Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 4 (2009)

Extensions to the Method of Multiplicities, with applications to Kakeya Sets and Mergers

Zeev Dvir* Swastik Kopparty † Shubhangi Saraf ‡ Madhu Sudan§

January 16, 2009

Abstract

We extend the "method of multiplicities" to get the following results, of interest in combinatorics and randomness extraction.

1. We show that every Kakeya set in $\mathbb{F}_q^n$, the n-dimensional vector space over the finite field on q elements, must be of size at least $q^n/2^n$. This bound is tight to within a $2 + o(1)$ factor for every n as $q \to \infty$.
2. We give improved "randomness mergers", i.e., seeded functions that take as input k (possibly correlated) random variables in $\{0,1\}^N$ and a short random seed and output a single random variable in $\{0,1\}^N$ that is statistically close to having entropy $(1-\delta)\cdot N$ when one of the k input variables is distributed uniformly. The seed we require is only $(1/\delta)\cdot\log k$-bits long, which significantly improves upon previous construction of mergers.

The "method of multiplicities", as used in prior work, analyzed subsets of vector spaces over finite fields by constructing somewhat low degree interpolating polynomials that vanish on every point in the subset *with high multiplicity*. The typical use of this method involved showing that the interpolating polynomial also vanished on some points outside the subset, and then used simple bounds on the number of zeroes to complete the analysis. Our augmentation to this technique is that we prove, under appropriate conditions, that the interpolating polynomial vanishes *with high multiplicity* outside the set. This novelty leads to significantly tighter analyses.

To get the extended method of multiplicities we provide a number of basic technical results about multiplicity of zeroes of polynomials that may be of general use. For instance, we strengthen the Schwartz-Zippel lemma to show that the expected multiplicity of zeroes of a non-zero degree d polynomial at a random point in S^n, for any finite subset S of the underlying field, is at most $d/|S|$ (a fact that does not seem to have been noticed in the CS literature before).

*IAS. zeev.dvir@gmail.com. Research supported in part by NSF Award CCF 0832797.
†MIT CSAIL. swastik@mit.edu. Research supported in part by NSF Award CCF 0829672.
‡MIT CSAIL. shibs@mit.edu. Research supported in part by NSF Award CCF 0829672.
§MIT CSAIL. madhu@mit.edu. Research supported in part by NSF Award CCF 0829672.

ISSN 1433-8092

1 Introduction

The goal of this paper is to improve on an algebraic method that has lately been applied, quite effectively, to analyze combinatorial parameters of subsets of vector spaces that satisfy some given algebraic/geometric conditions. This technique, which we refer to as as the *polynomial method* (of combinatorics), proceeds in three steps: Given the subset K satisfying the algebraic conditions, one first constructs a non-zero low-degree polynomial that vanishes on K. Next, one uses the algebraic conditions on K to show that the polynomial vanishes at other points outside K as well. Finally, one uses the fact that the polynomial is zero too often to derive bounds on the combinatorial parameters of interest. The polynomial method has seen utility in the computer science literature in works on "list-decoding" starting with Sudan [Sud97] and subsequent works. Recently the method has been applied to analyze "extractors" by Guruswami, Umans, and Vadhan [GUV07]. Most relevant to this current paper are its applications to lower bound the cardinality of "Kakeya sets" by Dvir [Dvi08], and the subsequent constructions of "mergers" by Dvir and Wigderson [DW08]. (We will elaborate on some of these results shortly.)

The *method of multiplicities*, as we term it, may be considered an extension of this method. In this extension one constructs polynomials that vanish with *high multiplicity* on the subset K. This requirement often forces one to use polynomials of higher degree than in the polynomial method, but it gains in the second step by using the high multiplicity of zeroes to conclude "more easily" that the polynomial is zero at other points. This typically leads to a tighter analysis of the combinatorial parameters of interest. This method has been applied widely in list-decoding starting with the work of Guruswami and Sudan [GS99] and continuing through many subsequent works, most significantly in the works of Parvaresh and Vardy [PV05] and Guruswami and Rudra [GR06] leading to rate-optimal list-decodable codes. Very recently this method was also applied to improve the lower bounds on the size of "Kakeya sets" by Saraf and Sudan [SS08].

The main contribution of this paper is an extension to this method, that we call the *extended method of multiplicities*, which develops this method (hopefully) fully to derive even tighter bounds on the combinatorial parameters. In our extension, we start as in the method of multiplicities to construct a polynomial that vanishes with high multiplicity on every point of K. But then we extend the second step where we exploit the algebraic conditions to show that the polynomial vanishes with *high multiplicity* on some points outside K as well. Finally we extend the third step to show that this gives better bounds on the combinatorial parameters of interest.

By these extensions we derive nearly optimal lower bounds on the size of Kakeya sets and qualitatively improved analysis of mergers. We also rederive algebraically a known bound on the list-size in the list-decoding of Reed-Solomon codes. We describe these contributions in detail next, before going on to describe some of the technical observations used to derive the extended method of multiplicities (which we believe are of independent interest).

1.1 Kakeya Sets over Finite Fields

Let $\mathbb{F}_q$ denote the finite field of cardinality q. A set $K \subseteq \mathbb{F}_q^n$ is said to be a *Kakeya set* if it "contains a line in every direction". In other words, for every "direction" $\mathbf{b} \in \mathbb{F}_q^n$ there should exist an "offset" $\mathbf{a} \in \mathbb{F}_q^n$ such that the "line" through $\mathbf{a}$ in direction $\mathbf{b}$, i.e., the set $\{\mathbf{a} + t\mathbf{b} | t \in \mathbb{F}_q\}$, is contained in

K. A question of interest in combinatorics/algebra/geometry, posed originally by Wolff [Wol99], is: "What is the size of the smallest Kakeya set, for a given choice of q and n?"

The trivial upper bound on the size of a Kakeya set is q^n and this can be improved to roughly $\frac{1}{2^{n-1}}q^n$ (precisely the bound is $\frac{1}{2^{n-1}}q^n + O(q^{n-1})$, see [SS08] for a proof of this bound due to Dvir). An almost trivial lower bound is $q^{n/2}$ (every Kakeya set "contains" at least q^n lines, but there are at most $|K|^2$ lines that intersect K at least twice). Till recently even the exponent of q was not known precisely (see [Dvi08] for details of work prior to 2008). Dvir [Dvi08] introduced the use of the polynomial method to analyze this problem, and showed that for every n, $|K| \geq c_n q^n$, for some constant c_n depending only on n. (The original result of Dvir gave a weaker bound of $|K| \geq c_n q^{n-1}$. The final bound comes from improvements due to Alon and Tao [Dvi08, Theorem 3].)

Subsequently the work [SS08] explored the growth of the constant c_n as a function of n. The result of [Dvi08] shows that $c_n \geq 1/n!$, and [SS08] improve this bound to show that $c_n \geq 1/(2.6)^n$. This still leaves a gap between the upper bound and the lower bound and we effectively close this gap.

Theorem 1 *If K is a Kakeya set in $\mathbb{F}_q^n$ then $|K| \geq \frac{1}{2^n}q^n$.*

Note that our bound is tight to within a $2 + o(1)$ multiplicative factor as long as $q = \omega(2^n)$ and in particular when $n = O(1)$ and $q \to \infty$.

1.2 Randomness Mergers

A general quest in the computational study of randomness is the search for simple primitives that manipulate random variables to convert their randomness into more useful forms. The exact notion of utility varies with applications. The most common notion is that of "extractors" that produce an output variable that is distributed statistically close to uniformly on the range. Other notions of interest include "condensers", "dispersers" etc. One such object of study (partly because it is useful to construct extractors) is a "randomness merger". A randomness merger takes as input k, possibly correlated, random variables $\mathsf{A}_1, \ldots, \mathsf{A}_k$, along with a short uniformly random seed B, which is independent of $\mathsf{A}_1, \ldots, \mathsf{A}_k$, and "merges" the randomness of $\mathsf{A}_1, \ldots, \mathsf{A}_k$. Specifically the output of the merger should be statistically close to a high-entropy-rate source of randomness provided at least one of the input variables $\mathsf{A}_1, \ldots, \mathsf{A}_k$ is uniform.

Mergers were first introduced by Ta-Shma [TS96] in the context of explicit constructions of extractors. A general framework was given in [TS96] that reduces the problem of constructing good extractors into that of constructing good mergers. Subsequently, in [LRVW03], mergers were used in a more complicated manner to create extractors which were optimal to within constant factors. The mergers of [LRVW03] had a very simple algebraic structure: the output of the merger was a random linear combination of the blocks over a finite vector space. The [LRVW03] merger analysis was improved in [DS07] using the connection to the finite field Kakeya problem and the (then) state of the art results on Kakeya sets.

The new technique in [Dvi08] inspired Dvir and Wigderson [DW08] to give a very simple, algebraic, construction of a merger which can be viewed as a derandomized version of the [LRVW03] merger. They associate the domain of each random variable A_i with a vector space $\mathbb{F}_q^n$. With the k-tuple of random variables $\mathsf{A}_1, \ldots, \mathsf{A}_k$, they associate a curve $C : \mathbb{F}_q \to \mathbb{F}_q^n$ of degree $\leq k$ which 'passes' through all the points $\mathsf{A}_1, \ldots, \mathsf{A}_k$ (that is, the image of C contains these points). They then select a random point $u \in \mathbb{F}_q$ and output $C(u)$ as the "merged" output. They show that if $q \geq \mathrm{poly}(k \cdot n)$ then the output of the merger is statistically close to a distribution of entropy-rate arbitrarily close to 1 on $\mathbb{F}_q^n$.

While the polynomial (or at least linear) dependence of q on k is essential to the construction above, the requirement $q \geq \mathrm{poly}(n)$ appears only in the analysis. In our work we remove this restriction to show:

Informal Theorem: *For every k, q the output of the Dvir-Wigderson merger is close to a source of entropy rate $1 - \log_q k$.*

The above theorem (in its more formal form given in Theorem 17) allows us to merge k sources using seed length which is only logarithmic in the number of sources and does not depend entirely on the length of each source. Earlier constructions of mergers required the seed to depend either linearly on the number of blocks [LRVW03, Zuc07] or to depend also on the *length* of each block [DW08]. [1]

1.3 List-Decoding of Reed-Solomon Codes

The Reed-Solomon list-decoding problem is the following: Given a sequence of points

$$(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n) \in \mathbb{F}_q \times \mathbb{F}_q,$$

and parameters k and t, find the list of all polynomials $p_1, \ldots, p_L$ of degree at most k that agree with the given set of points on t locations, i.e., for every $j \in \{1, \ldots, L\}$ the set $\{i | p_j(\alpha_i) = \beta_i\}$ has at least t elements. The associated combinatorial problem is: How large can the list size, L, be for a given choice of k, t, n, q (when maximized over all possible set of distinct input points)?

A somewhat nonstandard, yet reasonable, interpretation of the list-decoding algorithms of [Sud97, GS99] is that they give algebraic proofs, by the polynomial method and the method of multiplicities, of known combinatorial upper bounds on the list size, when $t > \sqrt{kn}$. Their proofs *happen* also to be algorithmic and so lead to algorithms to find a list of all such polynomials.

However, the bound given on the list size in the above works does not match the best known combinatorial bound. The best known bound to date seems to be that of Cassuto and Bruck [CB04] who show that, letting $R = k/n$ and $\gamma = t/n$, if $\gamma^2 > R$, then the list size L is bounded by $O(\frac{\gamma}{\gamma^2 - R})$ (in contrast, the Johnson bound and the analysis of [GS99] gives a list size bound of $O(\frac{1}{\gamma^2 - R})$, which is asymptotically worse for, say, $\gamma = (1 + O(1))\sqrt{R}$ and R tending to 0). In Theorem 20 we recover the bound of [CB04] using our extended method of multiplicities.

1.4 Technique: Extended method of multiplicities

The common insight to all the above improvements is that the extended method of multiplicities can be applied to each problem to improve the parameters. Here we attempt to describe the technical novelties in the development of the extended method of multiplicities.

[1] The result we refer to in [Zuc07, Theorem 5.1] is actually a condenser (which is stronger than a merger).

For concreteness, let us take the case of the Kakeya set problem. Given a set $K \subseteq \mathbb{F}_q^n$, the method first finds a non-zero polynomial $P \in \mathbb{F}_q[X_1, \ldots, X_n]$ that vanishes with high multiplicity m on each point of K. The next step is to prove that P vanishes with fairly high multiplicity ℓ at every point in $\mathbb{F}_q^n$ as well. This step turns out to be somewhat subtle (and is evidenced by the fact that the exact relationship between m and ℓ is not simple). Our analysis here crucially uses the fact that the (Hasse) derivatives of the polynomial P, which are the central to the notion of multiplicity of roots, are themselves polynomials, and also vanish with high multiplicity at points in K. This fact does not seem to have been needed/used in prior works and is central to ours.

A second important technical novelty arises in the final step of the method of multiplicities, where we need to conclude that if the degree of P is "small", then P must be identically zero. Unfortunately in our application the degree of P may be much larger than q (or nq, or even q^n). To prove that it is identically zero we need to use the fact that P vanishes *with high multiplicity* at every point in $\mathbb{F}_q^n$, and this requires some multiplicity-enhanced version of the standard Schwartz-Zippel lemma. We prove such a strengthening, showing that the expected multiplicity of zeroes of a degree d polynomial (even when $d \gg q$) at a random point in $\mathbb{F}_q^n$ is at most d/q (see Lemma 8). Using this lemma, we are able to derive much better benefits from the "polynomial method". Indeed we feel that this allows us to fully utilize the power of the polynomial ring $\mathbb{F}_q[\mathbf{X}]$ and are not limited by the power of the function space mapping $\mathbb{F}_q^n$ to $\mathbb{F}_q$.

Putting these ingredients together, the analysis of the Kakeya sets follows easily. The analysis of the mergers follows a similar path and may be viewed as a "statistical" extension of the Kakeya set analysis to "curve" based sets, i.e., here we consider sets S that have the property that for a noticeable fraction points $\mathbf{x} \in \mathbb{F}_q^n$ there exists a low-degree curve passing through $\mathbf{x}$ that has a noticeable fraction of its points in S. We prove such sets must also be large and this leads to the analysis of the Dvir-Wigderson merger.

Organization of this paper. In Section 2 we define the notion of the multiplicity of the roots of a polynomial, using the notion of the Hasse derivative. We present some basic facts about multiplicities and Hasse derivatives, and also present the multiplicity based version of the Schwartz-Zippel lemma. In Section 3 we present our lower bounds for Kakeya sets. In Section 4 we extend this analysis for "curves" and for "statistical" versions of the Kakeya property. This leads to our analysis of the Dvir-Wigderson merger in Section 5. Finally, in Section 6, we include the algebraic proof of the list-size bounds for the list-decoding of Reed-Solomon codes.

2 Preliminaries

In this section we formally define the notion of "mutliplicity of zeroes" along with the companion notion of the "Hasse derivative". We also describe basic properties of these notions, concluding with the proof of the "multiplicity-enhanced version" of the Schwartz-Zippel lemma.

2.1 Basic definitions

We start with some notation. We use $[n]$ to denote the set $\{1, \ldots, n\}$. For a vector $\mathbf{i} = \langle i_1, \ldots, i_n \rangle$ of non-negative integers, its *weight*, denoted $\mathrm{wt}(\mathbf{i})$, equals $\sum_{j=1}^n i_j$.

Let $\mathbb{F}$ be any field, and $\mathbb{F}_q$ denote the finite field of q elements. For $\mathbf{X} = \langle X_1, \ldots, X_n \rangle$, let $\mathbb{F}[\mathbf{X}]$ be the ring of polynomials in $X_1, \ldots, X_n$ with coefficients in $\mathbb{F}$. For a polynomial $P(\mathbf{X})$, we let $H_P(\mathbf{X})$ denote the homogeneous part of $P(\mathbf{X})$ of highest total degree.

For a vector of non-negative integers $\mathbf{i} = \langle i_1, \ldots, i_n \rangle$, let $\mathbf{X}^{\mathbf{i}}$ denote the monomial $\prod_{j=1}^{n} X_j^{i_j} \in \mathbb{F}[\mathbf{X}]$. Note that the (total) degree of this monomial equals $\mathrm{wt}(\mathbf{i})$. For n-tuples of non-negative integers $\mathbf{i}$ and $\mathbf{j}$, we use the notation

$$\binom{\mathbf{i}}{\mathbf{j}} = \prod_{k=1}^{n} \binom{i_k}{j_k}.$$

Note that the coefficient of $\mathbf{Z}^{\mathbf{i}}\mathbf{W}^{\mathbf{r}-\mathbf{i}}$ in the expansion of $(\mathbf{Z}+\mathbf{W})^{\mathbf{r}}$ equals $\binom{\mathbf{r}}{\mathbf{i}}$.

Definition 2 ((Hasse) Derivative) *For $P(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$ and non-negative vector $\mathbf{i}$, the $\mathbf{i}$th* (Hasse) derivative *of P, denoted $P^{(\mathbf{i})}(\mathbf{X})$, is the coefficient of $\mathbf{Z}^{\mathbf{i}}$ in the polynomial $\tilde{P}(\mathbf{X}, \mathbf{Z}) \stackrel{\mathrm{def}}{=} P(\mathbf{X}+\mathbf{Z}) \in \mathbb{F}[\mathbf{X}, \mathbf{Z}]$.*

Thus,

$$P(\mathbf{X}+\mathbf{Z}) = \sum_{\mathbf{i}} P^{(\mathbf{i})}(\mathbf{X})\mathbf{Z}^{\mathbf{i}}. \tag{1}$$

We are now ready to define the notion of the (zero-)multiplicity of a polynomial at any given point.

Definition 3 (Multiplicity) *For $P(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$ and $\mathbf{a} \in \mathbb{F}^n$, the* multiplicity *of P at $\mathbf{a} \in \mathbb{F}^n$, denoted $\mathrm{mult}(P, \mathbf{a})$, is the largest integer M such that for every non-negative vector $\mathbf{i}$ with $\mathrm{wt}(\mathbf{i}) < M$, we have $P^{(\mathbf{i})}(\mathbf{a}) = 0$ (if M may be taken arbitrarily large, we set $\mathrm{mult}(P, \mathbf{a}) = \infty$).*

Note that $\mathrm{mult}(P, \mathbf{a}) \geq 0$ for every $\mathbf{a}$. Also, $P(\mathbf{a}) = 0$ if and only if $\mathrm{mult}(P, \mathbf{a}) \geq 1$.

The above notations and definitions also extend naturally to a tuple $P(\mathbf{X}) = \langle P_1(\mathbf{X}), \ldots, P_m(\mathbf{X}) \rangle$ of polynomials with $P^{(\mathbf{i})} \in \mathbb{F}[\mathbf{X}]^m$ denoting the vector $\langle (P_1)^{(\mathbf{i})}, \ldots, (P_m)^{(\mathbf{i})} \rangle$. In particular, we define $\mathrm{mult}(P, \mathbf{a}) = \min_{j \in [m]} \{\mathrm{mult}(P_j, \mathbf{a})\}$.

The definition of multiplicity above is similar to the standard (analytic) definition of multiplicity with the difference that the standard partial derivative has been replaced by the Hasse derivative. The Hasse derivative is also a reasonably well-studied quantity (see, for example, [HKT08, pages 144-155]) and seems to have first appeared in the CS literature (without being explicitly referred to by this name) in the work of Guruswami and Sudan [GS99]. It typically behaves like the standard derivative, but with some key differences that make it more useful/informative over finite fields. For completeness we review basic properties of the Hasse derivative and multiplicity in the following subsections.

2.2 Properties of Hasse Derivatives

The following proposition lists basic properties of the Hasse derivatives. Parts (1)-(3) below are the same as for the analytic derivative, while Part (4) is not! Part (4) considers the derivatives of the derivatives of a polynomial and shows a different relationship than is standard for the analytic derivative. However crucial for our purposes is that it shows that the $\mathbf{j}$th derivative of the $\mathbf{i}$th derivative is zero if (though not necessarily only if) the $(\mathbf{i}+\mathbf{j})$-th derivative is zero.

Proposition 4 (Basic Properties of Derivatives) *Let* $P(\mathbf{X}), Q(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]^m$ *and let* $\mathbf{i}$, $\mathbf{j}$ *be vectors of nonnegative integers. Then:*

1. $P^{(\mathbf{i})}(\mathbf{X}) + Q^{(\mathbf{i})}(\mathbf{X}) = (P+Q)^{(\mathbf{i})}(\mathbf{X})$.
2. *If* P *is homogeneous of degree* d*, then* $P^{(i)}$ *is homogeneous of degree* $d - \mathrm{wt}(i)$.
3. $(H_P)^{(\mathbf{i})}(\mathbf{X}) = H_{P^{(\mathbf{i})}}(\mathbf{X})$
4. $\left(P^{(\mathbf{i})}\right)^{(\mathbf{j})}(\mathbf{X}) = \binom{\mathbf{i}+\mathbf{j}}{\mathbf{i}} P^{(\mathbf{i}+\mathbf{j})}(\mathbf{X})$.

Proof

Items 1 and 2 are easy to check, and item 3 follows immediately from them. For item 4, we expand $P(\mathbf{X}+\mathbf{Z}+\mathbf{W})$ in two ways. First expand

$$\begin{aligned} P(\mathbf{X}+(\mathbf{Z}+\mathbf{W})) &= \sum_{\mathbf{k}} P^{(\mathbf{k})}(\mathbf{X})(\mathbf{Z}+\mathbf{W})^{\mathbf{k}} \\ &= \sum_{\mathbf{k}} \sum_{\mathbf{i}+\mathbf{j}=\mathbf{k}} P^{(\mathbf{k})}(\mathbf{X}) \binom{\mathbf{k}}{\mathbf{i}} \mathbf{Z}^{\mathbf{j}}\mathbf{W}^{\mathbf{i}} \\ &= \sum_{\mathbf{i},\mathbf{j}} P^{(\mathbf{i}+\mathbf{j})}(\mathbf{X}) \binom{\mathbf{i}+\mathbf{j}}{\mathbf{i}} \mathbf{Z}^{\mathbf{j}}\mathbf{W}^{\mathbf{i}}. \end{aligned}$$

On the other hand, we may write

$$P((\mathbf{X}+\mathbf{Z})+\mathbf{W}) = \sum_{\mathbf{i}} P^{(\mathbf{i})}(\mathbf{X}+\mathbf{Z})\mathbf{W}^{\mathbf{i}} = \sum_{\mathbf{i}} \sum_{\mathbf{j}} \left(P^{(\mathbf{i})}\right)^{(\mathbf{j})}(\mathbf{X})\mathbf{Z}^{\mathbf{j}}\mathbf{W}^{\mathbf{i}}.$$

Comparing coefficients of $\mathbf{Z}^{\mathbf{j}}\mathbf{W}^{\mathbf{i}}$ on both sides, we get the result. ■

2.3 Properties of Multiplicities

We now translate some of the properties of the Hasse derivative into properties of the multiplicities.

Lemma 5 (Basic Properties of multiplicities) *If* $P(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$ *and* $\mathbf{a} \in \mathbb{F}^n$ *are such that* $\mathrm{mult}(P, \mathbf{a}) = m$*, then* $\mathrm{mult}(P^{(\mathbf{i})}, \mathbf{a}) \geq m - \mathrm{wt}(\mathbf{i})$.

Proof By assumption, for any $\mathbf{k}$ with $\mathrm{wt}(\mathbf{k}) < m$, we have $P^{(\mathbf{k})}(\mathbf{a}) = 0$. Now take any $\mathbf{j}$ such that $\mathrm{wt}(\mathbf{j}) < m - \mathrm{wt}(\mathbf{i})$. By item 3 of Proposition 4, $(P^{(\mathbf{i})})^{(\mathbf{j})}(\mathbf{a}) = \binom{\mathbf{i}+\mathbf{j}}{\mathbf{i}} P^{(\mathbf{i}+\mathbf{j})}(\mathbf{a})$. Since $\mathrm{wt}(\mathbf{i}+\mathbf{j}) = \mathrm{wt}(\mathbf{i}) + \mathrm{wt}(\mathbf{j}) < m$, we deduce that $(P^{(\mathbf{i})})^{(\mathbf{j})}(\mathbf{a}) = 0$. Thus $\mathrm{mult}(P^{(\mathbf{i})}, \mathbf{a}) \geq m - \mathrm{wt}(\mathbf{i})$. ■

We now discuss the behavior of multiplicities under composition of polynomial tuples. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_\ell)$ be formal variables. Let $P(\mathbf{X}) = (P_1(\mathbf{X}), \ldots, P_m(\mathbf{X})) \in \mathbb{F}[\mathbf{X}]^m$ and $Q(\mathbf{Y}) = (Q_1(Y), \ldots, Q_n(Y)) \in \mathbb{F}[\mathbf{Y}]^n$. We define the composition polynomial $P \circ Q(\mathbf{Y}) \in \mathbb{F}[\mathbf{Y}]^m$ to be the polynomial $P(Q_1(\mathbf{Y}), \ldots, Q_n(\mathbf{Y}))$. In this situation we have the following proposition.

Proposition 6 *Let $P(\mathbf{X}), Q(\mathbf{Y})$ be as above. Then for any $\mathbf{a} \in \mathbb{F}^\ell$,*

$$\mathrm{mult}(P \circ Q, \mathbf{a}) \geq \mathrm{mult}(P, Q(\mathbf{a})) \cdot \mathrm{mult}(Q - Q(\mathbf{a}), \mathbf{a}).$$

In particular, since $\mathrm{mult}(Q - Q(\mathbf{a}), \mathbf{a}) \geq 1$, we have $\mathrm{mult}(P \circ Q, \mathbf{a}) \geq \mathrm{mult}(P, Q(\mathbf{a}))$.

Proof Let $m_1 = \mathrm{mult}(P, Q(\mathbf{a}))$ and $m_2 = \mathrm{mult}(Q - Q(\mathbf{a}), \mathbf{a})$. Clearly $m_2 > 0$. If $m_1 = 0$ the result is obvious. Now assume $m_1 > 0$ (so that $P(Q(\mathbf{a})) = 0$).

$$\begin{aligned}
P(Q(\mathbf{a}+\mathbf{Z})) &= P\left(Q(\mathbf{a}) + \sum_{\mathbf{i}\neq 0} Q^{(\mathbf{i})}(\mathbf{a})\mathbf{Z}^{\mathbf{i}}\right) \\
&= P\left(Q(\mathbf{a}) + \sum_{\mathrm{wt}(\mathbf{i})\geq m_2} Q^{(\mathbf{i})}(\mathbf{a})\mathbf{Z}^{\mathbf{i}}\right) && \text{since } \mathrm{mult}(Q - Q(\mathbf{a}), \mathbf{a}) = m_2 > 0 \\
&= P\left(Q(\mathbf{a}) + h(\mathbf{Z})\right) && \text{where } h(\mathbf{Z}) = \textstyle\sum_{\mathrm{wt}(\mathbf{i})\geq m_2} Q^{(\mathbf{i})}(\mathbf{a})\mathbf{Z}^{\mathbf{i}} \\
&= P(Q(\mathbf{a})) + \sum_{\mathbf{j}\neq 0} P^{(\mathbf{j})}(Q(\mathbf{a}))h(\mathbf{Z})^{\mathbf{j}} \\
&= \sum_{\mathrm{wt}(\mathbf{j})\geq m_1} P^{(\mathbf{j})}(Q(\mathbf{a}))h(\mathbf{Z})^{\mathbf{j}} && \text{since } \mathrm{mult}(P, Q(\mathbf{a})) = m_1 > 0
\end{aligned}$$

Thus, since each monomial $\mathbf{Z}^{\mathbf{i}}$ appearing in h has $\mathrm{wt}(\mathbf{i}) \geq m_2$, and each occurrence of $h(\mathbf{Z})$ in $P(Q(\mathbf{a}+\mathbf{Z}))$ is raised to the power $\mathbf{j}$, with $\mathrm{wt}(\mathbf{j}) \geq m_1$, we conclude that $P(Q(\mathbf{a}+\mathbf{Z}))$ is of the form $\sum_{\mathrm{wt}(\mathbf{k})\geq m_1 \cdot m_2} c_{\mathbf{k}}\mathbf{Z}^{\mathbf{k}}$. This shows that $(P \circ Q)^{(\mathbf{k})}(\mathbf{a}) = 0$ for each $\mathbf{k}$ with $\mathrm{wt}(\mathbf{k}) < m_1 \cdot m_2$, and the result follows. ■

Corollary 7 *Let $P(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$ where $\mathbf{X} = (X_1, \ldots, X_n)$. Let $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n$. Let $P_{\mathbf{a},\mathbf{b}}(T)$ be the polynomial $P(\mathbf{a} + T \cdot \mathbf{b}) \in \mathbb{F}[T]$. Then for any $t \in \mathbb{F}$,*

$$\mathrm{mult}(P_{\mathbf{a},\mathbf{b}}, t) \geq \mathrm{mult}(P, \mathbf{a} + t \cdot \mathbf{b}).$$

Proof Let $Q(T) = \mathbf{a} + T\mathbf{b} \in \mathbb{F}[T]^n$. Applying the previous proposition to $P(\mathbf{X})$ and $Q(T)$, we get the desired claim. ■

2.4 Strengthening of the Schwartz-Zippel Lemma

We are now ready to state and prove the strengthening of the Schwartz-Zippel lemma. In the standard form this lemma states that the probability that $P(\mathbf{a}) = 0$ when $\mathbf{a}$ is drawn uniformly at random from S^n is at most $d/|S|$, where P is a non-zero degree d polynomial and $S \subseteq \mathbb{F}$ is a finite set. Using $\min\{1, \mathrm{mult}(P, \mathbf{a})\}$ as the indicator variable that is 1 if $P(\mathbf{a}) = 0$, this lemma can be restated as saying $\sum_{\mathbf{a}\in S^n} \min\{1, \mathrm{mult}(P, \mathbf{a})\} \leq d \cdot |S|^{n-1}$. Our version below strengthens this lemma by replacing $\min\{1, \mathrm{mult}(P, \mathbf{a})\}$ with $\mathrm{mult}(P, \mathbf{a})$ in this inequality.

Lemma 8 *Let $P \in \mathbb{F}[\mathbf{X}]$ be a nonzero polynomial of total degree at most d. Then for any finite $S \subseteq \mathbb{F}$,*

$$\sum_{\mathbf{a} \in S^n} \mathrm{mult}(P, \mathbf{a}) \leq d \cdot |S|^{n-1}.$$

Proof We prove it by induction on n.

For the base case when $n = 1$, we first show that if $\mathrm{mult}(P, a) = m$ then $(X - a)^m$ divides $P(X)$. To see this, note that by definition of multiplicity, we have that $P(a + Z) = \sum_i P^{(i)}(a) Z^i$ and $P^{(i)}(a) = 0$ for all $i < m$. We conclude that Z^m divides $P(a+Z)$, and thus $(X-a)^m$ divides $P(X)$. It follows that $\sum_{a \in S} \mathrm{mult}(P, a)$ is at most the degree of P.

Now suppose $n > 1$. Let

$$P(X_1, \ldots, X_n) = \sum_{j=0}^{t} P_j(X_1, \ldots, X_{n-1}) X_n^j,$$

where $0 \leq t \leq d$, $P_t(X_1, \ldots, X_{n-1}) \neq 0$ and $\deg(P_j) \leq d - j$.

For any $a_1, \ldots, a_{n-1} \in S$, let $m_{a_1,\ldots,a_{n-1}} = \mathrm{mult}(P_t, (a_1, \ldots, a_{n-1}))$. We will show that

$$\sum_{a_n \in S} \mathrm{mult}(P, (a_1, \ldots, a_n)) \leq m_{a_1,\ldots,a_{n-1}} \cdot |S| + t. \tag{2}$$

Given this, we may then bound

$$\sum_{a_1,\ldots,a_n \in S} \mathrm{mult}(P, (a_1, \ldots, a_n)) \leq \sum_{a_1,\ldots,a_{n-1} \in S} m_{a_1,\ldots,a_{n-1}} \cdot |S| + |S|^{n-1} \cdot t.$$

By the induction hypothesis applied to P_t, we know that

$$\sum_{a_1,\ldots,a_{n-1} \in S} m_{a_1,\ldots,a_{n-1}} \leq \deg(P_t) \cdot |S|^{n-2} \leq (d-t) \cdot |S|^{n-2}.$$

This implies the result.

We now prove Equation (2). Fix $a_1, \ldots, a_{n-1} \in S$ and let $\mathbf{i} = (i_1, \ldots, i_{n-1})$ be such that $\mathrm{wt}(\mathbf{i}) = m_{a_1,\ldots,a_{n-1}}$ and $P_t^{(\mathbf{i})}(X_1, \ldots, X_{n-1}) \neq 0$. Letting $(\mathbf{i}, 0)$ denote the vector $(i_1, \ldots, i_{n-1}, 0)$, we note that

$$P^{(\mathbf{i},0)}(X_1, \ldots, X_n) = \sum_{j=0}^{t} P_j^{(\mathbf{i})}(X_1, \ldots, X_{n-1}) X_n^j,$$

and hence $P^{(\mathbf{i},0)}$ is a nonzero polynomial.

Now by Lemma 5 and Corollary 7, we know that

$$\begin{aligned} \mathrm{mult}(P(X_1, \ldots, X_n), (a_1, \ldots, a_n)) &\leq \mathrm{wt}(\mathbf{i}, 0) + \mathrm{mult}(P^{(\mathbf{i},0)}(X_1, \ldots, X_n), (a_1, \ldots, a_n)) \\ &\leq m_{a_1,\ldots,a_{n-1}} + \mathrm{mult}(P^{(\mathbf{i},0)}(a_1, \ldots, a_{n-1}, X_n), a_n). \end{aligned}$$

Summing this up over all $a_n \in S$, and applying the $n = 1$ case of this lemma to the nonzero univariate degree-t polynomial $P^{(\mathbf{i},0)}(a_1, \ldots, a_{n-1}, X_n)$, we get Equation (2). This completes the proof of the lemma. ■

The following corollary simply states the above lemma in contrapositive form, with $S = \mathbb{F}_q$.

Corollary 9 *Let $P \in \mathbb{F}_q[\mathbf{X}]$ be a polynomial of total degree at most d. If $\sum_{\mathbf{a} \in \mathbb{F}_q^n} \mathrm{mult}(P, \mathbf{a}) > d \cdot q^{n-1}$, then $P(\mathbf{X}) = 0$.*

3 A lower bound on the size of Kakeya sets

We now give a lower bound on the size of Kakeya sets in $\mathbb{F}_q^n$. We implement the plan described in Section 1. Specifically, in Proposition 10 we show that we can find a somewhat low degree non-zero polynomial that vanishes with high multiplicity on any given Kakeya set, where the degree of the polynomial grows with the size of the set. Next, in Claim 12 we show that the homogenous part of this polynomial vanishes with fairly high multiplicity everywhere in $\mathbb{F}_q^n$. Using the strengthened Schwartz-Zippel lemma, we conclude that the homogenous polynomial is identically zero if the Kakeya set is too small, leading to the desired contradiction. The resulting lower bound (slightly stronger than Theorem 1) is given in Theorem 11.

Proposition 10 *Given a set $K \subseteq \mathbb{F}^n$ and non-negative integers m, d such that*

$$\binom{m+n-1}{n} \cdot |K| < \binom{d+n}{n},$$

there exists a non-zero polynomial $P = P_{m,K} \in \mathbb{F}[\mathbf{X}]$ of total degree at most d such that $\mathrm{mult}(P, \mathbf{a}) \geq m$ *for every $\mathbf{a} \in K$.*

Proof The number of possible monomials in P is $\binom{d+n}{n}$. Hence there are $\binom{d+n}{n}$ degrees of freedom in the choice for the coefficients for these monomials. For a given point $\mathbf{a}$, the condition that $\mathrm{mult}(P, \mathbf{a}) \geq m$ imposes $\binom{m+n-1}{n}$ homogeneous linear constraints on the coefficients of P. Since the total number of (homogeneous) linear constraints is $\binom{m+n-1}{n} \cdot |K|$, which is strictly less than the number of unknowns, there is a nontrivial solution.

■

Theorem 11 *If $K \subseteq \mathbb{F}_q^n$ is a Kakeya set, then $|K| \geq \left(\frac{q}{2-1/q}\right)^n$.*

Proof Let ℓ be a large multiple of q and let

$$m = 2\ell - \ell/q$$

$$d = \ell q - 1.$$

These three parameters (ℓ, m and d) will be used as follows: d will be the bound on the degree of a polynomial P which vanishes on K, m will be the multiplicity of the zeros of P on K and ℓ will be the multiplicity of the zeros of the homogenous part of P which we will deduce by restricting P to lines passing through K.

Note that by the choices above we have $d < \ell q$ and $(m - \ell)q > d - \ell$. We prove below later that

$$|K| \geq \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}} \geq \alpha^n$$

where $\alpha \to \frac{q}{2-1/q}$ as $\ell \to \infty$.

Assume for contradiction that $|K| < \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}}$. Then, by Proposition 10 there exists a non-zero polynomial $P(\mathbf{X}) \in \mathbb{F}[\mathbf{X}]$ of total degree exactly d^*, where $d^* \leq d$, such that $\mathrm{mult}(P, \mathbf{x}) \geq m$ for every $\mathbf{x} \in K$. Note that $d^* \geq \ell$ since $d^* \geq m$ (since P is nonzero and vanishes to multiplicity $\geq m$ at some point), and $m \geq \ell$ by choice of m. Let $H_P(\mathbf{X})$ be the homogeneous part of $P(\mathbf{X})$ of degree d^*. Note that $H_P(\mathbf{X})$ is nonzero. The following claim shows that H_P vanishes to multiplicity ℓ at each point of $\mathbb{F}_q^n$.

Claim 12 *For each* $\mathbf{b} \in \mathbb{F}_q^n$.

$$\mathrm{mult}(H_P, \mathbf{b}) \geq \ell.$$

Proof Fix $\mathbf{i}$ with $\mathrm{wt}(\mathbf{i}) = w \leq \ell - 1$. Let $Q(\mathbf{X}) = P^{(\mathbf{i})}(\mathbf{X})$. Let d' be the degree of the polynomial $Q(\mathbf{X})$, and note that $d' \leq d^* - w$.

Let $\mathbf{a} = \mathbf{a}(\mathbf{b})$ be such that $\{\mathbf{a} + t\mathbf{b} | t \in \mathbb{F}_q\} \subset K$. Then for all $t \in \mathbb{F}_q$, by Lemma 5, $\mathrm{mult}(Q, \mathbf{a} + t\mathbf{b}) \geq m - w$. Since $w \leq \ell - 1$ and $(m - \ell) \cdot q > d^* - \ell$, we get that $(m - w) \cdot q > d^* - w$.

Let $Q_{\mathbf{a},\mathbf{b}}(T)$ be the polynomial $Q(\mathbf{a} + T\mathbf{b}) \in \mathbb{F}_q[T]$. Then $Q_{\mathbf{a},\mathbf{b}}(T)$ is a univariate polynomial of degree at most d', and by Corollary 7, it vanishes at each point of $\mathbb{F}_q$ with multiplicity $m - w$. Since

$$(m - w) \cdot q > d^* - w \geq \deg(Q_{\mathbf{a},\mathbf{b}}(T)),$$

we conclude that $Q_{\mathbf{a},\mathbf{b}}(T) = 0$. Hence the coefficient of $T^{d'}$ in $Q_{\mathbf{a},\mathbf{b}}(T)$ is 0. Let H_Q be the homogenous component of Q of highest degree. Observe that the coefficient of $T^{d'}$ in $Q_{\mathbf{a},\mathbf{b}}(T)$ is $H_Q(\mathbf{b})$. Hence $H_Q(\mathbf{b}) = 0$.

However $H_Q(\mathbf{X}) = (H_P)^{(\mathbf{i})}(\mathbf{X})$ (by item 2 of Proposition 4). Hence $(H_P)^{(\mathbf{i})}(\mathbf{b}) = 0$. Since this is true for all $\mathbf{i}$ of weight at most $\ell - 1$, we conclude that $\mathrm{mult}(H_P, \mathbf{b}) \geq \ell$. ■

Applying Corollary 9, and noting that $\ell q^n > d^* q^{n-1}$, we conclude that $H_P(\mathbf{X}) = 0$. This contradicts the fact that $P(\mathbf{X})$ is a nonzero polynomial.

Hence,

$$|K| \geq \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}}$$

Now, by our choice of d and m,

$$\frac{\binom{d+n}{n}}{\binom{m+n-1}{n}} = \frac{\binom{\ell q-1+n}{n}}{\binom{2\ell-\ell/q+n-1}{n}} = \frac{\prod_{i=1}^{n}(\ell q - 1 + i)}{\prod_{i=1}^{n}(2\ell - \ell/q - 1 + i)}$$

Since this is true for all ℓ such that ℓ is a multiple of q, we get that

$$|K| \geq \lim_{\ell\to\infty} \prod_{i=1}^{n} \left(\frac{q - 1/l + i/l}{2 - 1/q - 1/l + i/l} \right) = \left(\frac{q}{2 - 1/q} \right)^n$$

■

4 Statistical Kakeya for curves

Next we extend the results of the previous section to a form conducive to analyze the mergers of Dvir and Wigderson [DW08]. The extension changes two aspects of the consideration in Kakeya sets, that we refer to as "statistical" and "curves". We describe these terms below.

In the setting of Kakeya sets we were given a set K such that for *every* direction, there was a line in that direction such that *every* point on the line was contained in K. In the *statistical* setting we replace both occurrences of the "every" quantifier with a weaker "for many" quantifier. So we consider sets that satisfy the condition that for many directions, there exists a line in that direction intersecting K in many points.

A second change we make is that we now consider curves of higher degree and not just lines. We also do not consider curves in various *directions*, but rather curves passing through a given set of special points. We start with formalizing the terms "curves", "degree" and "passing through a given point".

A *curve of degree k in $\mathbb{F}_q^n$* is a tuple of polynomials $C(X) = (C_1(X), \ldots, C_n(X)) \in \mathbb{F}_q[X]^n$ such that $\max_{i\in[n]} \deg(C_i(X)) = k$. A curve C naturally defines a map from $\mathbb{F}_q$ to $\mathbb{F}_q^n$. For $\mathbf{x} \in \mathbb{F}_q^n$, we say that a curve C *passes through* $\mathbf{x}$ if there is a $t \in \mathbb{F}_q$ such that $C(t) = \mathbf{x}$.

We now state and prove our statistical version of the Kakeya theorem for curves.

Theorem 13 (Statistical Kakeya for curves) *Let $\lambda > 0, \eta > 0$. Let $k > 0$ be an integer such that $\eta q > k$. Let $S \subseteq \mathbb{F}_q^n$ be such that $|S| = \lambda q^n$. Let $K \subseteq \mathbb{F}_q^n$ be such that for each $\mathbf{x} \in S$, there exists a curve $C_{\mathbf{x}}$ of degree at most k that passes through $\mathbf{x}$, and intersects K in at least ηq points. Then,*

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

In particular, if $\lambda \geq \eta$ we get that $|K| \geq \left(\frac{\eta q}{k+1}\right)^n$.

Observe that when $\lambda = \eta = 1$, and $k = 1$, we get the same bound as that for Kakeya sets as obtained in Theorem 11.

Proof Let ℓ be a large integer and let

$$d = \lambda \ell q - 1$$

$$m = k\frac{\lambda \ell q - 1 - (\ell - 1)}{\eta q} + \ell.$$

By our choice of m and d, we have $\eta q(m - (\ell - 1)) > k(d - (\ell - 1))$. Since $\eta q > k$, we have that for all w such that $0 \leq w \leq \ell - 1$, $\eta q(m - w) > k(d - w)$. Just as in the proof of Theorem 11, we will prove that

$$|K| \geq \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}} \geq \alpha^n$$

where $\alpha \to \frac{\lambda q}{k\left(\frac{\lambda q-1}{\eta q}\right)+1}$ as $\ell \to \infty$.

If possible, let $|K| < \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}}$. As before, by Proposition 10 there exists a non-zero polynomial $P(\mathbf{X}) \in \mathbb{F}_q[\mathbf{X}]$ of total degree d^*, where $d^* \leq d$, such that $\mathrm{mult}(P, \mathbf{a}) \geq m$ for every $\mathbf{a} \in K$. We will deduce that in fact P must vanish on all points in S with multiplicity ℓ. We will then get the desired contradiction from Corollary 9.

Claim 14 *For each* $\mathbf{x}_0 \in S$,

$$\mathrm{mult}(P, \mathbf{x}_0) \geq \ell.$$

Proof Fix any $\mathbf{i}$ with $\mathrm{wt}(\mathbf{i}) = w \leq \ell - 1$. Let $Q(\mathbf{X}) = P^{(\mathbf{i})}(\mathbf{X})$. Note that $Q(\mathbf{X})$ is a polynomial of degree at most $d^* - w$. By Lemma 5, for all points $\mathbf{a} \in K$, $\mathrm{mult}(Q, \mathbf{a}) \geq m - w$.

Let $C_{\mathbf{x}_0}$ be the curve of degree k through $\mathbf{x}_0$, that intersects K in at least ηq points. Let $t_0 \in \mathbb{F}_q$ be such that $C_{\mathbf{x}_0}(t_0) = \mathbf{x}_0$. Let $Q_{\mathbf{x}_0}(T)$ be the polynomial $Q \circ C_{\mathbf{x}_0}(T) \in \mathbb{F}_q[T]$. Then $Q_{\mathbf{x}_0}(T)$ is a univariate polynomial of degree at most $k(d^* - w)$. By Corollary 7, for all points $t \in \mathbb{F}_q$ such that $C_{\mathbf{x}_0}(t) \in K$, $Q_{\mathbf{x}_0}(T)$ vanishes at t with multiplicity $m - w$. Since the number of such points t is at least ηq, we get that $Q_{\mathbf{x}_0}(T)$ has at least $\eta q(m - w)$ zeros (counted with multiplicity). However, by our choice of parameters, we know that

$$\eta q(m - w) > k(d - w) \geq k(d^* - w) \geq \deg(Q_{\mathbf{x}_0}(T)).$$

Since the degree of $Q_{\mathbf{x}_0}(T)$ is strictly less than the number of its zeros, $Q_{\mathbf{x}_0}(T)$ must be identically zero. Thus we get $Q_{\mathbf{x}_0}(t_0) = Q(C_{\mathbf{x}_0}(t_0)) = Q(\mathbf{x}_0) = 0$ Hence $P^{(\mathbf{i})}(\mathbf{x}_0) = 0$. Since this is true for all $\mathbf{i}$ with $\mathrm{wt}(\mathbf{i}) \leq \ell - 1$, we conclude that $\mathrm{mult}(P, \mathbf{x}_0) \geq \ell$. ■

Thus P vanishes at every point in S with multiplicity ℓ. As $P(\mathbf{X})$ is a non-zero polynomial, Corollary 9 implies that $\ell|S| \leq d^* q^{n-1}$. Hence $\ell \lambda q^n \leq d q^{n-1}$, which contradicts the choice of d.

Thus $|K| \geq \frac{\binom{d+n}{n}}{\binom{m+n-1}{n}}$. By choice of d and m,

$$|K| \geq \frac{\binom{\lambda \ell q - 1 + n}{n}}{\binom{k\frac{\lambda \ell q - 1 - (\ell - 1)}{\eta q} + \ell + n - 1}{n}}.$$

Picking ℓ arbitrarily large, we conclude that

$$|K| \geq \lim_{\ell \to \infty} \frac{\binom{\lambda \ell q - 1 + n}{n}}{\binom{k\frac{\lambda \ell q - 1 - (\ell - 1)}{\eta q} + \ell + n - 1}{n}} = \lim_{\ell \to \infty} \left(\frac{\ell \lambda q - 1}{\ell k \left(\frac{\lambda q - 1}{\eta q}\right) + \ell} \right)^n = \left(\frac{\lambda q}{k \left(\frac{\lambda q - 1}{\eta q}\right) + 1} \right)^n.$$

■

5 Improved Mergers

In this section we state and prove our main result on randomness mergers.

5.1 Definitions and Theorem Statement

We start by recalling some basic quantities associated with random variables. The statistical distance between two random variables X and Y taking values from a finite domain Ω is defined as

$$\max_{S \subseteq \Omega} |\mathbf{Pr}[X \in S] - \mathbf{Pr}[Y \in S]| \, .$$

We say that X is ϵ-close to Y if the statistical distance between X and Y is at most ϵ, otherwise we say that X and Y are ϵ-far. The min-entropy of a random variable X is defined as

$$H_\infty(X) \triangleq \min_{x \in \mathrm{supp}(X)} \log_2 \left(\frac{1}{\mathbf{Pr}[X = x]} \right) .$$

We say that a random variable X is ϵ-close to having min-entropy m if there exists a random variable Y of min-entropy m such that X is ϵ-close to Y.

A "merger" of randomness takes a k-tuple of random variables and "merges" their randomness to produce a high-entropy random variable, provided the k-tuple is "somewhere random" as defined below.

Definition 15 (Somewhere-random source) *For integers k and N an (N,k)-*somewhere-random source *is a random variable $\mathsf{A} = (\mathsf{A}_1, \ldots, \mathsf{A}_k)$ taking values in S^k, where S is some finite set of cardinality 2^N, such that for some $i_0 \in [k]$, the distribution of A_{i_0} is uniform over S. (When N and k are clear from context we refer to the source as simply a "somewhere random source".)*

We are now ready to define a merger.

Definition 16 (Merger) *For positive integer k and set S of size 2^N, a function $f : S^k \times \{0,1\}^d \to S$ is called an (m, ϵ)-*merger *(of (N,k)-somewhere random sources), if for every (N,k) somewhere-random source $\mathsf{A} = (\mathsf{A}_1, \ldots, \mathsf{A}_k)$ taking values in S^k, and for B being uniformly distributed over $\{0,1\}^d$, the distribution of $f((\mathsf{A}_1, \ldots, \mathsf{A}_k), \mathsf{B})$ is ϵ-close to having min-entropy m.*

A merger thus has five parameters associated with it: N, k, m, ϵ and d. The general goal is to give explicit constructions of mergers of (N,k)-somewhere random sources for every choice of N and k, for as large an m as possible, and with ϵ and d being as small as possible. Known mergers attain $m = (1-\delta) \cdot N$ for arbitrarily small δ and our goal will be to achieve $\delta = o(1)$ as a function of N, while ϵ is an arbitrarily small positive real number. Thus our main concern is the growth of d as a function of N and k. Prior to this work, the best known bounds required either $d = \Omega(\log N + \log k)$ or $d = \Omega(k)$. We only require $d = \Omega(\log k)$.

Theorem 17 *For every $\epsilon, \delta > 0$ and integers N, k, there exists a $((1-\delta) \cdot N, \epsilon)$-merger of (N,k)-somewhere random sources, computable in polynomial time, with seed length*

$$d = \frac{1}{\delta} \cdot \log_2 \left(\frac{2k}{\epsilon} \right) .$$

5.2 The Curve Merger of [DW08] and its analysis

The merger that we consider is a very simple one proposed by Dvir and Wigderson [DW08], and we improve their analysis using our extended method of multiplicities. We note that they used the polynomial method in their analysis; and the basic method of multiplicities doesn't seem to improve their analysis.

The curve merger of [DW08], denoted f_{DW}, is obtained as follows. Let $q \geq k$ be a prime power, and let n be any integer. Let $\gamma_1, \ldots, \gamma_k \in \mathbb{F}_q$ be distinct, and let $c_i(T) \in \mathbb{F}_q[T]$ be the unique degree $k-1$ polynomial with $c_i(\gamma_i) = 1$ and for all $j \neq i$, $c_i(\gamma_j) = 0$. Then for any $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in (\mathbb{F}_q^n)^k$ and $u \in \mathbb{F}_q$, the curve merger f_{DW} maps $(\mathbb{F}_q^n)^k \times \mathbb{F}_q$ to $\mathbb{F}_q^n$ as follows:

$$f_{\mathrm{DW}}((\mathbf{x}_1, \ldots, \mathbf{x}_k), u) = \sum_{i=1}^{k} c_i(u)\mathbf{x}_i.$$

In other words, $f_{\mathrm{DW}}((\mathbf{x}_1, \ldots, \mathbf{x}_k), u)$ picks the (canonical) curve passing through $\mathbf{x}_1, \ldots, \mathbf{x}_k$ and outputs the uth point on the curve..

Theorem 18 *Let $q \geq k$ and A be somewhere-random source taking values in $(\mathbb{F}_q^n)^k$. Let B be distributed uniformly over $\mathbb{F}_q$. Let $\mathsf{C} = f_{\mathrm{DW}}(\mathsf{A}, \mathsf{B})$. Then for*

$$q \geq \left(\frac{2k}{\epsilon}\right)^{\frac{1}{\delta}},$$

C is ϵ-close to having min-entropy $(1-\delta) \cdot n \cdot \log_2 q$.

Theorem 17 easily follows from the above. We note that [DW08] proved a similar theorem assuming $q \geq \mathrm{poly}(n, k)$, forcing their seed length to grow logarithmically with n as well.

Proof of Theorem 17: Let $q = 2^d$, so that $q \geq \left(\frac{2k}{\epsilon}\right)^{\frac{1}{\delta}}$, and let $n = N/d$. Then we may identify identify $\mathbb{F}_q$ with $\{0,1\}^d$ and $\mathbb{F}_q^n$ with $\{0,1\}^N$. Take f to be the function f_{DW} given earlier. Clearly f is computable in the claimed time. Theorem 18 shows that f has the required merger property. ■

We now prove Theorem 18.

Proof of Theorem 18: Let $m = (1-\delta) \cdot n \cdot \log_2 q$. We wish to show that $f_{\mathrm{DW}}(\mathsf{A}, \mathsf{B})$ is ϵ-close to having min-entropy m.

Suppose not. Then there is a set $K \subseteq \mathbb{F}_q^n$ with $|K| \leq 2^m = q^{(1-\delta)\cdot n} \leq \left(\frac{\epsilon q}{2k}\right)^n$ such that

$$\Pr_{\mathsf{A},\mathsf{B}}[f(\mathsf{A}, \mathsf{B}) \in K] \geq \epsilon.$$

Suppose A_{i_0} is uniformly distributed over $\mathbb{F}_q^n$. Let A_{-i_0} denote the random variable $(\mathsf{A}_1, \ldots, \mathsf{A}_{i_0-1}, \mathsf{A}_{i_0+1}, \ldots, \mathsf{A}_k)$. By an averaging argument, with probability at least $\lambda = \epsilon/2$ over the choice of A_{i_0}, we have

$$\Pr_{\mathsf{A}_{-i_0},\mathsf{B}}[f(\mathsf{A}, \mathsf{B}) \in K] \geq \eta,$$

where $\eta = \epsilon/2$. Since A_{i_0} is uniformly distributed over $\mathbb{F}_q^n$, we conclude that there is a set S of cardinality at least λq^n such that for any $\mathbf{x} \in S$,

$$\Pr_{\mathsf{A},\mathsf{B}}[f(\mathsf{A},\mathsf{B}) \in K \mid \mathsf{A}_{i_0} = \mathbf{x}] \geq \eta.$$

Fixing the values of A_{-i_0}, we conclude that for each $\mathbf{x} \in S$, there is a $\mathbf{y} = \mathbf{y}(\mathbf{x}) = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ with $\mathbf{y}_{i_0} = x$ such that $\Pr_{\mathsf{B}}[f(\mathbf{y},\mathsf{B}) \in K] \geq \eta$. Define the degree $k-1$ curve $C_{\mathbf{x}}(T) = f(\mathbf{y}(\mathbf{x}), T) = \sum_{j=1}^{k} \mathbf{y}_j c_j(T)$. Then $C_{\mathbf{x}}$ passes through $\mathbf{x}$, since $C_{\mathbf{x}}(\gamma_{i_0}) = \sum_{j=1}^{k} \mathbf{y}_j c_j(\gamma_{i_0}) = \mathbf{y}_{i_0} = \mathbf{x}$, and $\Pr_{\mathsf{B}\in\mathbb{F}_q}[C_{\mathbf{x}}(\mathsf{B}) \in K] \geq \eta$ by definition of $C_{\mathbf{x}}$.

Thus S and K satisfy the hypothesis of Theorem 13. We now conclude that

$$|K| \geq \left(\frac{\lambda q}{(k-1)\left(\frac{\lambda q - 1}{\eta q}\right) + 1} \right)^n = \left(\frac{\epsilon q/2}{k - (k-1)/\eta q} \right)^n > \left(\frac{\epsilon q}{2k} \right)^n.$$

This is a contradiction, and the proof of the theorem is complete. ■

The Somewhere-High-Entropy case: It is possible to extend the merger analysis given above also to the case of *somewhere-high-entropy* sources. In this scenario the source is comprised of blocks, one of which has min entropy at least r. One can then prove an analog of Theorem 18 saying that the output of f_{DW} will be close to having min entropy $(1-\delta)\cdot r$ under essentially the same conditions on q. The proof is done by hashing the source using a random linear function into a smaller dimensional space and then applying Theorem 18 (in a black box manner). The reason why this works is that the merger commutes with the linear map (for details see [DW08]).

6 Bounds on the list size for list-decoding Reed-Solomon codes

In this section, we give a simple algebraic proof of an upper bound on the list size for list-decoding Reed-Solomon codes within the Johnson radius.

Before stating and proving the theorem, we need some definitions. For a bivariate polynomial $P(X,Y) \in \mathbb{F}[X,Y]$, we define its (a,b)-degree to be the maximum of $ai + bj$ over all (i,j) such that the monomial X^iY^j appears in $P(X,Y)$ with a nonzero coefficient. Let $N(k,d,\theta)$ be the number of monomials X^iY^j which have $(1,k)$-degree at most d and $j \leq \theta d/k$. We have the following simple fact.

Fact 19 *For any $k < d$ and $\theta \in [0,1]$, $N(k,d,\theta) > \theta \cdot (2-\theta) \cdot \frac{d^2}{2k}$.*

Now we prove the main theorem of this section. The proof is an enhancement of the original analysis of the Guruswami-Sudan algorithm using the extended method of multiplicities.

Theorem 20 (List size bound for Reed-Solomon codes) *Let $(\alpha_1,\beta_1),\ldots,(\alpha_n,\beta_n) \in \mathbb{F}^2$. Let $R, \gamma \in [0,1]$ with $\gamma^2 > R$. Let $k = Rn$. Let $f_1(X),\ldots,f_L(X) \in \mathbb{F}[X]$ be polynomials of degree at most k, such that for each $j \in [L]$ we have $|\{i \in [n] : f_j(\alpha_i) = \beta_i\}| > \gamma n$. Then $L \leq \frac{2\gamma}{\gamma^2 - R}$.*

Proof Let $\epsilon > 0$ be a parameter. Let $\theta = \frac{2}{\left(1+\frac{\gamma^2}{R}\right)}$. Let m be a large integer (to be chosen later), and let $d = (1+\epsilon)\cdot m \cdot \sqrt{\frac{nk}{\theta\cdot(2-\theta)}}$. We first interpolate a nonzero polynomial $P(X,Y) \in \mathbb{F}[X,Y]$ of $(1,k)$-degree at most d and Y-degree at most $\theta d/k$, that vanishes with multiplicity at least m at each of the points (α_i, β_i). Such a polynomial exists if $N(k,d,\theta)$, the number of monomials available, is larger than the number of homogeneous linear constraints imposed by the vanishing conditions:

$$\frac{m(m+1)}{2}\cdot n < N(k,d,\theta). \tag{3}$$

This can be made to hold by picking m sufficiently large, since by Fact 19,

$$N(k,d,\theta) > \theta\cdot(2-\theta)\frac{d^2}{2k} = \frac{(1+\epsilon)^2 m^2}{2}\cdot n.$$

Having obtained the polynomial $P(X,Y)$, we also view it as a univariate polynomial $Q(Y) \in \mathbb{F}(X)[Y]$ with coefficients in, $\mathbb{F}(X)$, the field of rational functions in X.

Now let $f(X)$ be any polynomial of degree at most k such that, letting $I = \{i \in [n] : f(\alpha_i) = \beta_i\}$, $|I| \geq A$. We claim that the polynomial $Q(Y)$ vanishes at $f(X)$ with multiplicity at least $m - d/A$. Indeed, fix an integer $j < m - d/A$, and let $R_j(X) = Q^{(j)}(f(X)) = P^{(0,j)}(X, f(X))$. Notice the degree of $R_j(X)$ is at most d. By Proposition 6 and Lemma 5,

$$\mathrm{mult}(R_j, \alpha_i) \geq \mathrm{mult}(P^{(0,j)}, (\alpha_i,\beta_i)) \geq \mathrm{mult}(P, (\alpha_i,\beta_i)) - j.$$

Thus

$$\sum_{i\in I}\mathrm{mult}(R_j,\alpha_i) \geq |I|\cdot(m-j) \geq A\cdot(m-j) > d.$$

By Lemma 8, we conclude that $R_j(X) = 0$. Since this holds for every $j < m - d/A$, we conclude that $\mathrm{mult}(Q, f(X)) \geq m - d/A$.

We now complete the proof of the theorem. By the above discussion, for each $j \in [L]$, we know that $\mathrm{mult}(Q, f_j(X)) \geq m - \frac{d}{\gamma n}$. Thus, by Lemma 8 (applied to the nonzero polynomial $Q(Y) \in \mathbb{F}(X)[Y]$ and the set of evaluation points $S = \{f_j(X) : j \in [L]\}$)

$$\deg(Q) \geq \sum_{j\in[L]}\mathrm{mult}(Q, f(X)) \geq \left(m - \frac{d}{\gamma n}\right)\cdot L.$$

Since $\deg(Q) \leq \theta d/k$, we get,

$$\theta d/k \geq \left(m - \frac{d}{\gamma n}\right)\cdot L.$$

Using $d = (1+\epsilon)\cdot m\cdot\sqrt{\frac{nk}{\theta\cdot(2-\theta)}}$ and $\theta = \frac{2}{1+\frac{\gamma^2}{R}}$, we get,

$$L \leq \frac{\theta}{k\cdot\frac{m}{d} - \frac{k}{\gamma n}} = \frac{\theta}{\frac{1}{1+\epsilon}\sqrt{\frac{k}{n}\cdot\theta\cdot(2-\theta)} - \frac{k}{\gamma n}} = \frac{1}{\frac{1}{1+\epsilon}\sqrt{R\left(\frac{2}{\theta}-1\right)} - \frac{R}{\theta\gamma}} = \frac{1}{\frac{\gamma}{1+\epsilon} - \left(\frac{\gamma}{2} + \frac{R}{2\gamma}\right)}.$$

Letting $\epsilon \to 0$, we get $L \leq \frac{2\gamma}{\gamma^2 - R}$, as desired. ■

References

[CB04] Yuval Cassuto and Jehoshua Bruck. A combinatorial bound on the list size. *Paradise Laboratory Technical report, California Institute of Technology*, 2004.

[DS07] Z. Dvir and A. Shpilka. An improved analysis of linear mergers. *Comput. Complex.*, 16(1):34–59, 2007. (Extended abstract appeared in RANDOM 2005).

[Dvi08] Z. Dvir. On the size of Kakeya sets in finite fields. *J. AMS (to appear)*, 2008.

[DW08] Zeev Dvir and Avi Wigderson. Kakeya sets, new mergers and old extractors. In *FOCS*, pages 625–633. IEEE Computer Society, 2008.

[GR06] Venkatesan Guruswami and Atri Rudra. Explicit capacity-achieving list-decodable codes. In Jon M. Kleinberg, editor, *STOC*, pages 1–10. ACM, 2006.

[GS99] Venkatesan Guruswami and Madhu Sudan. Improved decoding of Reed-Solomon and algebraic-geometric codes. *IEEE Transactions on Information Theory*, 45:1757–1767, 1999.

[GUV07] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. Unbalanced expanders and randomness extractors from parvaresh-vardy codes. In *IEEE Conference on Computational Complexity*, pages 96–108. IEEE Computer Society, 2007.

[HKT08] J. W. P. Hirschfeld, G. Korchmaros, and F. Torres. *Algebraic Curves over a Finite Field (Princeton Series in Applied Mathematics)*. Princeton University Press, 2008.

[LRVW03] Chi-Jen Lu, Omer Reingold, Salil Vadhan, and Avi Wigderson. Extractors: Optimal up to constant factors. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing"*, 2003.

[PV05] Farzad Parvaresh and Alexander Vardy. Correcting errors beyond the Guruswami-Sudan radius in polynomial time. In *FOCS*, pages 285–294. IEEE Computer Society, 2005.

[SS08] Shubhangi Saraf and Madhu Sudan. Improved lower bound on the size of kakeya sets over finite fields. *Analysis and PDE (to appear)*, 2008.

[Sud97] Madhu Sudan. Decoding of Reed-Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180–193, 1997.

[TS96] A. Ta-Shma. On extracting randomness from weak random sources. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA*, pages 276–285, 1996.

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

[Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(1):103–128, 2007.

ECCC ISSN 1433-8092
http://eccc.hpi-web.de/