Linear Hashing with ℓ_∞ guarantees and two-sided Kakeya bounds

Manik Dhar* Zeev Dvir†

Abstract

We show that a randomly chosen linear map over a finite field gives a good hash function in the ℓ_∞ sense. More concretely, consider a set $S \subset \mathbb{F}_q^n$ and a randomly chosen linear map $L : \mathbb{F}_q^n \to \mathbb{F}_q^t$ with q^t taken to be sufficiently smaller than $|S|$. Let U_S denote a random variable distributed uniformly on S. Our main theorem shows that, with high probability over the choice of L, the random variable $L(U_S)$ is close to uniform in the ℓ_∞ norm. In other words, *every* element in the range $\mathbb{F}_q^t$ has about the same number of elements in S mapped to it. This complements the widely-used Leftover Hash Lemma (LHL) which proves the analog statement under the statistical, or ℓ_1, distance (for a richer class of functions) as well as prior work on the expected largest 'bucket size' in linear hash functions [ADM+99]. By known bounds from the load balancing literature [RS98], our results are tight and show that linear functions hash as well as truly random function up to a constant factor in the entropy loss. Our proof leverages a connection between linear hashing and the finite field Kakeya problem and extends some of the tools developed in this area, in particular the polynomial method.

1 Introduction

Let $S \subset \{0,1\}^n$ be a set. In many scenarios, one is interested in 'hashing' the space $\{0,1\}^n$ into a smaller space so that the set S (on which we may have little or no information) is mapped in a way that is close to uniform. Specifically, we may need to find a function $H : \{0,1\}^n \to \{0,1\}^t$ so that the random variable $H(U_S)$ is close to the uniform distribution, where U_S denotes a random variable distributed uniformly on the set S. An important parameter here is the 'entropy-loss' given by $\log_2 |S| - t$. Clearly, this quantity has to be non negative, and, in practice, we would like it to be as small as possible.

An important result in this area is the celebrated Leftover Hash Lemma (LHL) of Impagliazzo, Levin and Luby [ILL89] which asserts that, the above scenario can be handled by choosing H at random from a family of universal hash functions (one in which for every $x \neq y$ the probability that $H(x) = H(y)$ is at most 2^{-t} over the choice of H).

Lemma 1.1 (Leftover Hash Lemma [ILL89])**.** *Let $S \subset \{0,1\}^n$ and suppose $H : \{0,1\}^n \to \{0,1\}^t$ is chosen from a family of universal hash functions with $t \leq \log_2 |S| - 2\log_2(1/\epsilon)$. Then the random variable[1] $(H, H(U_S))$ is ϵ-close to uniform in the ℓ_1-norm.*[2]

*Department of Computer Science, Princeton University. Email: manikd@princeton.edu.

†Department of Computer Science and Department of Mathematics, Princeton University. Email: zdvir@princeton.edu. Research supported by NSF grant DMS-1953807.

[1] In the notation $(H, H(U_S))$ we assume that the function H is represented by a string of bits of some fixed length.

[2] Typically, the conclusion of the lemma is stated with respect to the statistical distance (or total variation distance) which is defined to be 1/2 of the ℓ_1 distance

A few comments about the LHL are in order. The first is that, using a standard averaging argument, the LHL implies that, for any given set, most choices of H will be good, in the sense that $H(U_S)$ will be close to uniform in the ℓ_1 distance. It is also known that the *entropy loss* of the LHL, namely $2\log(1/\epsilon)$, is the smallest possible for any family of functions [RTS00]. Lastly, it is possible to generalize the LHL to handle arbitrary distributions of high min-entropy[3] (not just those uniform on a set). This follows from the fact that any distribution with min-entropy k is a convex combination of 'flat' distributions (those uniform on a set of size 2^k).

A convenient choice of a universal family of hash functions is that given by all linear maps over the finite field of two elements $\mathbb{F}_2$. That is, the LHL says that, if one picks a linear map $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ uniformly at random, then, w.h.p over the choice of L, the random variable $L(U_S)$ will be close to uniform in the ℓ_1-distance. Our main theorem shows that, with slightly larger entropy loss, one can give a stronger guarantee on the output, stated in ℓ_∞ distance to uniform. A reason to consider linear maps is their simplicity and ease of implementation (only requiring very basic bit operations) for applications. Since the full statement of the theorem is quite technical (stemming from our attempts to optimize the various constants) we start by giving an informal statement. The full statements of our results (also for other larger finite fields) are given in Section 2.

Theorem 1.2 (Main theorem (informal))**.** *Let $S \subset \mathbb{F}_2^n$ and let $t = \log_2 |S| - O(\log_2(\log_2 |S|/\tau\delta))$. Then, a $1-\delta$ fraction of all linear maps $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ are such that $L(U_S)$ is $\tau 2^{-t}$-close to uniform in the ℓ_∞ norm. That is, for all $y \in \mathbb{F}_2^t$ we have*

$$|\mathbf{Pr}[L(U_S) = y] - 2^{-t}| \leq \tau 2^{-t}.$$

An equivalent way to state this theorem comes from the observation that the set of elements in $\mathbb{F}_2^n$ mapping to a particular $y \in \mathbb{F}_2^t$ is always of the form $a_y + U$, where U is the kernel of the mapping L and $a_y \in \mathbb{F}_2^n$ is some shift. In this view, the theorem says that most $(n-t)$-dimensional subspaces $U \subset \mathbb{F}_2^n$ are such that *all* of their shifts intersect S in about the same number of points (up to a multiplicative factor of $1+\tau$). We devote Section 3 to a more detailed treatment of this view, which will be the one used in the proof. The question of bounding the maximal 'bucket size' (all elements mapping to a single $y \in \mathbb{F}_2^t$) in a random linear hash function was previously studied and we compare our results to the state-of-the-art in this area ([ADM+99]) in Section 2 after the formal statement of our results.

Our choice of the letter τ instead of ϵ as in the LHL is not accidental and is meant to highlight the fact that, in the ℓ_∞ setting, we can take τ to be greater than 1. When $\tau < 1$ the conclusion of our theorem, namely that $L(U_S)$ is $\tau/2^{-t}$-close to uniform in ℓ_∞, implies that $L(U_S)$ is also τ-close to uniform in ℓ_1. However, our theorem is still meaningful when $\tau > 1$, even though it says nothing about ℓ_1 distance. The advantage of taking τ to be large comes from the fact that it can reduce our entropy loss (this can be done up to a point, as is stated in the formal theorem statement below). To give an example of a scenario in which we can take large τ consider the case where the linear map L is used to derive a key for a digital signature scheme. We would like the key $L(U_S)$ to be close to uniform since we know that a uniform key prevents the adversary from producing a forgery with more than negligible probability. However, if we apply our theorem with large τ (say polynomial in n) we get that the probability of producing a forgery may increase by at most a factor of $1+\tau$ which still results in negligible probability of forgery. More generally, the case of large τ is relevant whenever we only care about events of small probability staying small. Another

[3] A distribution has min-entropy at least k if any output has probability at most 2^{-k}.

paper which focuses on these aspects of the LHL (that is, when we only care about low probability events) is [BDK+11].

The need for ℓ_∞ guarantees for hashing appears in many places in the literature. For example, in Cryptography, in the context of key generation for local data storage [CHS06] and batch verifying zero-knowledge proofs [KRV21] and in Computational in the context of uniformly generating a solution to NP-search problems (see Section 6.2.4.2 in [Gol08]). It is possible to guarantee ℓ_∞ hashing by using a larger and more complex classes of functions, for example high degree polynomials over a large finite field [ABI86]. For the applications in [CHS06] and [Gol08] our results allows one to use linear maps instead of polynomials, hence simplifying the proofs.

Our proofs leverage a connection between linear hashing and finite field Furstenberg sets (which generalize Kakeya sets). A k-dimensional Furstenberg set $S \subset \mathbb{F}_q^n$ is a set which has a large intersection with a k-flat (k-dimensional affine subspaces) in each direction. That is, for any k-dimensional subspace $U \subset \mathbb{F}_q^n$ there is a shift $s(U)$ such that the affine subspace $s(U) + U$ has a large intersection with the set S. The goal in this area is to prove lower bounds on the size of such sets. Surprisingly, such lower bounds play a role in explicit constructions of seeded extractors [DW08, DKSS13] which are randomness efficient variants of the LHL. However, the connection between Furstenberg sets and linear hashing we leverage in this paper is *unrelated* to the work on extractors mentioned above and is of a completely different nature. This connection was first observed in [DDL21b] and was used there to improve the best lower bounds on Furstenberg sets. Our work relies heavily on the methods developed in [DDL21b] (as well as other papers) and extends them in several respects. We devote Section 3 to a more complete discussion of this connection and, in particular, to explaining the phrase 'two-sided Kakeya bounds' from the title of the paper.

Acknowledgments: We are grateful to Or Ordentlich, Oded Regev and Barak Weiss for comments that led us to pursue this line of work. Their interest in theorems of this kind arose from trying to strengthen their breakthrough [ORW22] on lattice coverings, which uses the two dimensional Kakeya bounds of [KLSS11]. (A new paper by the same group of authors, using the results of the current paper, is in preparation).

Paper organization: The rest of the paper is organized as follows. In Section 2 we state our main theorems formally. In Section 2.1 we discuss the tightness of our results, compare them to prior work, and discuss possible generalizations. In Section 3 we discuss the connection to the theory of Furstenberg/Kakeya sets and introduce notations and definitions that will be used in the proofs. In Section 4 we give a high level overview of the proof. Section 5 contains the proofs of our main theorems with a lemma,giving an improved bound on Furstenberg sets, proved in Section 6.

2 Formal statement of our results

This section contains four variants of our main result. The four cases correspond to the distinction between large finite fields and $\mathbb{F}_2$ and between arbitrary τ and the special case $\tau > 1$ (in which we can get slightly better constants). We begin with the statement for large finite field and arbitrary τ.

Theorem 2.1. *Let $n \geq 5$ and let $S \subset \mathbb{F}_q^n$ be a set. Let $\tau > 0$ be a real number and $\delta \in (0,1)$ s.t*

$$q \geq 32 \max\left(\frac{n(1+\tau)}{(\tau\delta)^2}, n\right).$$

Suppose $q^r < |S| \leq q^{r+1}$ for some $4 \leq r \leq n-1$ and let $t = r-3$. Then a $1-\delta$ fraction of all surjective linear maps $L : \mathbb{F}_q^n \to \mathbb{F}_q^t$ are such that $L(U_S)$ is τ/q^t-close to the uniform distribution in the ℓ_∞ norm.

Notice that, in the setting above, the entropy loss, when measured in $\mathbb{F}_q$-dimension is at most 4. The restriction to the case of surjective linear maps is natural as these are maps that do not 'lose' entropy unnecessarily (one can consider all linear maps by increasing δ slightly).

The above theorem can be used to derive similar results for small fields, by treating blocks of coordinates as representing elements in an extensions field. We do this for every possible choice of basis to ensure that our theorem works for all surjective $\mathbb{F}_2$-linear maps. We only treat the case of $\mathbb{F}_2$ as this is the field most commonly used in applications (the same proof strategy will work for any finite field).

Theorem 2.2. *Let $S \subset \mathbb{F}_2^n$ be such that $|S| > 2^{20}\max(n^4(1+\tau)^4/(\tau\delta)^8, n^4)$ and let n, τ, δ satisfy $n \geq 5\lceil \log_2(\max(n(1+\tau)/(\tau\delta)^2, n))\rceil + 25$. Then there exists a natural number*

$$t \geq \log_2|S| - 4\log_2\left(\max\left(\frac{n(1+\tau)}{(\tau\delta)^2}, n\right)\right) - 20,$$

such that a $1-\delta$ fraction of all surjective linear maps $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ are such that $L(U_S)$ is $\tau 2^{-t}$-close to uniform in the ℓ_∞ norm.

When $|S|$ is small we can improve the previous theorem by replacing the n in the entropy loss by $\log_2|S|$. This is achieved using the following simple lemma, which allows us to first hash S into a universe of size roughly $|S|^2$ without any collisions.

Lemma 2.3. *Let $S \subset \mathbb{F}_2^n$ and*

$$t \geq \log_2(|S|(|S|-1)/2\delta).$$

Then, at least a $1-\delta$ fraction of all surjective linear maps $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ map S injectively into $\mathbb{F}_2^t$.

Proof. As surjective linear maps are a universal family of hash functions we have,

$$\mathbf{Pr}[L(x) = L(y)] \leq 1/2^t \leq \delta\frac{2}{|S|(|S|-1)}$$

for a random surjective linear map $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ and $x, y \in S, x \neq y$. By applying the union bound we see the probability that L is not injective is upper bounded by δ. □

Applying the above lemma followed by Theorem 2.2 immediately leads to a concrete instance of Theorem 1.2.

Theorem 2.4. *Let $S \subset \mathbb{F}_2^n, \tau, \delta \in (0,1)$ and $m = \log_2(|S|(|S|-1)/\delta)$ be such that*

$$|S| > 2^{20} m \max(2^8(1+\tau)^4/(\tau\delta)^8, 1) \tag{1}$$

$$m \geq 5\log_2(m\max(4(1+\tau)/(\tau\delta)^2, 1)) + 25, \tag{2}$$

then there exists a natural number

$$t \geq \log_2 |S| - 4\log_2\left(m \max\left(\frac{4(1+\tau)}{(\tau\delta)^2}, 1\right)\right) - 20,$$

such that a $1-\delta$ *fraction of all surjective linear maps* $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ *are such that* $L(U_S)$ *is* $\tau 2^{-t}$*-close to uniform in the* ℓ_∞ *norm.*

Proof. We apply Lemma 2.3 for $\delta/2$ and linear maps from $\mathbb{F}_2^n \to \mathbb{F}_2^m$ followed by applying Theorem 2.2 for $\delta/2$ and linear maps from $\mathbb{F}_2^m \to \mathbb{F}_2^t$. □

The conditions (1) and (2) are not very restrictive. In the setting $\tau = \delta = 1/n^C$ for some constant C conditions (1) and (2) are satisfied for $|S| \geq n^{C'}$ where C' only depends on C.

An interesting setting of parameters for Theorem 2.4 is that of $\tau = 1/\delta^2$. In this case (when δ is sufficiently small), the two terms in the 'max' function above are about the same and we get an entropy loss of $4\log_2(4\log_2(|S|(|S|-1)/\delta))$. With this entropy loss, we get that the output $L(U_S)$ is $(1/\delta)^2 \cdot 2^{-t}$ close to uniform in the ℓ_∞ norm. Or, in other words, for $1-\delta$ fraction of linear maps L, the probability of any event under $L(U_S)$ is at most a multiplicative factor of $1/\delta^2$ larger than its probability under the uniform distribution. In this setting (1) and (2) are satisfied by ensuring $|S|$ is larger than some fixed universal constant.

When $\tau > 1$ we can improve the constant in the above two theorems slightly. We start with the case of large finite field. In the following theorem, the bound on the size of q does not contain the constant 32 appearing in Theorem 2.1. The dependence of q on τ changes from $\frac{1+\tau}{\tau^2}$ to $\frac{1+\tau}{(\tau-\sqrt{\tau})^2}$ which are asymptotically the same when τ grows. Hence, when τ is sufficiently large, the saving in q is roughly a factor of 32. The price we pay for this improvement is the need for n to be at least 20 (as opposed to 5) and an upper bound $\delta < 1/10$.

Theorem 2.5. *Let* $n \geq 20$ *and let* $S \subset \mathbb{F}_q^n$ *be a set. Let* $\tau > 1$ *be a real number and* $\delta \in (0, 1/10)$ *s.t*

$$q \geq \max\left(n\frac{1+\tau}{(\tau-\sqrt{\tau})^2\delta^2}, n\right).$$

Suppose $q^r < |S| \leq q^{r+1}$ *for some* $4 \leq r \leq n-1$ *and let* $t = r-3$*. Then a* $1-\delta$ *fraction of all surjective linear maps* $L : \mathbb{F}_q^n \to \mathbb{F}_q^t$ *are such that* $L(U_S)$ *is* τ/q^t*-close to the uniform distribution in the* ℓ_∞ *norm.*

As before, this can be used to prove a version over $\mathbb{F}_2$ for large τ with improved constants.

Theorem 2.6. *Let* $S \subset \mathbb{F}_2^n$ *and let* $\delta \leq 1/10$*,* $\tau > 1$ *be such that* $|S| > \max(n^4(1+\tau)^4/((\tau - \sqrt{\tau})\delta)^8, n^4)$ *and* n, τ, δ *satisfy* $n \geq 20\lceil\log_2(\max(n(1+\tau)/((\tau-\sqrt{\tau})\delta)^2, n))\rceil$*. Then there exists a natural number*

$$t \geq \log_2 |S| - 4\log_2\left(\max\left(n\frac{1+\tau}{(\tau-\sqrt{\tau})^2\delta^2}, n\right)\right),$$

such that a $1-\delta$ *fraction of all surjective linear maps* $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ *have the property that* $L(U_S)$ *is* $\tau 2^{-t}$*-close to uniform in the* ℓ_∞ *norm.*

We can again use Lemma 2.3 to improve the entropy loss in the previous theorem.

Theorem 2.7. *Let $S \subset \mathbb{F}_2^n$ and let $\delta \leq 1/10$, $\tau > 1$ and $m = \log_2(|S|(|S|-1)/\delta)$ be such that*

$$|S| > m^4 \max(2^8(1+\tau)^4/((\tau-\sqrt{\tau})\delta)^8, 1)$$
$$m \geq 20\lceil \log_2(m \max(4(1+\tau)/((\tau-\sqrt{\tau})\delta)^2, 1))\rceil$$

. Then there exists a natural number

$$t \geq \log_2 |S| - 4\log_2\left(m \max\left(\frac{4(1+\tau)}{(\tau-\sqrt{\tau})^2\delta^2}, 1\right)\right),$$

such that a $1-\delta$ fraction of all surjective linear maps $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ have the property that $L(U_S)$ is $\tau 2^{-t}$-close to uniform in the ℓ_∞ norm.

2.1 Some comments

Tightness of our results: It is natural to ask whether our results are tight. Fixing the parameter δ to be constant for the sake of simplicity, can we possibly improve on the entropy loss stated in Theorem 2.4? The answer is a resounding No! Even for a truly random function, the results of [RS98] show that we need an entropy loss of at least $\log_2(\log_2 |S|/\tau^2)$ (up to a additive constant) to acheive the conclusion of Theorem 2.4. Hence, up to a reasonably small constant factor (of about 32), linear functions hash as well as random functions.

Prior results on linear hash functions: Properties of random linear hashes w.r.t. the ℓ_∞ norm have been studied in earlier works [CW79, MV84, ADM+99] with [ADM+99] being the state-of-the-art. The results in this area are typically stated as upper bounds on the expected 'maximal bucket size' (that is, the maximum size of $L^{-1}(y)$ over all $y \in \mathbb{F}_q^t$). We will see that earlier results only give bounds for $\tau \gg 1$ (as far as we know, our paper is the first to give ℓ_∞ guarantees for small τ).

Theorem 5 of [ADM+99] is the most relevant to this work and shows that, when $\log_2 |S| - t = \log_2(t)$ the expected maximal bucket size is $O(t \log_2(t))$. A Markov argument shows then, that, with probability at least $1-\delta$, the maximal bucket size is at most $O(t \log_2(t)/\delta)$ which is a factor of $\log_2(t)/\delta$ larger than the trivial bound of $|S|/2^t = t$. Note that $\log_2(t)/\delta \gg 1$.

Theorem 2.4 for small τ shows that when $\log_2 |S| - t \approx O(\log_2(\log_2(|S|^2/\delta)(\tau\delta)^{-2}))$, the maximal bucket size will be at most a factor of $1+\tau$ larger than the trivial bound of $|S|/2^t$ with probability $1-\delta$ over the choice of the linear function. Hence, the results of [ADM+99] deal with the case of smaller entropy loss ($\log_2(t) \approx \log_2(\log_2(|S|))$ instead of $O(\log_2(\log_2(|S|^2/\delta)(\tau\delta)^{-2}))$) but are a multiplicative factor of $\log_2(t)/\delta \gg 1$ away from uniform instead of $\tau + 1$ which can be made arbitrarily close to 1 (by reducing τ and increasing the entropy loss).

We can also make comparisons in the regime of large τ. As stated earlier for $\tau = 1/\delta^2$ Theorem 2.4 shows that the maximal bucket size will be at most a factor of $1 + 1/\delta^2$ larger than the trivial bound of $|S|/2^t$ with probability $1-\delta$ over the choice of the linear function. In this setting for $\delta \gg 1/\log_2(t)$, we lose a constant factor in the entropy loss ($\log_2 \log_2 |S|$ in [ADM+99] and $O(\log_2 \log_2 |S|)$ for our result) and gain in the bucket size bound ($\log_2(t)\delta$ times $|S|/2^t$ in [ADM+99] and $1+1/\delta^2$ times $|S|/2^t$ for our result). Although it should be noted that the results in [ADM+99] are incomparable in the sense that they compute the expected value of the bucket size while our results only give bounds on the bucket size with high probability.

Other families of universal hash functions: In this section we look at whether our results can hold for other universal families of hash functions.

We first show that our results can not hold for all families of universal hash functions by means of an example. The family we will consider is linear maps from $\mathbb{F}_{q^2}^2$ to $\mathbb{F}_{q^2}$ which do form a universal family. We will show that that known results from [ADM+99] prove that this family needs at least an entropy loss of $\Omega(\log_2|S|)$ to get the distance guarantees of Theorem 2.4. This also shows that we need high dimensionality to get good linear hash function over large fields.

Theorem 8 of [ADM+99] proves that for any finite field $\mathbb{F}_{q^2}$ where q is a prime power if we consider the set of linear maps from $\mathbb{F}_{q^2}^2$ to $\mathbb{F}_{q^2}$ then there exists a set S_0 of size q^2 such that for every linear map the maximal bucket size is at least q.

This implies that for any S_0' of size $q^{2+\eta}, \eta < 1$ which contains S_0, every linear map $L : \mathbb{F}_{q^2}^2 \to \mathbb{F}_{q^2}$ will have a maximal bucket size of at least q. In other words $L(U_{S_0'})$ will be at least $1/q^{1+\eta} \gg C/q^2$ away from uniform in ℓ_∞ distance. Equivalently, even for an entropy loss of $\eta\log_2(q) = \Omega(\log_2|S_0'|) \gg O(\log_2\log_2|S|)$, linear maps from $L : \mathbb{F}_{q^2}^2 \to \mathbb{F}_{q^2}$ do not guarantee that the image $L(U_{S_0'})$ will be C/q^2 close to uniform for any fixed constant C. This also means that we need at least an entropy loss of $\Omega(\log_2|S|)$ to get the distance guarantees of Theorem 2.4.

Other families of universal hash function could still achieve the guarantees of Theorem 2.4. In particular, for a prime p consider the family of hash functions $h_{a,b} : \{0,1,\ldots,p-1\} \to \{0,\ldots,m-1\}$ for $a \in \{1,\ldots,p-1\}, b \in \{0,\ldots,p-1\}$ defined as $h_{a,b}(x) = (ax+b \mod p) \mod m$. From [CW79], we know that this family is universal. By following the framework in Section 3, it can be checked that proving ℓ_∞-guarantees for this family is a generalization of the notoriously difficult Arithmetic Kakeya problem [GR19].

The case of high min-entropy: As was mentioned before, The LHL holds not just for 'flat' distributions of the form U_S, but for any distribution with high min-entropy. This more general version can be derived easily from the LHL for sets using a convex combination argument. As far as we can tell, this argument fails in the case of ℓ_∞ and so we cannot automatically derive a min-entropy analog of our results. While we do believe that our proof techniques could be made to handle this more general case (e.g., as is the case in [DDL21b]), we leave it for future work.

3 Connection to prior work on Kakeya and Furstenberg sets

In this section we will explain the connection between Theorem 2.1 and the finite field Kakeya/Furstenberg problem. Along the way we will introduce notations and definitions that will be used later on in the proofs.

We will now describe an equivalent formulation of Theorem 2.1 in terms of the kernel of the linear map $L : \mathbb{F}_q^n \to \mathbb{F}_q^t$ appearing in the theorem. This will allow us to highlight its connection to the finite field Kakeya problem. To do so, we introduce some notations. For $1 \le k \le n$ we denote by $\mathcal{L}_k(\mathbb{F}_q^n)$ the set of k-dimensional flats in $\mathbb{F}_q^n$ and by $\mathcal{L}_k^*(\mathbb{F}_q^n)$ the set of k-dimensional subspaces (flats passing through the origin). Let $S \subset \mathbb{F}_q^n$ be a set. For $k \in [n]$, we denote by

$$E_k(S) = |S|/q^{n-k}$$

the expectation of $|R \cap S|$ with R chosen uniformly in $\mathcal{L}_k(\mathbb{F}_q^n)$. When S is clear from the context we omit it and simply write E_k.

Definition 3.1. *We say that $R \in \mathcal{L}_k(\mathbb{F}_q^n)$ is τ-balanced w.r.t a set $S \subset \mathbb{F}_q^n$ if we have:*

$$||R \cap S| - E_k(S)| \leq \tau \cdot E_k(S).$$

Otherwise, we say that R is τ-unbalanced w.r.t S.

Definition 3.2. *We say that $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ is τ-shift-balanced w.r.t S if, for all $a \in \mathbb{F}_q^n$, the flat $R = A + a$ is τ-balanced w.r.t S.*

Notice that if $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ is τ-shift-balanced w.r.t S and $A' \in \mathcal{L}_{k'}^*(\mathbb{F}_q^n)$ contains A (with $k' > k$) then A' is also τ-shift-balanced w.r.t S.

We will now express Theorem 2.1 using this new notation. Suppose $L : \mathbb{F}_q^n \to \mathbb{F}_q^t$ is an onto linear map and let $A = \ker(L)$ be its $k = n - t$ dimensional kernel. Notice that, for each $y \in \mathbb{F}_q^t$,

$$\mathbf{Pr}[L(U_S) = y] = \frac{|(A + a) \cap S|}{|S|},$$

for some $a \in \mathbb{F}_q^n$ for which $L(a) = y$. Therefore,

$$|\mathbf{Pr}[L(U_S) = y] - q^{-t}| \leq \tau q^{-t},$$

if and only if $A+a$ is τ-balanced w.r.t S. Hence, Theorem 2.1 is equivalent to the following theorem.

Theorem 3.3. *Let $n \geq 5$ and let $S \subset \mathbb{F}_q^n$ be a set such that $|S| > q^4$. Let $\tau > 0, \delta \in (0, 1)$ be a real number s.t $q \geq 32 \max(n(1 + \tau)/(\tau\delta)^2, n)$. Let $4 \leq r \leq n - 1$ be an integer s.t $q^r < |S| \leq q^{r+1}$ and let $k = n - r + 3$. Then a $1 - \delta$ fraction of all subspaces in $\mathcal{L}_k^*(\mathbb{F}_q^n)$ are τ-shift-balanced w.r.t S.*

We now take a moment to explain the expression 'two-sided Kakeya bounds' from the title and the connection to prior work on Kakeya sets. A Kakeya set in $\mathbb{F}_q^n$ is a set containing a line in each direction. The main question, asked by Wolff in [Wol99], is to lower bound the size of such sets. This question has now been completely resolved in the series of papers [Dvi09, DKSS13, BB21]. We will be mostly interested in the high dimensional variants of this problem, asking about sets containing k-dimensional flats in all directions, or more generally, sets that have large intersection with a flat in each direction (these are called Furstenberg sets). These type of questions have been also studied extensively, with tight bounds obtained in some cases [EOT10, KLSS11, EE16, DDL21a, DDL21b].

We start by recalling some definitions from that domain.

Definition 3.4 (m-rich flats)**.** *We call a flat $R \in \mathcal{L}_k(\mathbb{F}_q^n)$ m-rich w.r.t a set $S \subset \mathbb{F}_q^n$ if $|R \cap S| \geq m$.*

Definition 3.5 ((k, m, β)-Furstenberg sets)**.** *We call a set $K \subset \mathbb{F}_q^n$ a (k, m, β)-Furstenberg set if K has an m-rich k-flat for at least a β fraction of directions. That is, for at least a β-fraction of all $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ there exists $a \in \mathbb{F}_q^n$ so that $a + A$ is m-rich w.r.t K.*

Prior works on Kakeya/Furstenberg sets were focused on giving *lower bounds* on the size of $(k, m, 1)$-Furstenberg sets. For example, in [DDL21b], it was shown that, if S is a $(k, m, 1)$-Furstenberg set then $|S| > (1 - \epsilon)mq^{n-k}$, assuming q is sufficiently large as a function of n and ϵ (in particular, q has to be exponential in n). Notice that this is the best possible since any set of size mq^{n-k} is $(k, m, 1)$-Furstenberg. Stated in the counter-positive direction, this theorem shows that: If

$$|S| \leq (1 - \epsilon)mq^{n-k} \tag{3}$$

then there exists a k-dimensional subspace R such that all shifts of R have less than m-points in common with S. Notice that (3) gives us that

$$E_k(S) \leq (1-\epsilon)m.$$

So, what we discover is that, the results in [DDL21b] simply say that, for every S, there is a subspace R such that all shifts of R have intersection with S that is not much larger from the expectation E_k. Hence, Theorem 3.3 can be viewed as a two-sided generalization of this statement by showing that, in fact, there exists R such that all shifts of R have roughly the expected intersection with S.

4 Proof Overview

We now give a short sketch of the proof of Theorem 2.1. The proof of Theorem 2.5 (the case of $\tau > 1$) is essentially the same as the proof of Theorem 2.1 with a different setting of a single parameter and so we will not discuss it here. We will also not discuss the two theorems dealing with the case of $\mathbb{F}_2$ as they will follow from the large field case by a simple encoding argument.

As discussed in Section 3, Theorem 2.1 is equivalent to Theorem 3.3 which is stated in the language of shift-balanced sub-spaces. Given a set $S \subset \mathbb{F}_q^n$, the theorem claims that there are many sub-spaces $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ that are τ-shift-balanced. Let us instead try and prove the easier claim that *there exists* at least one such subspace. We will prove this by contradiction. Suppose there are no τ-shift balanced sub-spaces A. Then, for each $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ we can find a shift $f(A) \in \mathbb{F}_q^n$ so that the flat $T_A = f(A) + A$ is τ-unbalanced.

At a very high level, the contradiction will follow by combining the following three statements:

- (Concentration Statement) A random $k-2$ flat is $\tau/2$-balanced with high probability.
- (Anti concentration statement) If T is a τ-unbalanced k-flat and R is a randomly chosen $k-2$-flat in T then R is $\tau/2$-unbalanced with high probability.
- (Kakeya statement) Given a collection of k-flats, T_A, one in each direction $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$. A randomly chosen $k-2$-flat in a randomly chosen T_A 'behaves like' a truly random $k-2$-flat.

Before we discuss the proofs of these statements, let us see how they can be combined to derive a contradiction. Consider the distribution on $k-2$-flats obtained by sampling $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$ uniformly at random and then choosing a random $k-2$-flat R inside $f(A)+A$, where $f(A)$ is defined above so that $f(A)+A$ is τ-unbalanced. By the anti-concentration statement, this distribution outputs a $\tau/2$-unbalanced R with high probability. Now, from the Kakeya statement we get that this should (in some way) also be the behaviour of a truly random $k-2$-flat, contradicting the concentration statement. This is essentially the structure of the proof, with the 'behaves like' portion of the Kakeya statement replaced by a quantitative bound on the probability of landing in a given small set (the set of unbalanced $k-2$ flats).

Let us now discuss the proofs of the three statements. The first two (concentration and anti-concentration), follow easily from Chebyshev's inequality and pair-wise independence and so we will only be concerned with the proof of the third one. We can generalize the Kekeya statement as follows, given a collection of k-flats T_A, one in each direction A, what can be said about the distribution of a random r-flat R in a random T_A where we allow r to be in the range $\{0, 1, \ldots, k\}$. To recover the original (one dimensional) Kakeya problem all we have to do is set $k = 1$ and $r = 0$.

Now, we are asking about the distribution of a random point R on a line T_A chosen so that its direction is uniformly random and its shift is arbitrary. The finite field Kakeya conjecture (proved in [Dvi09]) says that the distribution of R has large support. In [DW08, DKSS13], motivated by applications to extractors, it was shown that, in fact, the distribution of R has high min-entropy. These results can be easily 'lifted up' to the case where $k > 1$ and $r = k - 1$ but, alas, the known (and tight) quantitative bounds on the min entropy are not sufficient for our purposes. Specifically, it is possible for the distribution of R in this case to be contained in a set of density 2^{-n} inside $\mathbb{F}_q^n$, which is much too small for our purposes. This motivates us to take $r = k - 2$, which reduces to understanding the case of $k = 2$ and $r = 0$. That is, given a family of 2-flats T_A, one in each direction, what can be say about the behaviour of a random point R on a random T_A? Luckily, in this case, the results of [KLSS11, DDL21b] can be used to show that the distribution of R has support with density approaching one.

To prove our theorem we need to extend the results of [KLSS11, DDL21b] in several ways, including going from support size to min entropy, reducing the field size from exponential to polynomial and handling the case of 'many' directions in stead of 'all' (which corresponds to the parameter δ being less than one). The required lemma is stated below and proved in Section 6.

Lemma 4.1 (Furstenberg lemma)**.** *For any $\gamma, \beta \in [0,1], n \in \mathbb{N}$, q a prime power every $(2, \gamma q^2, \beta)$-Furstenberg set $K \subseteq \mathbb{F}_q^n$ has size at least,*

$$|K| \geq \beta \gamma^n q^n \left(1 + \frac{1}{q}\right)^{-n}.$$

Note, this lemma has been proven in [ORW22] with a slightly worse lower bound of $\beta\gamma^n q^n \left(1 + \frac{2}{q}\right)^{-n}$ which is enough to prove Theorem 2.1 leading to slightly worse constants in the field size requirement and hence the entropy loss of the theorem. The proof in [ORW22] uses a combinatorial reduction to reduce the case of arbitrary β to constant β. We give a new argument to prove this lemma directly.

Our proof of this lemma follows along the lines of prior works in this area and uses the polynomial method. One important ingredient is a new variant of the celebrated Schwartz-Zippel lemma which allows us to improve the dependence on β above from β^n to just β (See Corollary 6.12). We believe this lemma could have applications in other situations where the polynomial method is used. For instance in an upcoming work [Dha] extensions of these arguments will be used to prove maximal Kakeya bounds in the general setting of the integers modulo a composite number.

5 Proof of Theorems 2.1 and 2.5

We prove Theorems 2.1 and 2.5 by contradiction. We will prove the equivalent versions of the theorems stated using τ-shift-balanced subspaces (Theorem 3.3 and similarly for Theorem 2.5 even though it was not stated separately). Suppose the Theorems are not true. Then there exists a function with parameters as in the Theorems:

$$f : \mathcal{L}_k^*(\mathbb{F}_q^n) \to \mathbb{F}_q^n$$

such that, for a δ fraction of $A \in \mathcal{L}_k^*(\mathbb{F}_q^n)$, the flat $f(A) + A$ is τ-unbalanced w.r.t. S. Notice that $f(A)$ can be taken to be any point on the flat $f(A) + A$ (the choice doesn't matter for this proof).

For a real number $\sigma > 0$, let

$$B^{\sigma}_{k-2} \subset \mathcal{L}_{k-2}(\mathbb{F}^n_q)$$

denote the set of $k-2$-flats that are σ-unbalanced w.r.t. S. We will eventually set σ to one of two values: To prove Theorem 2.1 we will set $\sigma = \tau/2$ and, to prove Theorem 2.5 (when $\tau > 1$) we will set $\sigma = \sqrt{\tau}$. Notice that, in both cases, we have $\tau - \sigma > 0$.

For a k-flat $T \in \mathcal{L}_k(\mathbb{F}^n_q)$ we let $\mathcal{L}_{k-2}(T)$ be the set of $k-2$-flats contained in T and let

$$B^{\sigma}_{k-2}(T) = B^{\sigma}_{k-2} \cap \mathcal{L}_{k-2}(T)$$

denote the set of σ-unbalanced $k-2$-flats w.r.t. S that are contained in T.

Notice first that, by our assumption on r, we have

$$4 \leq k \leq n-1 \tag{4}$$

Our first claim shows that a random $k-2$ flat is balanced w.h.p. This gives the 'concentration' part of the argument laid out in the proof overview.

Claim 5.1. *If R is chosen uniformly in $\mathcal{L}_{k-2}(\mathbb{F}^n_q)$ then*

$$\mathbf{Pr}[R \in B^{\sigma}_{k-2}] \leq \frac{1}{\sigma^2 q}.$$

Proof. Since $k \geq 3$ we can use pairwise independence and Chebyshev. The probability that $|R \cap S|$ deviates from its expectation E_{k-2} by at least σE_{k-2} is at most

$$\frac{\mathbf{V}(|R \cap S|)}{(\sigma E_{k-2})^2} \leq \frac{1}{\sigma^2 E_{k-2}} \leq \frac{1}{\sigma^2 q},$$

where we use the fact that $E_{k-2} = |S|/q^{n-k+2} \geq q$ for $k = n - r + 3$. □

The next claim gives the 'anti concentration' part of the proof overview, showing that a random $k-2$-flat in an unbalanced k-flat is unbalanced w.h.p.

Claim 5.2. *Let $T \in \mathcal{L}_k(\mathbb{F}^n_q)$ be τ-unbalanced w.r.t S. Suppose R is chosen uniformly at random from $\mathcal{L}_{k-2}(T)$. Then*

$$\mathbf{Pr}[R \in B^{\sigma}_{k-2}(T)] \geq 1 - \frac{1+\tau}{(\tau-\sigma)^2 q}.$$

Proof. As before, the size of $R \cap S$ is a sum of p.w independent indicator variables with expectation:

$$\mathbb{E}[|R \cap S|] = \frac{|S \cap T|}{|T|} q^{k-2} = |S \cap T|/q^2. \tag{5}$$

Since T is τ-unbalanced, we have that

$$||S \cap T| - E_k| \geq \tau E_k. \tag{6}$$

Therefore, dividing by q^2 and using (5) we have that

$$|\mathbb{E}[|R \cap S|] - E_{k-2}| \geq \tau E_{k-2}. \tag{7}$$

We will separate into two cases: case 1 is when

$$\mathbb{E}[|R \cap S|] \leq (1-\tau)E_{k-2}. \tag{8}$$

In this case (which can only happen if $\tau < 1$), using Chebyshev, the probability that R is σ-balanced is bounded from above by,

$$\begin{aligned} \mathbf{Pr}[|R \cap S| - E_{k-2} \geq -\sigma E_{k-2}] &\leq \\ \mathbf{Pr}[|R \cap S| - \mathbb{E}[|R \cap S|] \geq (\tau-\sigma)E_{k-2}] &\leq \\ \frac{\mathbf{V}(|R \cap S|)}{(\tau-\sigma)^2 E_{k-2}^2} \leq \frac{\mathbb{E}(|R \cap S|)}{(\tau-\sigma)^2 E_{k-2}^2} &\leq \frac{1-\tau}{(\tau-\sigma)^2 q}. \end{aligned}$$

In the second case we have,

$$\mathbb{E}[|R \cap S|] \geq (1+\tau)E_{k-2}.$$

In this case the probability that R is σ-balanced is bounded above by,

$$\begin{aligned} \mathbf{Pr}[|R \cap S| - E_{k-2} \leq \sigma E_{k-2}] \leq \\ \mathbf{Pr}\left[||R \cap S| - \mathbb{E}[|R \cap S|]| \geq \mathbb{E}[|R \cap S|] - (1+\sigma)E_{k-2}\right] \leq \\ \mathbf{Pr}\left[||R \cap S| - \mathbb{E}[|R \cap S|]| \geq \mathbb{E}[|R \cap S|] \cdot \frac{\tau-\sigma}{1+\tau}\right] \leq \\ \frac{\mathbf{V}(|R \cap S|)}{(\tau-\sigma)^2/(1+\tau)^2 \mathbb{E}[|R \cap S|]^2} \leq \frac{(1+\tau)^2}{(\tau-\sigma)^2 \mathbb{E}[|R \cap S|]} \leq \frac{1+\tau}{(\tau-\sigma)^2 E_{k-2}} \leq \frac{1+\tau}{(\tau-\sigma)^2 q}. \end{aligned}$$

Hence, the probability that R is σ balanced is bounded by $(1+\tau)/((\tau-\sigma)^2 q)$ and so we are done. □

We next define three important sets:

- $(\mathcal{L}^*_{k-2}(T))$: For $T \in \mathcal{L}_k(\mathbb{F}_q^n)$, we define $\mathcal{L}^*_{k-2}(T)$ to be the set of subspaces in $\mathcal{L}^*_{k-2}(\mathbb{F}_q^n)$ which on translation can lie in T (or equivalently are parallel to T).
- $(\mathcal{L}_{k-2}(T, \parallel W))$: For $T \in \mathcal{L}_k(\mathbb{F}_q^n)$ and $W \in \mathcal{L}^*_{k-2}(T)$ we let $\mathcal{L}_{k-2}(T, \parallel W)$ be the set of $k-2$ flats in T which are parallel to W (notice that there are exactly q^2 such flats and that their disjoint union is T).
- $(C^{\sigma,c}_{k-2}(T))$: For $T \in \mathcal{L}_k(\mathbb{F}_q^n)$ we let $C^{\sigma,c}_{k-2}(T)$ be the set of flats W in $\mathcal{L}^*_{k-2}(T)$ such that at least a $1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}$ fraction of the flats in $\mathcal{L}_{k-2}(T, \parallel W)$ are in $B^{\sigma}_{k-2}(T)$ (we will set $c \geq 1$ to two different values for Theorem 2.1 and Theorem 2.5).

The previous lemma can now be used to prove that, if T is unbalanced, then many W's are in fact in the set $C^{\sigma,c}_{k-2}(T)$ defined above (this is essentially a Markov style averaging argument).

Claim 5.3. *Let $T \in \mathcal{L}_k(\mathbb{F}_q^n)$ be τ-unbalanced w.r.t S. Suppose W is chosen uniformly at random from $\mathcal{L}^*_{k-2}(T)$. Then*

$$\mathbf{Pr}[W \in C^{\sigma,c}_{k-2}(T)] \geq 1 - 1/c.$$

Proof. Let us say the claim is false then with probability less than $1 - 1/c$, $W \in C^{\sigma,c}_{k-2}(T)$ for a uniformly random $W \in \mathcal{L}^*_{k-2}(T)$. Equivalently, with probability greater than $1/c$, $W \notin C^{\sigma,c}_{k-2}(T)$. We can sample a uniformly random chosen $R \in \mathcal{L}_{k-2}(T)$ by first picking a direction $W \in \mathcal{L}^*_{k-2}(T)$ at random and then taking R to be a random shift of W inside T. The above assumption will then give us that:

$$\begin{aligned}\mathbf{Pr}[R \in B^{\sigma}_{k-2}(T)] &\leq \mathbf{Pr}[W \notin C^{\sigma,c}_{k-2}(T)]\left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right) + \mathbf{Pr}[W \in C^{\sigma,c}_{k-2}(T)] \\ &\leq 1 - \mathbf{Pr}[W \notin C^{\sigma,c}_{k-2}(T)]\frac{c(1+\tau)}{(\tau-\sigma)^2 q} < 1 - \frac{1+\tau}{(\tau-\sigma)^2 q}.\end{aligned}$$

This contradicts Claim 5.2. □

Given $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$, let $\mathcal{L}_{k-2}(\| W)$ be the set of $k-2$ flats parallel to W and $\mathcal{L}^*_k(\| W)$ be the set of k-dimensional subspaces containing W. Let

$$B^{\sigma}_{k-2}(\| W) = B^{\sigma}_{k-2} \cap \mathcal{L}_{k-2}(\| W)$$

denote the set of σ-unbalanced flats parallel to W.

The next claim shows that there is a 'good' choice of $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ to which we should restrict our attention (that is, we will consider only $k-2$ flats parallel to W).[4] This W should preserve the typical behavior of a random W in two respects: one is that B^{σ}_{k-2} should still have low density when restricted to flats parallel to W. The other is that W hits $C^{\sigma,c}_{k-2}(f(A)+A)$ for many $A \in \mathcal{L}^*_k(\| W)$.

Claim 5.4. *There exists $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ such that*

1. $|B^{\sigma}_{k-2}(\| W)| \leq \frac{1}{\sigma^2 q(1-\sqrt{1-\delta(1-1/c)})} \cdot |\mathcal{L}_{k-2}(\| W)| \leq \frac{2c}{(c-1)\sigma^2\delta} q^{n-k+1}$
2. $\mathbf{Pr}_{A \sim \mathcal{L}^*_k(\| W)}[W \in C^{\sigma,c}_{k-2}(f(A)+A)] \geq 1 - \sqrt{1-\delta(1-1/c)} \geq \frac{(c-1)\delta}{c+c\sqrt{1-\delta/2}} \geq \frac{\delta(c-1)}{2c}$.

Proof. Notice, that B^{σ}_{k-2} is a disjoint union of $B^{\sigma}_{k-2}(\| W)$ over all $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$. Suppose W is chosen uniformly at random from $\mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$. Then, by Claim 5.1 and Markov, we have that the probability that W does not satisfy 1. above is less than $1 - \sqrt{1-\delta(1-1/c)}$.

Consider the bi-partite graph G between $\mathcal{L}^*_k(\mathbb{F}^n_q)$ and $\mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ where the edges correspond to pairs $(A, W) \in \mathcal{L}^*_k(\mathbb{F}^n_q) \times \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ such that $W \subset A$. Let μ be the distribution over the pairs $(A, W) \in \mathcal{L}^*_k(\mathbb{F}^n_q) \times \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ which is uniform over the edges of G. As the graph G is regular on both sides sampling from μ is equivalent to sampling A uniformly from $\mathcal{L}^*_k(\mathbb{F}^n_q)$ and W uniformly from $\mathcal{L}^*_{k-2}(A)$. It also is equivalent to uniformly sampling $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ and sampling A from $\mathcal{L}^*_k(\| W)$. By Claim 5.3 and the fact that at least for a δ fraction of $A \in \mathcal{L}^*_k(\mathbb{F}^n_q)$, $f(A)+A$ is τ-unbalanced we have,

$$\mathbf{Pr}_{(A,W)\sim\mu}[W \in C^{\sigma,c}_{k-2}(f(A)+A)] \geq \delta(1-1/c).$$

By averaging, again, we get that the probability that W fails to satisfy

$$\mathbf{Pr}_{A \sim \mathcal{L}^*_k(\| W)}[W \notin C^{\sigma,c}_{k-2}(f(A)+A)] \leq \sqrt{1-\delta(1-1/c)}$$

is at most $\sqrt{1-\delta(1-1/c)}$. By a union bound we now see that there exists a $W \in \mathcal{L}^*_{k-2}(\mathbb{F}^n_q)$ which satisfies the two properties in the claim. □

[4] This part of the proof corresponds to the statement in the proof overview arguing that the case of general k and r can be reduced to the case of $r = 0$ and $k \mapsto k - r$.

Fix $W = \hat{W}$ satisfying the two numbered items of Claim 5.4. Let $G_{\hat{W}}$ be the random variable which outputs the random 2-flat

$$f(\mathsf{span}\{U, \hat{W}\}) + U$$

for uniformly random $U \in \mathcal{L}_2^*(\mathbb{F}_q^n)$. Notice that there is a small probability that $\mathsf{span}\{U, \hat{W}\}$ is not k dimensional. In this case we set $f(\mathsf{span}\{U, \hat{W}\}) = 0$.

We will now show that $G_{\hat{W}}$ has large intersections with a small set with high probability. The particular structure of the random variable $G_{\hat{W}}$ allows us to state this using the notion of a Furstenberg set.

Claim 5.5. *There exists a set $K \subset \mathbb{F}_q^n$ such that*

1. $|K| \leq \frac{2c}{(c-1)\sigma^2\delta} q^{n-1}$.
2. *K is $\left(2, \left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right) q^2, \delta \frac{c-1}{2c}(1 - 1/q - 1/q^2)\right)$-Furstenberg.*

Proof. We take

$$K = \bigcup_{R \in B_{k-2}^{\sigma}(\| \hat{W})} R,$$

to be the union of all σ-unbalanced $k-2$-flats parallel to $\hat{W}$. To show that 1. holds, we use the first item of Claim 5.4 and the fact that each R has q^{k-2} points.

For a uniformly random $U \in \mathcal{L}_2^*(\mathbb{F}_q^n)$, $F = f(\mathsf{span}(U, \hat{W})) + U$ gives us a sample from $G_{\hat{W}}$. Let $T = f(\mathsf{span}(U, \hat{W})) + \mathsf{span}(U, \hat{W})$. If $\hat{W} \in C_{k-2}^{\sigma,c}(T)$ then that means at least

$$\left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right) q^2$$

many flats in $\mathcal{L}_{k-2}(T, \| \hat{W})$ are in $B_{k-2}^{\sigma}(T)$ and hence contained in K. F will intersect with each of these flats (and hence K) in distinct points. This is because T is a $f(\mathsf{span}(U, \hat{W}))$-shift of the span of U and $\hat{W}$ and $U \cap \hat{W} = \{0\}$ so T is a disjoint union of shifts of $\hat{W}$ by elements in $F = f(\mathsf{span}(U, \hat{W})) + U$. This implies that F is $(1 - c(1+\tau)/(\tau-\sigma)^2 q)q^2$-rich w.r.t. K. Finally, note that conditioned on the event that $\mathsf{span}(U, \hat{W})$ is k-dimensional T has the same distribution as $f(A) + A$ where A is uniformly distributed over $\mathcal{L}_k^*(\| \hat{W})$. This means

$$\mathbf{Pr}\left[G_{\hat{W}} \text{ is } \left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right) q^2\text{-rich w.r.t } K\right] \geq$$
$$\mathbf{Pr}_{A \sim \mathcal{L}_k^*(\| \hat{W})}\left[\hat{W} \in C_{k-2}^{\sigma,c}(f(A) + A)\right] \cdot \mathbf{Pr}_{U \in \mathcal{L}_2^*(\mathbb{F}_q^n)}\left[\dim \mathsf{span}\{U, \hat{W}\} = k\right].$$

We note $\mathbf{Pr}_{U \in \mathcal{L}_2^*(\mathbb{F}_q^n)}[\dim \mathsf{span}\{U, \hat{W}\} = k]$ is at least $1 - 1/q - 1/q^2$. If we generate U by picking two random vectors then the first one being in U has probability at most $1/q^{n-k+2} \leq 1/q^2$ and the second being in the space spanned by the first and U has probability at most $1/q^{n-k+1} \leq 1/q$. Now using Claim 5.3 and the equation above we have,

$$\mathbf{Pr}\left[G_{\hat{W}} \text{ is } \left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right) q^2\text{-rich}\right] \geq \frac{(c-1)\delta}{2c}\left(1 - \frac{1}{q} - \frac{1}{q^2}\right).$$

The above equation implies 2. as $G_{\hat{W}}$ by definition takes a uniformly chosen $U \in \mathcal{L}_2^*(\mathbb{F}_q^n)$ and outputs a flat parallel to U. ∎

To finish the proof of the theorem we need a bound for $(2, \gamma q^2, \beta)$-Furstenberg Sets. We will use Lemma 4.1 which we prove in the next section. We restate the Lemma here for convenience.

Lemma 4.1. *For any $\gamma, \beta \in [0,1], n \in \mathbb{N}$, q a prime power every $(2, \gamma q^2, \beta)$-Furstenberg set $K \subseteq \mathbb{F}_q^n$ has size at least,*

$$|K| \geq \beta \gamma^n q^n \left(1 + \frac{1}{q}\right)^{-n}.$$

Given this lemma, we substitute the values

$$\gamma = \left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right), \beta = \frac{\delta(c-1)}{2c}\left(1 - \frac{1}{q} - \frac{1}{q^2}\right)$$

and the bound on $|K|$ given by Claim 5.5 into the lemma above. We get the bound,

$$\frac{2c}{(c-1)\sigma^2\delta} q^{n-1} \geq |K| \geq q^n \left(1 - \frac{c(1+\tau)}{(\tau-\sigma)^2 q}\right)^n \left(1 + \frac{1}{q}\right)^{-n} \frac{\delta(c-1)}{2c}\left(1 - \frac{1}{q} - \frac{1}{q^2}\right). \tag{9}$$

To prove Theorem 2.1 use $q \geq 32\max(n(1+\tau)/(\tau\delta)^2, n)$, $c = 4$, $\sigma = \tau/2$ and $\delta \leq 1$ in (9) and re-arrange to get,

$$\frac{8}{9n} \geq \left(1 - \frac{1}{2n}\right)^n \left(1 + \frac{1}{32n}\right)^{-n} \left(1 - \frac{1}{32n} - \frac{1}{32^2 n^2}\right).$$

Using $(1 - x/n)^n \geq e^{-x}(1 - x^2/n)$ for $x < n$, $(1 + x/n)^n \leq e^x$ and $n \geq 5$ then implies,

$$\frac{8}{45} > e^{-1/5}(1 - 1/20)e^{-1/32}(1 - 1/160 - 1/(160)^2)$$

which leads to a contradiction proving Theorem 2.1.

To prove Theorem 2.5 use $q \geq \max(n(1+\tau)/(\tau - \sqrt{\tau})^2\delta^2, n)$,$\sigma = \sqrt{\tau}$, $\delta \leq 1/10$ and set $c = 10$ in (9) to get,

$$\frac{400}{81n} \geq \frac{400(\tau - \sqrt{\tau})^2}{81\tau(\tau+1)n} \geq \left(1 - \frac{1}{10n}\right)^n \left(1 + \frac{1}{n}\right)^{-n} \left(1 - \frac{1}{n} - \frac{1}{n^2}\right).$$

Using $(1 - x/n)^n \geq e^{-x}(1 - x^2/n)$ for $x < n$, $(1 + x/n)^n \leq e^x$ and $n \geq 20$ gives us,

$$\frac{400}{81 \cdot 20} > e^{-1/10}(1 - 1/(100 \cdot 20))e^{-1}(1 - /20 - 1/400)$$

which leads to a contradiction proving Theorem 2.5. □

5.1 The case of $\mathbb{F}_2$

In this section we prove Theorem 2.2 using Theorem 2.1. The same argument can be used to derive Theorem 2.6 from Theorem 2.5.

Theorem 2.2. *Let $S \subset \mathbb{F}_2^n$ be such that $|S| > 2^{20}\max(n^4(1+\tau)^4/(\tau\delta)^8, n^4)$ and let n, τ, δ satisfy $n \geq 5\lceil \log_2(\max(n(1+\tau)/(\tau\delta)^2, n))\rceil + 25$. Then there exists a natural number*

$$t \geq \log_2 |S| - 4\log_2(\max(n(1+\tau)/(\tau\delta)^2, n)) - 20,$$

such that a $1 - \delta$ fraction of all surjective linear maps $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$ have the property that $L(U_S)$ is $\tau 2^{-t}$-close to uniform in the ℓ_∞ norm.

Proof. Take

$$\ell = \lceil \log_2(32 \max(n(1+\tau)/(\tau\delta)^2, n)) \rceil.$$

and set

$$q = 2^\ell.$$

Let

$$n' = \lceil n/\ell \rceil$$

so that we have

$$2^n \leq q^{n'}.$$

We now identify $\mathbb{F}_2^n$ with an $\mathbb{F}_2$-linear subspace of $\mathbb{F}_q^{n'}$, e.g., by identifying $\mathbb{F}_q^n$ with $\mathbb{F}_2^{n\ell}$ as $\mathbb{F}_2$-vector spaces and then identifying $\mathbb{F}_2^n$ with the first $n \leq n'\ell$ coordinates. The above embedding of $\mathbb{F}_2^n$ in $\mathbb{F}_q^{n'}$ allows us to think of the set S as sitting in $\mathbb{F}_q^{n'}$ and so we can apply Theorem 2.1 if we check that all the conditions are met. We first see that, by our choice of ℓ, the bound on $q \geq 32 \max(n'(1+\tau)/(\tau\delta)^2, n')$ is met (notice that $n' \leq n$). We also need to check that $|S| > q^4$ which holds from our assumption $|S| \geq 2^{20} \max(n^4(1+\tau)^4/(\tau\delta)^8, n^4)$. $n' \geq 5$ is also satisfied.

Hence we can apply Theorem 2.1 in our setting. Let r be such that

$$q^r < |S| \leq q^{r+1}$$

and set

$$t' = r - 3.$$

We get that for a $1-\delta$ fraction of all surjective linear maps $L' : \mathbb{F}_q^{n'} \to \mathbb{F}_q^{t'}$ satisfy the property that $L'(U_S)$ is $\tau q^{-t'}$-close to uniform in the ℓ_∞ distance. Since an $\mathbb{F}_q$-linear map is also an $\mathbb{F}_2$-linear map, we can think of L' as an $\mathbb{F}_2$-linear map from $\mathbb{F}_2^{n'\ell}$ to $\mathbb{F}_2^{t'\ell}$. Setting

$$t = t'\ell$$

and let L be the restriction of L' to the subspace we previously identified with $\mathbb{F}_2^n$ (which contains S) we get that for any such $L : \mathbb{F}_2^n \to \mathbb{F}_2^t$, $L(U_S)$ is $\tau 2^{-t}$-close to uniform in the ℓ_∞ distance (clearly $L(U_S)$ and $L'(U_S)$ have the same distribution).

We now bound the 'entropy loss' or $\log_2 |S| - t$. Notice that

$$\log_2 |S| \leq (r+1)\ell$$

and that

$$t = t'\ell = (r-3)\ell.$$

Combining the last two inequalities we get that

$$\log_2 |S| - t \leq 4\ell \leq 4 \log_2(\max(n(1+\tau)/(\tau\delta)^2, n) - 20.$$

We are not done yet as not all surjective linear maps from $\mathbb{F}_2^n$ to $\mathbb{F}_2^t$ will be restrictions of surjective linear maps from $\mathbb{F}_q^{n'}$ to $\mathbb{F}_q^{t'}$. We now overcome this obstacle using a random rotation argument. We started out with embedding $S \subseteq \mathbb{F}_2^n$ in a bigger space $\mathbb{F}_2^{n'\ell}$. If we can show a $(1-\delta)$-fraction of surjective linear maps from $\mathbb{F}_2^{n'\ell}$ to $\mathbb{F}_2^t$ satisfy the desired property we are also done. We also let $\phi : \mathbb{F}_2^{n'\ell} \to \mathbb{F}_q^{n'}$ be the $\mathbb{F}_2$-linear isomorphism between $\mathbb{F}_2^{n'\ell}$ and $\mathbb{F}_q^{n'}$ we had implicitly chosen in the beginning.

Let H be the set of surjective linear maps from $\mathbb{F}_2^{n'\ell}$ to $\mathbb{F}_2^t$ which are also surjective linear maps from $\mathbb{F}_q^{n'}$ to $\mathbb{F}_q^{t'}$. We just showed that a $1-\delta$ fraction of the maps in H satisfy the desired property. Let M be a random invertible linear map in $\mathrm{GL}_{n'\ell}(\mathbb{F}_2)$. We note $\phi \circ M$ is also a valid $\mathbb{F}_2$-linear isomorphism between $\mathbb{F}_2^n n'\ell$ and $\mathbb{F}_q^{n'}$. If we repeated our earlier argument with this isomorphism we will have proven that a $1-\delta$ fraction of the maps in $M \cdot H = \{L \circ M | L \in H\}$ satisfy the desired property. But under a random rotation we see that each surjective linear map from $\mathbb{F}_2^{n'\ell}$ to $\mathbb{F}_2^t$ will be included in an equal number of $M \cdot H$. This proves that there is at least a $1-\delta$ fraction of surjective linear maps from $\mathbb{F}_2^{n'\ell}$ to $\mathbb{F}_2^t$ which satisfy the desired property. □

6 Proof of Lemma 4.1 using the polynomial method

We will be using the polynomial method to lower bound the sizes of $(2, \gamma q^2, \beta)$-Furstenberg sets in $\mathbb{F}_q^n$ which are needed to prove our hashing guarantees. As stated earlier, these bounds have been proven in [ORW22] using a combinatorial reduction. The bounds from [ORW22] can be directly used to prove our hashing theorems with slightly worse constants.

We will give a new proof to lower bound these set sizes by extending ideas developed in [Dha21] to prove bounds for Kakeya sets over rings of integers modulo a composite number. The advantages are three fold: we get slightly better constants, the argument here gives significantly better bounds for $(1, \gamma q, \beta)$-Furstenberg sets (although not important for our application) and as mentioned earlier these ideas will also be used to resolve the maximal Kakeya conjecture over rings of integers modulo a composite number in an upcoming work [Dha].

In this section we develop improvements to the polynomial method argument to get the desired dependence on β. In a nutshell, our improvement comes from picking a carefully chosen set of monomials, instead of just taking all monomials up to a specified degree. This section will be divided into three sub-sections. First, we review basic definitions and results on the polynomial method (with multiplicities) as developed in [DKSS13]. Then, we devote a section to understanding ranks of sub-matrices of a special matrix which maps a polynomial to its evaluations (with derivatives) on a given set of points. Finally, we put everything together to prove Lemma 4.1.

6.1 Multiplicities and Hasse derivative

We first review the definitions of multiplicities and Hasse derivatives that will be needed in the proof (see [DKSS13] for a more detailed discussion). We will allow the definitions to be over an arbitrary field $\mathbb{F}$ since we will need to apply them both for $\mathbb{F} = \mathbb{F}_q$ (which is the usual case) and also for $\mathbb{F} = \mathbb{F}_q(t_1, t_2)$ (the field of rational function in t_1, t_2 with coefficients in $\mathbb{F}_q$). Working over this extension field is natural when handling two-dimensional flats and already appears in [KLSS11].

Definition 6.1 (Hasse Derivatives)**.** *Let $\mathbb{F}$ be a field. Given a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ and an $\mathbf{i} \in \mathbb{Z}_{\geq 0}^n$ the $\mathbf{i}$th* Hasse derivative *of f is the polynomial $f^{(\mathbf{i})}$ in the expansion*

$$f(x+z) = \sum_{\mathbf{j} \in \mathbb{Z}_{\geq 0}^n} f^{(\mathbf{j})}(x) z^{\mathbf{j}}$$

where $x = (x_1, ..., x_n)$, $z = (z_1, ..., z_n)$ and $z^{\mathbf{j}} = \prod_{k=1}^n z_k^{j_k}$.

Hasse derivatives satisfy the following useful property (see [DKSS13] for a proof). We will only need this property to show that, if $f^{(\mathbf{i}+\mathbf{j})}$ vanishes at a point then so does $(f^{(\mathbf{i})})^{(\mathbf{j})}$.

Lemma 6.2. *Given a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ and $\mathbf{i}, \mathbf{j} \in \mathbb{Z}^n_{\geq 0}$, we have*

$$(f^{(\mathbf{i})})^{(\mathbf{j})} = f^{(\mathbf{i}+\mathbf{j})} \prod_{k=1}^{n} \binom{i_k + j_k}{i_k}$$

We make precise what it means for a polynomial to vanish on a point $a \in \mathbb{F}^n$ with multiplicity. First we recall for a point $\mathbf{j}$ in the non-negative lattice $\mathbb{Z}^n_{\geq 0}$, its weight is defined as $\mathrm{wt}(\mathbf{j}) = \sum_{i=1}^{n} j_i$.

Definition 6.3 (Multiplicity)**.** *For a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ and a point $a \in \mathbb{F}^n$ we say f vanishes on a with* multiplicity $m \in \mathbb{Z}_{\geq 0}$*, if m is the largest integer such that all Hasse derivatives of f of weight strictly less than m vanish on a. We use $\mathsf{mult}(f, a)$ to refer to the multiplicity of f at a.*

Note that the number of Hasse derivatives over $\mathbb{F}[x_1, \ldots, x_n]$ with weight strictly less than m is $\binom{n+m-1}{n}$. Hence, requiring that a polynomial vanishes to order m at a single point a enforces the same number of homogeneous linear equations on the coefficients of the polynomial. We will use the following simple property concerning multiplicities of composition of polynomials (see [DKSS13] for a proof).

Lemma 6.4. *Given a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ and a tuple $H = (h_1, \ldots, h_n)$ of polynomials in $\mathbb{F}[y_1, \ldots, y_m]$, and $a \in \mathbb{F}^m$ we have,*

$$\mathsf{mult}(f \circ H, a) \geq \mathsf{mult}(f, H(a)).$$

We will now state the multiplicity version of the Schwartz-Zippel bound [Sch79, Zip79] (see [DKSS13] for a proof). We denote by $\mathbb{F}[x_1, .., x_n]_{\leq d}$ the space of polynomials of total degree at most d with coefficients in $\mathbb{F}$.

Lemma 6.5 (Schwartz-Zippel with multiplicities)**.** *Let $\mathbb{F}$ be a field, $d \in \mathbb{Z}_{\geq 0}$ and let $f \in \mathbb{F}[x_1, .., x_n]_{\leq d}$ be a non-zero polynomial. Then, for any finite subset $U \subseteq \mathbb{F}$,*

$$\sum_{a \in U^n} \mathsf{mult}(f, a) \leq d|U|^{n-1}.$$

6.2 The EVAL matrix, its submatrices and their ranks

If M is a matrix over an extension field of $\mathbb{F}_q$, we define the $\mathbb{F}_q$-rank of M, denoted by $\mathsf{rank}_{\mathbb{F}_q} M$, to be the size of the largest subset of columns of M which are $\mathbb{F}_q$-linearly independent (in other words, no non-zero $\mathbb{F}_q$-linear combination of those columns is 0). For convenience, we define the coefficient matrix of a matrix with entries in $\mathbb{F}_q[t_1, t_2]$. This will help us argue about the $\mathbb{F}_q$-rank of a matrix over an extension, by connecting it with the rank of a matrix with entries in $\mathbb{F}_q$.

Definition 6.6 (Coefficient matrix of E)**.** *Let E be an $n_1 \times n_2$ matrix with entries in $\mathbb{F}_q[t_1, t_2]_{\leq d}$. The* coefficient matrix of E*, denoted by $\mathsf{Coeff}(E)$, is a $\binom{d+2}{2} n_1 \times n_2$ matrix with entries in $\mathbb{F}_q$ whose rows are labelled by elements in $((i, j), k) \in \mathbb{Z}^2_{\geq 0} \times [n_1]$ and whose entry in row $((i, j), k)$ and column ℓ is given by the coefficient of $t_1^i t_2^j$ of the polynomial in the (k, ℓ)'th entry of E.*

In other words, to construct $\mathsf{Coeff}(E)$ we replace each entry with a (column) vector of its coefficients. For example:

$$E = \begin{bmatrix} t_1 & t_2+1 \\ 2+4t_1 & t_1+3t_2 \end{bmatrix}, \mathsf{Coeff}(E) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 2 & 0 \\ 4 & 1 \\ 0 & 3 \end{bmatrix}.$$

By construction we have,

$$\mathsf{rank}_{\mathbb{F}_q} E = \mathsf{rank}_{\mathbb{F}_q} \mathsf{Coeff}(E).$$

Our main object of interest is the matrix encoding the evaluation of a subset of monomials (with their derivatives) on a subset of points.

Definition 6.7 ($\mathsf{EVAL}^m(S,W)$ matrix)**.** *Let $\mathbb{F}$ be a field, and let $n, m \in \mathbb{N}$. Given a set $S \subset \mathbb{F}^n$ and a set of monomials $W \subset \mathbb{F}[x_1, \ldots, x_n]$, we let $\mathsf{EVAL}^m(S,W)$ denote an $|S|\binom{m-1+n}{n} \times |W|$ matrix whose columns are indexed by W and rows are indexed by tuples $(x, \mathbf{j}) \in S \times \mathbb{Z}^n_{\geq 0}$ s.t. $wt(\mathbf{j}) < m$. The $((x,\mathbf{j}), f)$th entry of this matrix is,*

$$f^{(\mathbf{j})}(x).$$

In other words, the $(x,\mathbf{j})$th row of the matrix consists of the evaluation of the $\mathbf{j}$'th Hasse derivative of all $f \in W$ at x. Equivalently, the f'th column of the matrix consists of the evaluations of weight strictly less than m Hasse derivatives of f at all points in S.

We let,

$$\mathcal{V} = \{u't_1 + v't_2 | u', v' \in \mathbb{F}_q\}^n = \{ut_1 + vt_2 | u, v \in \mathbb{F}_q^n\} \subseteq (\mathbb{F}_q[t_1, t_2])^n$$

denote the set of n-tuples of homogeneous linear forms in t_1, t_2 and

$$\mathcal{V}_{\text{full}} = \{ut_1 + vt_2 \in \mathcal{V} \mid \dim_{\mathbb{F}_q} \mathsf{span}\{u, v\} = 2\} \subseteq \mathcal{V}$$

denote the subset of $\mathcal{V}$ in which the coefficient vectors of t_1 and of t_2 are linearly independent.

Let $W_{d,n}$ denote the set of monomials in n-variables $x_1, \ldots, x_n$ of degree at most d. Our first lemma shows that the $\mathbb{F}_q$-rank of $\mathsf{EVAL}^m(\mathcal{V}, W_{d,n})$ is maximal whenever d is not too large. This is essentially the Schwartz-Zippel lemma since it means that a polynomial of bounded degree could be recovered from its evaluations (up to high enough order) on a product set.

Lemma 6.8 (Rank of $\mathsf{EVAL}^m(\mathcal{V}, W_{d,n})$)**.** *Let $m \in \mathbb{N}$ then for all $d < mq^2$ we have,*

$$\text{rank}_{\mathbb{F}_q} \mathsf{EVAL}^m(\mathcal{V}, M_{d,n}) = |W_{d,n}| = \binom{d+n}{d}.$$

Proof. Recall $\mathcal{V} = \{u't_1 + v't_2 \in \mathbb{F}_q(t_1, t_2) | u, v \in \mathbb{F}_q\}^n$. Any $\mathbb{F}_q$-linear combination of columns in $\mathsf{EVAL}^m(S, W_{d,n})$ for some subset $S \subseteq \mathcal{V}$ corresponds to looking at the evaluation of the weight $< m$ Hasse derivatives on S of a degree at most d polynomial in $\mathbb{F}_q[x_1, \ldots, x_n]$. To be precise if we take the linear combination of columns corresponding to the monomials $f_1, f_2, \ldots, f_\ell$ with coefficients $\alpha_1, \ldots, \alpha_\ell \in \mathbb{F}_q$ the column vector we get will be the evaluation of the weight $< m$ Hasse derivatives

of $\sum_{i=1}^{\ell} \alpha_i f_i \in \mathbb{F}_q[x_1, \ldots, x_n]$ over S. Note that the polynomial we are considering has coefficients only in $\mathbb{F}_q$ while the evaluations are being done over the field $\mathbb{F}_q(t_1, t_2)$.

For any $d < mq^2$ any $\mathbb{F}_q$-linear combination of columns in $\mathsf{EVAL}^m(\mathcal{V}, W_{d,n})$ being 0 will be equivalent to a degree at most $mq^2 - 1$ polynomial vanishing on $\mathcal{V}$ with multiplicity m. By Lemma 6.5 we see that a non-zero polynomial of degree at most d vanishing on $\mathcal{V}$ (which is a product set of size q^{2n}) with multiplicity at least m satisfies

$$dq^{2(n-1)} \geq mq^{2n}$$

which leads to a contradiction (as $d < mq^2$). This means $\mathsf{EVAL}^m(\mathcal{V}, W_{d,n})$ has $\mathbb{F}_q$-rank $|W_{d,n}|$ for $d < mq^2$. Note this proof would also show that the $\mathbb{F}_q(t_1, t_2)$-rank of $\mathsf{EVAL}^m(\mathcal{V}, W_{d,n})$ is $|W_{d,n}|$ for $d < mq^2$. □

We next show that the same rank bound holds even if we restrict the rows to only come from the smaller set $\mathcal{V}_{\text{full}}$.

Lemma 6.9. $\mathsf{EVAL}^m(\mathcal{V}_{\text{full}}, W_{d,n})$ has $\mathbb{F}_q$-rank $|W_{d,n}|$ for $d < mq^2$.

Proof. This lemma will need the fact that we are only computing the $\mathbb{F}_q$ (and not $\mathbb{F}_q(t_1, t_2)$) rank. Consider any $\mathbb{F}_q$-linear combination f of monomials in $W_{d,n}$. It suffices to show that if f vanishes with multiplicity at least m over $\mathcal{V}_{\text{full}}$ then it vanishes with multiplicity at least m over $\mathcal{V}$. $\mathcal{V} \setminus \mathcal{V}_{\text{full}}$ contains elements of the form ut_1 or $u(ct_1 + t_2)$ where $u \in \mathbb{F}_q^n$ and $c \in \mathbb{F}_q$. First we consider $u \in \mathbb{F}_q^n \setminus \{0\}$. We can pick a $v \in \mathbb{F}_q^n$ such that v and u are linearly independent. $ut_1 + vt_2$ now is an element in $\mathcal{V}_{\text{full}}$. This means f vanishes on $ut_1 + vt_2$ with multiplicity at least m. f is a polynomial in $\mathbb{F}_q[x_1, \ldots, x_n]$ which means all its Hasse derivatives are also $\mathbb{F}_q$-polynomials. Therefore, for any $\mathbf{i} \in \mathbb{Z}_{\geq 0}^n$ we get $f^{(\mathbf{i})}(ut_1)$ by setting $t_2 = 0$ in $f^{(\mathbf{i})}(ut_1 + vt_2)$. This implies that f vanishes on ut_1 with multiplicity at least m. Setting $t_1 = 0$ then shows that f vanishes on 0 with multiplicity at least m. Again as t_1 is a formal variable and $f \in \mathbb{F}_q[x_1, \ldots, x_n]$ we can replace t_1 with $ct_1 + t_2$ to get f vanishes on $u(ct_1 + t_2)$ with multiplicity at least m. □

Our final lemma, which is the heart of this section, shows that any δ-fraction of the rows in $\mathsf{EVAL}^m(\mathcal{V}_{\text{full}}, W_{d,n})$ have rank at least δ times the rank of the full matrix. This is not true for an arbitrary matrix and uses the fact that the general linear group acts on the set of rows in a transitive way.

Lemma 6.10 (Rank of $\mathsf{EVAL}^m(S, W_{d,n})$)**.** *Let $m \in \mathbb{N}$ and $S \subseteq \mathcal{V}_{full}$ with $|S| \geq \delta|\mathcal{V}_{full}|, \delta \in [0, 1]$ then for all $d < mq^2$ we have,*

$$\text{rank}_{\mathbb{F}_q} \mathsf{EVAL}^m(S, W_{d,n}) \geq \delta \cdot |W_{d,n}| = \delta \binom{d+n}{d}.$$

Proof. Consider $S \subseteq \mathcal{V}_{\text{full}}$ such that $|S| = \delta|\mathcal{V}_{\text{full}}|$. For any $M \in \text{GL}_n(\mathbb{F}_q)$ we let M act on $ut_1 + vt_2$ where $u, v \in \mathbb{F}_q^n$ as $M \cdot (ut_1 + vt_2) = Mut_1 + Mvt_2$. Let $M \cdot S = \{M \cdot y | y \in S\}$.

Claim 6.11.

$$\mathsf{rank}_{\mathbb{F}_q} \mathsf{EVAL}^m(S, W_{d,n}) = \mathsf{rank}_{\mathbb{F}_q} \mathsf{EVAL}^m(M \cdot S, W_{d,n}).$$

Proof. We will prove this statement by constructing an isomorphism between the column-space of the two matrices. An element in the column space of $\mathsf{EVAL}^m(S, W_{d,n})$ is the evaluation of the weight strictly less than m Hasse derivatives on S of a polynomial $f(x) \in \mathbb{F}_q[x_1, \ldots, x_n]$ of degree at most d. We map such a vector to the evaluation of the weight strictly less than m Hasse derivatives on $M \cdot S$ of the polynomial $f(M^{-1}x)$ which will also be of degree at most d. The choice of f in the beginning can be ambiguous but if there are two polynomials $f(x)$ and $g(x)$ having the same evaluation of weight strictly less than m Hasse derivatives over S then $f(x) - g(x)$ vanishes on S with multiplicity at least m. By Lemma 6.2, $f(M^{-1}x) - g(M^{-1}x)$ vanishes on $M \cdot S$ with multiplicity m which implies $f(M^{-1}x)$ and $g(M^{-1}x)$ evaluate to the same weight strictly less than m Hasse derivatives over $M \cdot S$. The inverse map can be similarly constructed. □

The above claim shows it suffices to show the rank bound for any $M \cdot S$ where $M \in \mathrm{GL}_n(\mathbb{F}_q)$. We do this by a probabilistic method argument.

The previous Lemma implies that $\mathsf{Coeff}(\mathsf{EVAL}^m(\mathcal{V}_{\text{full}}, W_{d,n}))$ has $\mathbb{F}_q$-rank $|W_{d,n}|$. As this is a matrix with $\mathbb{F}_q$ entries this means that there exists a $|W_{d,n}| = \binom{d+n}{n}$ subset of rows R of $\mathsf{Coeff}(\mathsf{EVAL}^m(\mathcal{V}_{\text{full}}, W_{d,n}))$ which are linearly independent. These rows are indexed by tuples

$$(x, \mathbf{i}, (j,k)) \in \mathcal{V}_{\text{full}} \times \mathbb{Z}^n_{\geq 0} \times \mathbb{Z}^2_{\geq 0}$$

with $\mathrm{wt}(\mathbf{i}) < m$ and $j + k \leq d$. The $(x, \mathbf{i}, (j,k))$th row is the coefficient of $t_1^j t_2^k$ in the evaluation of the $\mathbf{i}$th Hasse Derivative at x of the monomials in $W_{d,n}$.

We pick an $M \in \mathrm{GL}_n(\mathbb{F}_q)$ uniformly at random. We now calculate the expected fraction of the rows from R which appear in $\mathsf{Coeff}(\mathsf{EVAL}^m(M \cdot S, M_{d,n}))$.

A row in R indexed by $(x, i, (j,k)) \in \mathcal{V}_{\text{full}} \times \mathbb{Z}^n_{\geq 0} \times \mathbb{Z}^2_{\geq 0}$ will appear in $\mathsf{Coeff}(\mathsf{EVAL}^m(M \cdot S, W_{d,n}))$ if and only if $x \in M \cdot S$. As the action of $\mathrm{GL}_n(\mathbb{F}_q)$ on $\mathcal{V}_{\text{full}}$ we see that this happens with probability at least δ. This means the expected fraction of rows in R appearing in $\mathsf{Coeff}(\mathsf{EVAL}^m(M \cdot S, W_{d,n}))$ is at least δ. This ensures that there is some matrix M such that $\mathsf{Coeff}(\mathsf{EVAL}^m(M \cdot S, W_{d,n}))$ and hence $\mathsf{EVAL}^m(M \cdot S, W_{d,n})$ has $\mathbb{F}_q$-rank at least $\delta|W_{d,n}|$. □

We note the above lemma could be proven in a more general setting where we wanted to compare the $\mathbb{F}_q$ rank of $\mathsf{EVAL}^m(G, W_{d,n})$ for $G = S \subseteq \mathbb{F}^n$ and $G = S' \subseteq S$ a large subset of S as long as the general linear group acts transitively on S. For instance, this style of argument was also used in [Dha21] to obtain a better dependence on β for $(1, m, \beta)$-Furstenberg sets[5] over $\mathbb{Z}/p^k\mathbb{Z}$.

We will use a simple corollary of this lemma.

Corollary 6.12. *Let $r \in \mathbb{N}$ and $S \subseteq \mathcal{V}_{full}$ with $|S| \geq \delta|\mathcal{V}_{full}|, \delta \in [0,1]$ then for any $d < rq^2$ there exists a set $P_S(d,r), |P_S(d,r)| = \delta\binom{d+n}{n}$ of monomials of degree at most d such that no non-zero $\mathbb{F}_q$-linear combination of monomials in $P_S(d,r)$ vanishes with multiplicity at least r over all points in S.*

6.3 Proving the bound on Furstenberg sets

We first give a brief description of the polynomial method argument as was used for example in [KLSS11]. Given a $(k, \gamma q^2, \beta)$-Furstenberg set K we take a polynomial Q of degree at most d (where d depends on β, γ and q) which vanishes with high multiplicity on K. If $|K|$ small, such a polynomial can be found by solving a system of linear constraints. For at least a β fraction of the

[5] Denoted as (m, β)-Kakeya sets in [Dha21]

flats $A \in \mathcal{L}_2^*(\mathbb{F}_q^n)$ there is a shift $a + A, a \in \mathbb{F}_q^n$ such that $a + A$ is γq^2-rich w.r.t. K. By restricting Q to $a + A$ and using Lemma 6.5 we then show that Q vanishes identically on $a + A$ which implies that the highest degree homogeneous part of Q vanishes identically on A. This will imply that the highest degree homogenous part of Q vanishes on a β fraction of $A \in \mathcal{L}_2^*(\mathbb{F}_q^n)$. Another application of the Lemma 6.5 then gives us a size bound for $|K|$ by arguing that $\deg(Q)$ cannot be too small (here, the dependency between β and d comes into play). The size of K is lower bounded by the number of at most degree d monomials $\binom{d+n}{n}$. The dependence of β on d leads to a loss of β^n in the final bound.

In [ORW22] they overcome this problem by using random rotations to reduce to the case of constant β. We overcome this loss by instead using Corollary 6.12 to start out with a subset of monomials of degree at most d' (here d' will not depend on β) of size $\beta\binom{d'+n}{n}$ such that any $\mathbb{F}_q$-linear combination of those will not vanish on the β fraction of flats in $\mathcal{L}_k^*(\mathbb{F}_q^n)$ which have γq^2-rich shifts w.r.t. K. Now the standard polynomial method argument will let us prove Lemma 4.1.

Lemma 4.1 (Size of $(2, \gamma q^2, \beta)$-Furstenberg sets)**.** *For any $\gamma \in [0, 1], \beta \in [0, 1], n \in \mathbb{N}$, q a prime power every $(2, \gamma q^2, \beta)$-Furstenberg set $K \subseteq \mathbb{F}_q^n$ has size at least,*

$$|K| \geq \beta \gamma^n q^n \left(1 + \frac{1}{q}\right)^{-n}.$$

Proof. Let

$$t = \lceil \gamma q^2 \rceil / q \geq \gamma q.$$

As K is a $(2, \gamma q^2, \beta)$-Furstenberg set then there exists a subset $\mathcal{F} \subseteq \mathcal{L}_2^*(\mathbb{F}_q^n)$ of size at least $\beta|\mathcal{L}_2^*(\mathbb{F}_q^n)|$ such that for every $A \in \mathcal{F}$ there exists a $tq = \lceil \gamma q^2 \rceil$-rich shift $a + A$ for some $a \in \mathbb{F}_q^n$. To $\mathcal{F}$ we can also associate a set of elements

$$\mathcal{F}' = \{ut_1 + vt_2 | u, v \in \mathbb{F}_q^n, \mathsf{span}\{u, v\} \in \mathcal{F}\} \subseteq \mathbb{F}_q(t_1, t_2)^n \subset \mathcal{V}_{\text{full}}.$$

Note, in general for each flat $A \in \mathcal{L}_2^*(\mathbb{F}_q^n)$ there are $(q^2-1)(q^2-q)$ elements in $\mathcal{V}_{\text{full}}$ corresponding to it (because there are $(q^2 - 1)(q^2 - q)$ ordered pairs of vectors which span A) and each element $ut_1 + vt_2 \in \mathcal{V}_{\text{full}}$ corresponds to a unique choise of basis. Thus, we have

$$|\mathcal{F}'| \geq \beta |\mathcal{V}_{\text{full}}|.$$

Let ℓ be an integer parameter (we will later send ℓ to infinity) and take

$$m = (q^2 + t - 1)\ell$$

and

$$d = q^2 t\ell - 1$$

for $\ell \in \mathbb{N}$. As $d < q^2 t\ell$, using Corollary 6.12 we can find a set $P_{\mathcal{F}'}(d, t\ell)$ of monomials of degree at most d so that

$$|P_{\mathcal{F}'}(d, t\ell)| \geq \beta \binom{d+n}{n}$$

and such that no $\mathbb{F}_q$-linear combination of monomials in $P_{\mathcal{F}'}(d, t\ell)$ vanishes over all points in $\mathcal{F}'$ with multiplicity at least $t\ell$. If

$$\beta \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}} > |K|,$$

then we can find (by solving a system of homogeneous linear equations) a non-zero polynomial $Q \in \mathbb{F}_q[x_1, \ldots, x_n]$ of degree at most d spanned by monomials in $P_{\mathcal{F}'}(d, t\ell)$ vanishing with multiplicity at least m on every point in K. Let Q^H be the highest degree homogenous part of Q.

Claim 6.13. *For any $x \in \mathcal{F}'$ we have mult$(Q^H, x) \geq t\ell$.*

Proof. Let $\mathbf{j} \in \mathbb{Z}^n_{\geq 0}$ be such that $\mathrm{wt}(\mathbf{j}) < t\ell$. By Lemma 6.2, $Q^{(\mathbf{j})}$ vanishes on K with multiplicity at least $m - \mathrm{wt}(\mathbf{j})$. $Q^{(\mathbf{j})}$ is also of degree at most $d - \mathrm{wt}(\mathbf{j})$

By construction for every element $ut_1 + vt_2 \in \mathcal{F}'$ there exists an element $c_{u,v} \in \mathbb{F}_q^n$ such that $c_{u,v} + \{ut_1 + vt_2 | t_1, t_2 \in \mathbb{F}_q\}$ is qt-rich w.r.t. K. By Lemma 6.4 we have that the bivariate polynomial $Q^{(\mathbf{j})}(c_{u,v} + ut_1 + vt_2)$ vanishes on qt many points in $\mathbb{F}_q^2$ with multiplicity at least $m - \mathrm{wt}(\mathbf{j})$.

By Lemma 6.5, we have that, if $Q^{(\mathbf{j})}(c_{u,v} + ut_1 + vt_2)$ is non-zero then,

$$(d - \mathrm{wt}(\mathbf{j}))q \geq (m - \mathrm{wt}(\mathbf{j}))qt.$$

Rearranging gives us,

$$d + \mathrm{wt}(\mathbf{j})(t-1) \geq mt.$$

Substituting $d = q^2t\ell - 1$, $m = (q^2 + t - 1)\ell$ and using the fact that $\mathrm{wt}(\mathbf{j}) < t\ell$ gives us,

$$q^2t\ell - 1 + (t-1)t\ell > q^2t\ell + (t-1)t\ell.$$

This leads to a contradiction. This means that $Q^{(\mathbf{j})}(c_{u,v} + ut_1 + vt_2)$ is identically 0. We note, $Q^{(\mathbf{j})}(c_{u,v} + ut_1 + vt_2) \in \mathbb{F}_q[t_1, t_2]$ and its highest degree homogeneous part is $(Q^H)^{(\mathbf{j})}(ut_1 + vt_2)$. This means $(Q^H)^{(\mathbf{j})}(ut_1 + vt_2) = 0$ for all $\mathbf{j}$ such that $\mathrm{wt}(\mathbf{j}) < t\ell$. This proves the claim. □

As Q^H is a non-zero polynomial with coefficients in $\mathbb{F}_q$ spanned by monomials in $P_{\mathcal{F}'}(d, t\ell)$ and it vanishes with multiplicity at least $t\ell$ on every point in $\mathcal{F}'$, we get a contradiction to Corollary 6.12. Therefore, we can conclude that

$$\beta \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}} \leq |K|.$$

Substituting $d = q^2t\ell - 1$, $m = (q^2 + t - 1)\ell$ gives us,

$$|K| \geq \beta \frac{(q^2t\ell - 1 + n)(q^2t\ell - 2 + n)\ldots(q^2t\ell)}{((q^2 + t - 1)\ell + n - 1)((q^2 + t - 1)\ell + n - 2)\ldots((q^2 + t - 1)\ell)}.$$

Letting $\ell \to \infty$ gives us,

$$|K| \geq \beta t^n \left(1 + \frac{t-1}{q^2}\right)^{-n}.$$

As $q \geq t \geq \gamma q$ the proof of the lemma is complete. □

We note the arguments in this section easily generalizes for $(k, \gamma q^k, \beta)$-Furstenberg sets for all $k \geq 1$ to prove the following theorem.

Theorem 6.14 (Size of $(k, \gamma q^k, \beta)$-Furstenberg Sets)**.** *For any $\gamma \in [0,1], \beta \in [0,1], n \in \mathbb{N}$, q a prime power every $(k, \gamma q^k, \beta)$-Furstenberg set $K \subseteq \mathbb{F}_q^n$ has size at least,*

$$|K| \geq \beta \gamma^n q^n \left(1 + \frac{1}{q^{k-1}}\right)^{-n}.$$

The [ORW22] reduction gives a quantitatively worse bound of $\beta\gamma^n q^n (2^n \log_2(2en)e)^{-1}$ compared to $\beta\gamma^n q^n 2^{-n}$ for $k = 1$. Note that for $k \geq 5$, Theorem 3.3 gives us much better bounds for $q >> n$.

References

[ABI86] Noga Alon, László Babai, and Alon Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7(4):567–583, 1986.

[ADM+99] Noga Alon, Martin Dietzfelbinger, Peter Bro Miltersen, Erez Petrank, and Gábor Tardos. Linear Hash Functions. *Journal of Association for Computing Machinery*, 46(5):667–683, 1999.

[BB21] Ting-Wei Chao Boris Bukh. Sharp density bounds on the finite field Kakeya problem. *Discrete Analysis*, 2021.

[BDK+11] Boaz Barak, Yevgeniy Dodis, Hugo Krawczyk, Olivier Pereira, Krzysztof Pietrzak, François-Xavier Standaert, and Yu Yu. Leftover Hash Lemma, Revisited. In *Advances in Cryptology - CRYPTO 2011*, pages 1–20. Springer, 2011.

[CHS06] Ran Canetti, Shai Halevi, and Michael Steiner. Mitigating Dictionary Attacks on Password-Protected Local Storage. In *Advances in Cryptology - CRYPTO 2006*, pages 160–179. Springer, 2006.

[CW79] J.Lawrence Carter and Mark N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18(2):143–154, 1979.

[DDL21a] Manik Dhar, Zeev Dvir, and Ben Lund. Furstenberg sets in finite fields: Explaining and improving the Ellenberg-Erman proof. *preprint arXiv:1909.02431*, 2021.

[DDL21b] Manik Dhar, Zeev Dvir, and Ben Lund. Simple proofs for Furstenberg sets over finite fields. *Discrete Analysis*, 22, 2021.

[Dha] Manik Dhar. Maximal Kakeya and (m, ϵ)-Kakeya bounds over $\mathbb{Z}/N\mathbb{Z}$ for general N. in preparation.

[Dha21] Manik Dhar. The Kakeya set conjecture for $\mathbb{Z}/N\mathbb{Z}$ for general N. *preprint arXiv:2110.14889*, 2021.

[DKSS13] Zeev Dvir, Swastik Kopparty, Shubhangi Saraf, and Madhu Sudan. Extensions to the Method of Multiplicities, with Applications to Kakeya Sets and Mergers. *SIAM Journal on Computing*, 42(6):2305–2328, 2013.

[Dvi09] Z. Dvir. On the size of Kakeya sets in finite fields. *Journal of American Mathematical Society*, 22:1093–1097, 2009.

[DW08] Z. Dvir and A. Wigderson. Kakeya Sets, New Mergers and Old Extractors. In *FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 625–633. IEEE Computer Society, 2008.

[EE16] Jordan Ellenberg and Daniel Erman. Furstenberg sets and Furstenberg schemes over finite fields. *Algebra & Number Theory*, 10(7):1415–1436, 2016.

[EOT10] Jordan S Ellenberg, Richard Oberlin, and Terence Tao. The Kakeya set and maximal conjectures for algebraic varieties over finite fields. *Mathematika*, 56(1):1–25, 2010.

[Gol08] Oded Goldreich. *Computational Complexity: A Conceptual Perspective*. Cambridge University Press, 2008.

[GR19] Ben Green and Imre Z Ruzsa. On the arithmetic kakeya conjecture of katz and tao. *Periodica Mathematica Hungarica*, 78(2):135–151, 2019.

[ILL89] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-Random Generation from One-Way Functions. In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, STOC '89, page 12–24. Association for Computing Machinery, 1989.

[KLSS11] Swastik Kopparty, Vsevolod F Lev, Shubhangi Saraf, and Madhu Sudan. Kakeya-type sets in finite vector spaces. *Journal of Algebraic Combinatorics*, 34(3):337–355, 2011.

[KRV21] Inbar Kaslasi, Ron D. Rothblum, and Prashant Nalini Vasudevanr. Public-Coin Statistical Zero-Knowledge Batch Verification Against Malicious Verifiers. In *Advances in Cryptology: EUROCRYPT 2021, Part III*, page 219–246. Springer, 2021.

[MV84] Kurt Mehlhorn and Uzi Vishkin. Randomized and deterministic simulations of PRAMs by parallel machines with restricted granularity of parallel memories. *Acta Informatica*, 21(4):339–374, 1984.

[ORW22] Or Ordentlich, Oded Regev, and Barak Weiss. New bounds on the density of lattice coverings. *Journal of the American Mathematical Society*, 35(1):295–308, 2022.

[RS98] Martin Raab and Angelika Steger. "Balls into Bins" - A Simple and Tight Analysis. In *Randomization and Approximation Techniques in Computer Science*, pages 159–170. Springer, 1998.

[RTS00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for Dispersers, Extractors, and Depth-Two Superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[Sch79] Jacob T Schwartz. Probabilistic algorithms for verification of polynomial identities. In *International Symposium on Symbolic and Algebraic Manipulation*, pages 200–215. Springer, 1979.

[Wol99] T. Wolff. Recent work connected with the Kakeya problem. *Prospects in mathematics (Princeton, NJ, 1996)*, pages 129–162, 1999.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In Edward W. Ng, editor, *Symbolic and Algebraic Computation*, pages 216–226. Springer, 1979.