a vertex. Then by the methods of Section 4.1, we see that this is on the order of

$$\begin{aligned} & \frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{x_1, x_2 \in \mathbb{F}_q^d} \lambda_{\theta}^m(x - \theta x') \lambda_{\phi}^n(x - \phi x') \quad (5.9) \\ &= \frac{1}{\text{Stab}(m)\text{Stab}(n)} \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{x_1, x_2 \in \mathbb{F}_q^d} \lambda_{\theta}^{m-2}(x - \theta x') \lambda_{\phi}^{n-2}(x - \phi x') \sum_{\substack{u_1, u_1' \in E \\ u_1 - \theta u_1' \\ = x - \theta x'}} \sum_{\substack{u_2, u_2' \in E \\ u_2 - \phi u_2' \\ = x - \phi x'}} 1. \quad (5.10) \end{aligned}$$

For every congruent pair of embeddings  $h, h'$  the expression above counts,  $u_1$  and  $u_2$  are the vertices  $h(v_1)$  and  $h(v_2)$ , and  $u_1'$  and  $u_2'$  are the vertices  $h'(v_1)$  and  $h'(v_2)$ . See Figure 6 for an example where  $\dim(S_1) = 2$  and  $\dim(S_2) = 3$ .

We now weight the count of each pair of congruent embedding  $h, h' : V(S_1 \cup S_2) \rightarrow E$  by how many pairs of congruent embeddings  $\tilde{h}, \tilde{h}' : V(C) \rightarrow E$  it extends to. This is exactly

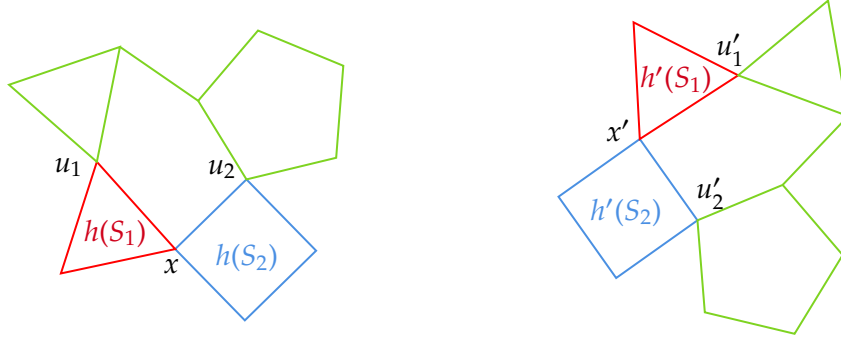


FIGURE 6. A potential extension  $\tilde{h}$  and  $\tilde{h}'$  of two embeddings  $h, h' : V(S_1 \cup S_2) \rightarrow E$ .

$$\sum_{\delta'} F_{\delta'}(u_1, u_2) F_{\delta'}(u'_1, u'_2), \quad (5.11)$$

completing the proof of the proposition.  $\square$

We now prove a modified version of simplex unbalancing for cycles to reduce to the case where all but one simplex in  $C$  has dimension 1. We have the following definition for convenience of notation, and the following Lemma:

**Definition 5.6.** For a cycle of simplices  $C$ , denote  $p(C)$  as the infimum of  $s \in \mathbb{R}$  such that  $\sum_{\delta \in \mathbb{D}} v_C^2(\delta) \lesssim E^{2|V(C)|} q^{-c(C)}$  whenever  $|E| \gtrsim q^s$ .

**Lemma 5.7.** Given a cycle of simplices  $C$ , consider two distinct simplices  $S_1$  and  $S_2$  each of dimension at least 2, so each have at least one free vertex. Consider the cycle of simplices  $C'$  formed by deleting one free vertex of  $S_1$  and adding one free vertex to  $S_2$ . Then we have that

$$\left( \frac{\text{Stab}(\dim(S_1))}{\text{Stab}(\dim(S'_1))} \cdot \frac{\text{Stab}(\dim(S_2))}{\text{Stab}(\dim(S'_2))} \right) q^{c(C')} = q^{c(C)}.$$

*Proof.* This proof follows identically to Lemma 4.8 as  $c(\mathcal{T})$  and  $c(C)$  are both defined to be the sum of  $c(S)$  for all simplices  $S$  in the structure.  $\square$

**Lemma 5.8 (Simplex Unbalancing for Cycles).** Given a cycle of simplices  $C$ , consider two distinct simplices  $S_1$  and  $S_2$  each of dimension at least 2. Consider the cycle of simplices  $C_1$  formed by deleting all free vertices of  $S_2$  and adding the number of free vertices deleted to  $S_1$ . Similarly, consider the cycle of simplices  $C_2$  formed by deleting all free vertices of  $S_1$  and adding the number of free vertices deleted to  $S_2$ . Then  $p(C) \leq \max(p(C_1), p(C_2))$ .

*Proof.* Throughout this proof, we denote the modification of the simplices  $S_1$  and  $S_2$  in  $C_1$  as  $S_1^+$  and  $S_2^-$  respectively. Similarly, we denote the modification of  $S_1$  and  $S_2$  in  $C_2$  as  $S_1^-$  and  $S_2^+$ .

Suppose that  $S_1$  and  $S_2$  are nonadjacent, of dimensions  $m$  and  $n$  respectively. Define the function  $\psi(a_1, a_2) : \mathbb{R}^2 \rightarrow \mathbb{R}$  as the expression  $D_{\text{nonadj}}(C, S_1, S_2)$ , where we replace the exponents on  $\lambda_\theta$  and  $\lambda_\psi$  with  $a_1$  and  $a_2$  respectively. The expression  $D_{\text{nonadj}}(C, S_1, S_2)$  can be written as  $\psi(m-1, n-1)$ , and by an application of Corollary 2.5, we see that

$$\psi(m-1, n-1) \leq \psi(m+n-2, 0)^{s_1} \psi(0, m+n-2)^{s_2} \quad (5.12)$$

for some positive  $s_i$  such that  $s_1 + s_2 = 1$ . Note that the only difference between  $\psi(m+n-2, 0)$  and  $D_{\text{nonadj}}(C_1, S_1^+, S_2^-)$  is the stabilizer terms out front. This is because in the process of applying Hadamard three-lines above, we changed the dimension of  $S_1$  and  $S_2$  by changing the power on the  $\lambda_\theta$  terms, but didn't change the stabilizer terms accordingly. Therefore, we have that

$$\psi(k_1+k_2, 0) = \left( \frac{\text{Stab}(\dim(S_1) + \dim(S_2) - 1)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(1)}{\text{Stab}(\dim(S_2))} \right) D_{\text{nonadj}}(C_1, S_1^+, S_2^-). \quad (5.13)$$

By a repeated application of Lemma 5.7, we have that

$$\left( \frac{\text{Stab}(\dim(S_1) + \dim(S_2) - 1)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(1)}{\text{Stab}(\dim(S_2))} \right) \frac{1}{q^{c(C_1)}} = \frac{1}{q^{c(C)}}. \quad (5.14)$$

After noting that  $|V(C_1)| = |V(C)|$ , we have that for  $s > p(C_1)$ ,

$$\psi(k_1+k_2, 0) \lesssim \left( \frac{\text{Stab}(\dim(S_1) + k_2)}{\text{Stab}(\dim(S_1))} \cdot \frac{\text{Stab}(\dim(S_2) - k_2)}{\text{Stab}(\dim(S_2))} \right) \left( \frac{|E|^{2|V(C_1)|}}{q^{c(C_1)}} \right) \quad (5.15)$$

$$\lesssim \frac{|E|^{2|V(C)|}}{q^{c(C)}}. \quad (5.16)$$

By an analogous argument, we also have that for  $s > p(C_2)$ , that  $\psi(0, k_1+k_2) \lesssim E^{2|V(C_2)|} q^{-c(C_2)}$ . Therefore, whenever  $s > \max(p(C_1), p(C_2))$ , we may conclude that  $\sum_{\delta \in \mathbb{D}} \nu_C^2(\delta) \approx \psi(k_1, k_2) \lesssim E^{2|V(C)|} q^{-c(C)}$ .

The proof for adjacent simplices  $S_1$  and  $S_2$  follows similarly but with  $D_{\text{adj}}$  instead.  $\square$

Given any cycle of simplices  $C$  with at least three simplices, by applying simplex unbalancing repeatedly and noting that this operation preserves the sum of the dimensions of simplices in  $C$ , we arrive at Proposition 5.3. All that remains is to find an  $s$  such that  $E$  contains a positive proportion of congruence classes of  $C'$  whenever  $|E| \gtrsim q^s$ .

**5.2. Technical obstructions.** For the rest of this section, we describe our attempts to bound the number of pairs of congruent embeddings of  $C'$  in  $E$ , and the obstacles that prevent us from applying a generalization of the methods in Section 3 and [AIM24]. We assume that  $N$ , the dimension of the large simplex in  $C'$ , is

larger than  $d$  so that we don't have additional stabilizer terms in the sums that we're working with. This is to add clarity to where technical obstructions arise, but the same issues occur with the stabilizer terms present.

We define  $P_{t_1, \dots, t_{k-1}}(u, u')$  as the number of paths  $(x_1, \dots, x_k)$  in  $E$  of length  $k-1$  such that  $x_1 = u$ ,  $x_2 = u'$ , and  $\|x_{i+1} - x_i\| = t_i$  for all  $1 \leq i \leq k-1$ . Let  $C'$  be a cycle of simplices with one  $N$ -dimensional simplex and  $k-1$  1-dimensional simplices. As in Section 4.3, to find an  $s$  such that  $E$  contains a positive proportion of congruence classes of  $C'$  whenever  $|E| \gtrsim q^s$ , we bound the sum

$$\sum_{\delta \in \mathbb{D}} \nu_C^2(\delta) \approx \sum_{t_1, \dots, t_{k-1}} \sum_{\theta, w} \lambda_\theta^{N-1}(w) \sum_{\substack{u_1, u'_1, u_2, u'_2 \\ u_1 - \theta u'_1 \\ = u_2 - \theta u'_2 \\ = w}} P_{t_1, \dots, t_{k-1}}(u_1, u_2) P_{t_1, \dots, t_{k-1}}(u'_1, u'_2) \quad (5.17)$$

As in Section 4.3, we may assume that  $t_1, \dots, t_k$  range over nonzero elements in  $\mathbb{F}_q^d$  rather than all of  $\mathbb{F}_q^d$ . We fix a congruence class of paths by fixing  $t_1, \dots, t_{k-1}$ , and drop the  $t_i$ s from the notation from  $P$  for notational convenience. Then, similar to the goal of Section 3, we want to show that

$$\sum_{\theta, w} \lambda_\theta^{N-1}(w) \sum_{\substack{u, u', v, v' \\ u - \theta u' \\ = v - \theta v' \\ = w}} P(u, v) P(u', v') \lesssim \frac{|E|^{2|V(C)|}}{q^{c(C)+k-1}} \quad (5.18)$$

Following the work of Section 3 and of [AIM24], we would like to define a function of the form

$$\beta_\theta(w) = \sum_{\substack{u, u', v, v' \\ u - \theta u' \\ = v - \theta v' \\ = w}} P(u, v) P(u', v') \quad (5.19)$$

We first calculate the zero Fourier coefficient of  $\beta_\theta$  below.

$$\widehat{\beta}_\theta(0) = \frac{1}{q^d} \sum_{\substack{u, v, u', v' \\ \|u-v\| = \|u'-v'\|}} P(u, v) P(u', v') \quad (5.20)$$

$$= \frac{1}{q^d} \sum_{u, v} P(u, v) \sum_{\substack{u', v' \\ \|u-v\| = \|u'-v'\|}} P(u', v') \quad (5.21)$$

$$\approx \frac{1}{q^d} \cdot \frac{|E|^k}{q^{k-1}} \cdot \frac{|E|^k}{q^k} = \frac{|E|^{2k}}{q^{d+2k-1}} \quad (5.22)$$

The first equality of (5.22) follows from the fact that  $E$  contains a statistically correct number of cycles of length  $k$  for  $|E| \gtrsim q^{(d+2)/2}$  by [IJM21] (with the technical condition that  $k > 4$ ), and that  $E$  contains a statistically correct number of paths of length  $k-1$  for a lower threshold by [Ben+16].

We have that

$$\sum_{\theta, w} \lambda_{\theta}^{N-1}(w) \sum_{\substack{u, u', v, v' \\ u-\theta u' \\ =v-\theta v' \\ =w}} P(u, v)P(u', v') = \sum_{\theta, w} \lambda_{\theta}^{N-1}(w)\beta(w) \quad (5.23)$$

This is a sum of a very similar form as the sum  $\sum_{\theta, w} \lambda_{\theta}^n(w)\Gamma_{T_k, \theta}(w)$  in Section 3. Therefore, we try to apply the method of proof of Lemma 3.4 to bound this term. However, this is not effective, since if we attempt to break up  $\sum_{\theta, w} \lambda_{\theta}^{N-1}(w)\beta(w)$  in the same way as in Lemma 3.4, the term  $\sum_{w, \theta} \lambda_{\theta}^{N-1}\widehat{\beta}_{\theta}(0)$  is too large by a factor of  $q^{d-1}$ .

Another idea is to instead consider the function  $\alpha_{\theta} : \mathbb{F}_q^d \times \mathbb{F}_q^d \rightarrow \mathbb{R}$  defined as

$$\alpha_{\theta}(w_1, w_2) = \sum_{\substack{u, u', v, v' \\ u-\theta u' = w_1 \\ v-\theta v' = w_2}} f(u, v)f(u', v') \quad (5.24)$$

We note that  $\alpha_{\theta}(w, w) = \beta_{\theta}(w)$ . The benefit of  $\alpha$  is that after taking the Fourier transform of  $\alpha$  in both variables,  $\widehat{\alpha}_{\theta}(0, 0)$  is smaller than  $\widehat{\beta}_{\theta}(0)$ . More precisely, in a very similar manner as for  $\Gamma_{T_k, \theta}$  in Section 3, we see that

$$\widehat{\alpha}_{\theta}(m_1, m_2) = q^{2d}\widehat{P}(m_1, m_2)\overline{\widehat{P}(\theta^{-1}m_1, \theta^{-1}m_2)} \quad (5.25)$$

We now see that

$$\widehat{\alpha}_{\theta}(0, 0) \approx \frac{|E|^{2k}}{q^{2d+2k-2}}, \quad (5.26)$$

which is a factor of  $q^{d-1}$  smaller than  $\widehat{\beta}_{\theta}(0)$ . Therefore, we define  $B_n$  as  $\sum_{\theta, w} \lambda_{\theta}^n(w)\alpha_{\theta}(w, w)$  to break up our sum instead as

$$B_n = \sum_{\theta, w} \lambda_{\theta}^n(w)\alpha_{\theta}(w, w) \leq \left| \sum_{\theta, w} \lambda_{\theta}^{n-1}(w) \left( \lambda_{\theta}(w) - \frac{|E|^2}{q^d} \right) (\alpha_{\theta}(w, w) - \widehat{\alpha}_{\theta}(0, 0)) \right| \\ + \sum_{w, \theta} \lambda_{\theta}^n(w)\widehat{\alpha}_{\theta}(0, 0) + \frac{|E|^2}{q^d} B_{n-1} \quad (5.27)$$

$$=: I + II + III. \quad (5.28)$$

As  $\widehat{\alpha}_{\theta}(0, 0)$  is sufficiently small, the term *II* is of the proper size. Term *I* can be bounded nontrivially by pulling absolute values inside, taking the supremum of  $\lambda_{\theta}^{n-1}(w)$ , and applying Cauchy Schwarz. We handle the term of the form

$$\sum_w (\alpha_{\theta}(w, w) - \widehat{\alpha}_{\theta}(0, 0))^2$$

by dominating the sum over  $w$  by sums over  $w_1, w_2$ , and then running an analogous argument to 3.6. This gives us that

$$\sum_{\theta, w} (\alpha_\theta(w, w) - \widehat{\alpha}_\theta(0, 0))^2 \approx q^{2d} \left( \frac{|E|^{4k-4}}{q^{4k-4}} \right). \quad (5.29)$$

Plugging in bounds from 2.8 gives us that  $I$  is on the order of  $E^{2|V(C)|} q^{-c(C)-k+1}$  for an  $s$  slightly larger than that of an  $N$  simplex, matching what we morally expect.

As we bound term *III* by induction, it remains to bound the base case  $B_1$ , which we are unfortunately unable to bound effectively. In the spirit of how we bound  $A_1$  in Lemma 3.4, one may easily generalize Lemma 2.6 to functions  $\mathbb{F}_q^d \times \mathbb{F}_q^d \rightarrow \mathbb{R}$  by modifying the proof seen in [Ben+17]. However, our bound for  $B_1$  is still too big by a factor of  $q^{d-1}$ . If we reduce the base case to  $B_0$ , the  $L_1$  norm of  $\alpha$  is too large by a factor of  $q^{d-1}$ . Finally, if we break up  $B_1$  slightly differently in hopes to extract cancellation from the negative term, we are still off by the same factor.

The recurring issue is that we need for one of  $\alpha$  or  $\beta$ , both its  $L_1$  norm and zero Fourier coefficient to be small. While  $\widehat{\alpha}_\theta(0, 0)$  is smaller than  $\widehat{\beta}_\theta(0)$ , the  $L_1$  norm of  $\alpha_\theta$  is much larger than  $\beta_\theta$ . We have this difference because we get an extra factor of  $q^d$  in the denominator by taking the Fourier transform of both coordinates for  $\alpha_\theta$ . What we believe needs to happen is for the negative contribution of term  $I$  (without the absolute values) to cancel with part of term *III* of 5.28. Therefore, we need to somehow extract cancellation between terms  $I$  and *III*, potentially involving results on number-theoretic sums. We also propose a potential workaround in Section 6.2.

## 6. FUTURE WORK

**6.1. Edge gluing.** We end this paper by discussing a natural generalization of the work in the past two sections, considering structures of simplices where two simplices may share an edge or some higher dimensional face rather than a vertex. The Hadamard three-lines framework established in Section 4 can be applied to such a situation, allowing us to reduce to a simple class of simplex structures that encourages the use of a combination of group action and inductive techniques. To illustrate this, we work through a simple example.

Consider the graph  $G$  formed by two 3-simplices  $S_1, S_2$  sharing an edge, as shown in Figure 7. We would like to find an  $s$  such that for  $E \subset \mathbb{F}_q^3$ ,  $E$  contains a positive proportion of congruence classes of  $G$  whenever  $|E| \gtrsim q^s$ . Note that there are on the order of  $q^{11}$  congruence classes of  $G$  in  $\mathbb{F}_q^d$ . Denoting  $\mathbb{D}$  as the set of congruence classes of  $G$ , and  $v_G(\delta)$  as the number of embeddings of  $G$  in  $E$  in the congruence class  $\delta$ , by the Cauchy Schwarz setup as presented in Section 4 it is sufficient to find an  $s$  such that for  $|E| \gtrsim q^s$ ,

$$\sum_{\delta \in \mathbb{D}} v_G^2(\delta) \lesssim \frac{|E|^{12}}{q^{11}} \quad (6.1)$$



$G$ , two 3-simplices sharing an edge

$H$ , a 4-simplex and a 1-simplex sharing an edge

FIGURE 7. An example of Hadamard three-lines simplification for edge-gluing.

For two embeddings  $h_1, h_2 : V(G) \rightarrow E$  of  $G$ ,  $h_1$  and  $h_2$  are congruent if and only if there exist  $\theta, \phi \in O_3(\mathbb{F}_q)$  and  $w_1, w_2 \in \mathbb{F}_q^3$  such that  $h_1(v) = \theta h_2(v) + w_1$  for all  $v \in S_1$  and  $h_1(v) = \phi h_2(v) + w_2$  for all  $v \in S_2$ . Therefore, we see by a similar argument as in Section 4.1 that

$$\sum_{\delta \in \mathbb{D}} v_G^2(\delta) \approx \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{\substack{x, y \in \mathbb{F}_q^d \\ x - \theta x' = y - \theta y' \\ x - \phi x' = y - \phi y'}} \lambda_\theta^2(x - \theta x') \lambda_\phi^2(x - \phi x') \quad (6.2)$$

Note that in general stabilizer terms are introduced (as in Section 4.1), but in this example the stabilizer terms are trivial. Applying Hadamard three-lines in the style of simplex unbalancing gives us that

$$\sum_{\delta \in \mathbb{D}} v_G^2(\delta) \lesssim \sum_{\theta, \phi \in O_d(\mathbb{F}_q)} \sum_{\substack{x, y \in \mathbb{F}_q^d \\ x - \theta x' = y - \theta y' \\ x - \phi x' = y - \phi y'}} \lambda_\theta^3(x - \theta x') \lambda_\phi^1(x - \phi x'), \quad (6.3)$$

where the right hand side is on the order of  $\sum_{\delta \in \mathbb{D}} v_H^2(\delta)$ , where  $H$  is the graph shown in Figure 7. Therefore, to find an  $s$  for  $G$ , it suffices to find an  $s$  such that for  $|E| \gtrsim q^s$ ,

$$\sum_{\delta \in \mathbb{D}} v_H^2(\delta) \lesssim \frac{|E|^{12}}{q^{11}}. \quad (6.4)$$

In more complicated simplex structures where all simplices are connected either by vertices or edges, an inductive Hadamard three-lines approach similar to Section 4 can reduce to the case where  $\mathcal{T}$  is a large simplex  $S$  where at a single vertex, standard (1-simplex) trees are attached, and at a single edge, tree structures formed by gluing 2-links (shown in Figure 8) are attached.

Therefore, 2-links and trees of 2-links form a natural analogue to paths and trees in the edge-gluing setting. Their inductive nature suggests that there exists an  $s$  such that  $E \subset \mathbb{F}_q^d$ ,  $|E| \gtrsim q^s$  implies that  $E$  contains a statistically correct number of 2-links of arbitrary length in any congruence class, similar to chains and trees

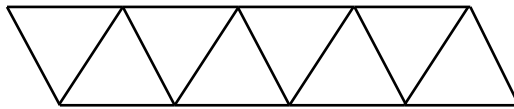


FIGURE 8. A 2-link of length 8

of simplices as in the work of [Ben+16] and [Bri+23]. Considering simplex structures where simplices can share up to  $n$ -dimensional faces, a similar Hadamard three-lines approach can be used to reduce to a large simplex with  $k$ -links attached (defined analogously) for  $1 \leq k \leq n$ .

In attempting to bound  $\sum_{\delta \in \mathbb{D}} v_H^2(\delta)$  in a similar manner to Section 5.2, we run into similar technical obstructions. We propose a different approach in the following subsection.

**6.2. Thresholds.** Intuitively, it is clear that if  $E$  contains a positive proportion of congruence classes in  $\mathbb{F}_q^d$  of a graph  $G$  whenever  $|E| \gtrsim q^s$ , then for  $|E| \gtrsim q^{s+\epsilon}$ ,  $E$  should contain "almost all" of the congruence classes of  $G$  in  $\mathbb{F}_q^d$ . However, the rate at which this threshold phenomenon acts, and the precise definition of "almost all," is of importance as it may provide a different approach to finding such an  $s$  for graphs  $G$  and  $H$  sharing a vertex, edge, or more complicated subgraph that avoids the obstacles presented in Sections 5 and 6.

Suppose we have two arbitrary graphs  $G$  and  $H$ , and we pick an  $s$  such that for  $E \subset \mathbb{F}_q^d$ ,  $|E| \gtrsim q^s$ ,  $E$  contains a positive proportion of congruence classes of  $G$  in  $\mathbb{F}_q^d$ , and  $E$  contains a statistically correct number of embeddings of every congruence class of  $H$  in  $\mathbb{F}_q^d$ . This is the setup that we have in Section 3, where  $G$  is a simplex and  $H$  is a  $k$ -weak simplex tree. Being able to quantify how many congruence classes of  $G$  that  $E$  contains when  $|E| \gtrsim q^{s+\epsilon}$  could open the door to showing that  $E$  contains a positive proportion of congruence classes of  $G$  attached to  $H$  at a vertex/edge/some subgraph via a probabilistic argument and concentration inequalities. An approach such as this would avoid the need to work with the explicit sums corresponding to  $G$  and  $H$  attached at some subgraph, bypassing the obstacles found in Sections 5 and 6 that occur while trying to separate out the individual contributions from  $G$  and  $H$  in these sums.

## 7. ACKNOWLEDGEMENTS

We thank Prof. Brian McDonald for several helpful conversations. We also thank Jacques Hadamard for his great result as we celebrate his 159 birthday this year.

This paper was written as part of the SMALL REU Program 2024. We appreciate the support of Williams College, as well as funding from NSF grant DMS-2241623, Duke University, Emmanuel College Cambridge, the Finnerty Fund, Princeton University, University of Michigan, and Williams College. A.I. was supported in part by the NSF under grant no. 2154232. A.I. wishes to thank the Isaac Newton Institute (INI) in Cambridge, Great Britain, for their support and hospitality. This

work was completed during the INI program, entitled "Multivariate approximation, discretization, and sampling recovery."

## REFERENCES

- [AIM24] Esen Aksoy, Alex Iosevich, and Brian MacDonald. *Interpolation of point configurations in the discrete plane*. 2024. arXiv: [2408.07010](https://arxiv.org/abs/2408.07010) [math.CO]. URL: <https://arxiv.org/pdf/2408.07010>.
- [Ben+16] M. Bennett, J. Chapman, D. Covert, D. Hart, A. Iosevich, and J. Pakianathan. "Long paths in the distance graph over large subsets of vector spaces over finite fields". In: *Journal of the Korean Mathematical Society* 53.1 (2016), pp. 115–126. ISSN: 0304-9913, 2234-3008. DOI: [10.4134/JKMS.2016.53.1.115](https://doi.org/10.4134/JKMS.2016.53.1.115).
- [Ben+17] M. Bennett, D. Hart, A. Iosevich, J. Pakianathan, and M. Rudnev. "Group actions and geometric combinatorics in  $\mathbb{F}_q^d$ ". In: *Forum Mathematicum* 29.1 (2017), pp. 91–110. DOI: [10.1515/forum-2015-0251](https://doi.org/10.1515/forum-2015-0251).
- [Bri+23] Paige Bright, Xinyu Fang, Barrett Heritage, Alex Iosevich, and Maxwell Sun. *Improved bounds for embedding certain configurations in subsets of vector spaces over finite fields*. 2023. arXiv: [2308.09215](https://arxiv.org/abs/2308.09215) [math.CO]. URL: <https://arxiv.org/abs/2308.09215>.
- [Cha+12] Jeremy Chapman, M. Burak Erdoan, Derrick Hart, Alex Iosevich, and Doowon Koh. "Pinned distance sets, k-simplices, Wolffs exponent in finite fields and sum-product estimates". In: *Mathematische Zeitschrift* 271 (2012), pp. 63–93. DOI: [10.1007/s00209-011-0852-4](https://doi.org/10.1007/s00209-011-0852-4).
- [Du+23] Xiumin Du, Yumeng Ou, Kevin Ren, and Ruixiang Zhang. *New improvement to Falconer distance set problem in higher dimensions*. 2023. arXiv: [2309.04103](https://arxiv.org/abs/2309.04103) [math.CA]. URL: <https://arxiv.org/abs/2309.04103>.
- [Erd46] Paul Erds. "On Sets of Distances of  $n$  points". In: *American Mathematical Monthly* 53.5 (1946), pp. 248–250. DOI: [10.2307/2305092](https://doi.org/10.2307/2305092).
- [Fal85] K. J Falconer. "On the Hausdorff Dimensions of Distance Sets". In: *Mathematika* 32.2 (1985), pp. 206–212. DOI: [10.1017/S0025579300010998](https://doi.org/10.1017/S0025579300010998).
- [GK15] Larry Guth and Nets Hawk Katz. "On the Erdos distinct distance problem in the plane". In: *Annals of Mathematics* 181.1 (2015), pp. 155–190. DOI: [10.4007/annals.2015.181.1.2](https://doi.org/10.4007/annals.2015.181.1.2).
- [Gre+15] A. Greenleaf, A. Iosevich, B. Liu, and E. Palsson. "A group-theoretic viewpoint on Erdos-Falconer problems and the Mattila integral". In: *Revista Matemática Iberoamericana* 31.3 (2015), pp. 799–810. DOI: [10.4171/RMI/854](https://doi.org/10.4171/RMI/854).
- [Gut+20] Larry Guth, Alex Iosevich, Yumeng Ou, and Hong Wang. "On Falconer's distance set problem in the plane". In: *Inventiones mathematicae* 219 (2020), pp. 779–830. DOI: [10.1007/s00222-019-00917-x](https://doi.org/10.1007/s00222-019-00917-x).

- [Har+11] Derrick Hart, Alex Iosevich, Doowon Koh, and Misha Rudnev. “Averages over hyperplanes, sum-product theory in vector spaces over finite fields and the Erdos-Falconer distance conjecture”. In: *Transactions of the American Mathematical Society* 363.6 (2011), pp. 3255–3275.
- [IJM21] Alex Iosevich, Gail Jardine, and Brian McDonald. “Cycles of arbitrary length in distance graphs on  $\mathbb{F}_q^d$ ”. In: *Proceedings of the Steklov Institute of Mathematics* 314.1 (2021), pp. 27–43. ISSN: 00815438. DOI: [10.1134/S0081543821040027](https://doi.org/10.1134/S0081543821040027).
- [IP19] Alex Iosevich and Hans Parshall. “Embedding distance graphs in finite field vector spaces”. In: *Journal of the Korean Mathematical Society* 56.6 (2019), pp. 1515–1528. ISSN: 0304-9913, 2234-3008. DOI: [10.4134/JKMS.j180776](https://doi.org/10.4134/JKMS.j180776).
- [IR07] Alex Iosevich and Misha Rudnev. “Erdos distance problem in vector spaces over finite fields”. In: *Transactions of the American Mathematical Society* 359.12 (2007), pp. 6127–6142.
- [McD20] Alex McDonald. “Congruence classes of large configurations in vector spaces over finite fields”. In: *Functiones et Approximatio Commentarii Mathematici* 62.1 (2020), pp. 131–141. DOI: [10.7169/facm/1814](https://doi.org/10.7169/facm/1814).
- [Mur+22] Brendan Murphy, Giorgis Petridis, Thang Pham, Misha Rudnev, and Sophie Stevens. “On the pinned distances problem in positive characteristic”. In: *Journal of the London Mathematical Society* 105.1 (Jan. 2022), pp. 469–499. ISSN: 1469-7750. DOI: [10.1112/jlms.12524](https://doi.org/10.1112/jlms.12524). URL: <http://dx.doi.org/10.1112/jlms.12524>.
- [SS03] Elias Stein and Rami Shakarchi. *Fourier Analysis*. Princeton University Press, 2003.
- [SV08] Jozsef Solymosi and Van H. Vu. “Near Optimal Bounds for the Erdos Distinct Distances Problem in High Dimensions”. In: *Combinatorica* 28 (2008), pp. 113–125. DOI: [10.1007/s00493-008-2099-1](https://doi.org/10.1007/s00493-008-2099-1).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN  
 Email address: [timcheek@umich.edu](mailto:timcheek@umich.edu)

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE  
 Email address: [jc2407@cam.ac.uk](mailto:jc2407@cam.ac.uk)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SANTA BARBARA  
 Email address: [picogilman@ucsb.edu](mailto:picogilman@ucsb.edu)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ROCHESTER  
 Email address: [alex.iosevich@rochester.edu](mailto:alex.iosevich@rochester.edu)

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY  
 Email address: [kj5388@princeton.edu](mailto:kj5388@princeton.edu)

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH  
 Email address: [palsson@vt.edu](mailto:palsson@vt.edu)

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY  
*Email address:* [vismay.sharan@yale.edu](mailto:vismay.sharan@yale.edu)

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE  
*Email address:* [jms13@williams.edu](mailto:jms13@williams.edu)

DEPARTMENT OF MATHEMATICS, DUKE UNIVERSITY  
*Email address:* [mariehelene.tome@duke.edu](mailto:mariehelene.tome@duke.edu)