Electronic Colloquium on Computational Complexity, Report No. 44 (2010)

Affine Dispersers from Subspace Polynomials

Eli Ben-Sasson* Swastik Kopparty †

March 12, 2010

Abstract

Dispersers and *extractors* for affine sources of dimension d in $\mathbb{F}_p^n$ — where $\mathbb{F}_p$ denotes the finite field of prime size p — are functions $f : \mathbb{F}_p^n \to \mathbb{F}_p$ that behave pseudorandomly when their domain is restricted to any particular affine space $S \subseteq \mathbb{F}_p^n$ of dimension at least d. For dispersers, "pseudorandom behavior" means that f is nonconstant over S, i.e., $|\{f(s) \mid s \in S\}| > 1$. For extractors, it means that $f(s)$ is distributed almost uniformly over $\mathbb{F}_p$ when s is distributed uniformly over S. Dispersers and extractors for affine sources have been considered in the context of deterministic extraction of randomness from structured sources of imperfect randomness. Previously, explicit constructions of affine dispersers were known for every $d = \Omega(n)$, due to Barak, Kindler, Shaltiel, Sudakov, and Wigderson [2005] and explicit affine extractors for the same dimension were obtained by Bourgain [2007].

The main result of this paper is an efficient deterministic construction of affine dispersers for *sublinear* dimension $d = \Omega(n^{4/5})$. Additional results include a new and simple affine extractor for dimension $d > 2n/5$, and a simple disperser for multiple independent affine sources. The main novelty in this paper lies in the method of proof, which makes use of classical algebraic objects called *subspace polynomials*. In contrast, the papers mentioned above relied on the sum-product theorem for finite fields and other recent results from additive combinatorics.

*Computer Science Department, Technion — Israel Institute of Technology, Haifa, 32000, Israel. Email: eli@cs.technion.ac.il.Research of both co-authors supported in part by a European Community International Reintegration Grant, an Alon Fellowship, and grants by the Israeli Science Foundation (grant number 679/06) and by the US-Israel Binational Science Foundation.

†CSAIL, MIT, Cambridge, MA 02139. Email: swastik@mit.edu. Research supported in part by NSF Award CCR-0514915.

ISSN 1433-8092

Contents

1 Introduction

A *one-output-bit seedless disperser* (often called a deterministic disperser) for a family $\mathcal{F}$ of subsets of $\{0,1\}^m$ is a function $\mathrm{Disp} : \{0,1\}^m \to \{0,1\}$ satisfying the property that on any subset $X \in \mathcal{F}$, $X \subset \{0,1\}^m$ (the set X is often called a "source" in the derandomization literature) the function Disp takes more than one value, i.e., $|\{\mathrm{Disp}(x) : x \in X\}| > 1$. An *extractor* for $\mathcal{F}$ is a function $\mathrm{Extr} : \{0,1\}^m \to \{0,1\}$ satisfying the stronger requirement that for every $X \in \mathcal{F}$, if x is picked uniformly over X, then $\mathrm{Extr}(x)$ is nearly-uniformly distributed. We think of dispersers and extractors as behaving *pseudorandomly* on sources $X \in \mathcal{F}$ because in typical settings where the size of $\mathcal{F}$ is not too large, a random function is indeed an extractor and hence also a disperser. Extractors and dispersers have been intensively studied in recent years in the context of extracting randomness from imperfect sources of randomness. The goal of these studies has been to obtain extractors and dispersers computable in polynomial time, and today several constructions of seedless dispersers for various structured families of subsets are known, including for "bit-fixing" and "samplable" sources [Chor et al., 1985, Gabizon et al., 2006, Trevisan and Vadhan, 2000, Kamp and Zuckerman, 2007]. We refer the reader to [Barak et al., 2005] for more information on seedless dispersers and extractors.

A particularly interesting family of structured subsets that has been considered in this context, and is also the focus of our paper, is the family of affine subspaces over a fixed finite field $\mathbb{F}_p$ of size p (think of $p = 2$). Extractors and dispersers for this family of sources are known as affine extractors and dispersers. Affine extractors for spaces of dimension greater than $m/2$ are relatively easy to construct [Ben-Sasson et al., 2001]. However, for spaces of dimension smaller than $m/2$ the problem becomes much harder, and to date, only two explicit pseudorandom constructions are known [Barak et al., 2005, Bourgain, 2007][1]. Both these works give constructions that are shown to behave pseudorandomly on all affine spaces of dimension $\geq \epsilon m$, where $\epsilon > 0$ is any fixed constant. The work of Barak et al. [2005] constructs affine dispersers and that of Bourgain [2007] constructs affine extractors. Both constructions use recent sum-product theorems over finite fields [Bourgain et al., 2004, 2006] and related results from additive combinatorics, along with several other non-trivial ideas.

1.1 Results

Our main result (Theorem 2.2) is the explicit construction of an affine disperser for spaces of dimension $o(m)$. Specifically, our disperser works for spaces of dimension at least $6m^{4/5}$. The structure of our main affine disperser is as follows. The m coordinates are grouped into r blocks, each with an equal number k of coordinates, and each block is interpreted as specifying an element of the finite field $\mathbb{F}_{p^k}$. The r elements thus obtained in $\mathbb{F}_{p^k}$ are now substituted into a certain polynomial over $\mathbb{F}_{p^k}$, and its output, which is an element of $\mathbb{F}_{p^k}$, is projected onto $\mathbb{F}_p$ via a nontrivial $\mathbb{F}_p$-linear mapping of $\mathbb{F}_{p^k}$ to $\mathbb{F}_p$.

The techniques we use allow for a host of results with a similar flavor. The simplest-to-state result is a "univariate" affine extractor below the $m/2$ barrier. By "univariate" we mean that the function

[1] A related, though incomparable, result of Gabizon and Raz [2005] constructs extractors for affine sources over "large" finite fields, where "large" means $p > m^2$, see also DeVos and Gabizon [2009] for recent improvements along this line of research.

we use to compute the extractor is naturally viewed as a univariate polynomial. Let $\phi : \mathbb{F}_p^m \to \mathbb{F}_{p^m}$ be any $\mathbb{F}_p$-linear isomorphism and $\pi : \mathbb{F}_{p^m} \to \mathbb{F}_p$ be any nontrivial $\mathbb{F}_p$-linear map[2]. We show that the function $f : \mathbb{F}_p^m \to \mathbb{F}_p$ defined by

$$f(x) = \pi\left((\phi(x))^{1+p+p^2}\right) \tag{1}$$

is an extractor for dimension at least $2m/5 + O(1)$, as long as m is odd. Another pseudorandom univariate construction appearing in this paper is

$$f(x) = \pi\left((\phi(x))^{1+p+p^2+p^3}\right) \tag{2}$$

which we prove is an affine disperser for dimension greater than $n/3 + O(1)$. We conjecture that this construction is in fact an extractor and that both univariate constructions are merely the first two members of a larger family of univariate extractors (see Conjecture 2.8).

We point out that if m is even, then $\mathbb{F}_{p^m}$ has a subfield $\mathbb{F}_{p^{m/2}}$ which is also a $m/2$-dimensional subspace of $\mathbb{F}_{p^m}$ for which the above mentioned constructions will not be a disperser. Indeed, when x belongs to $\mathbb{F}_{p^{m/2}}$ then so does every power of x, hence some nontrivial $\mathbb{F}_p$-linear map π will be constant on both $\left\{x^{1+p+p^2} \;\middle|\; x \in \mathbb{F}_{p^{m/2}}\right\}$ and $\left\{x^{1+p+p^2+p^3} \;\middle|\; x \in \mathbb{F}_{p^{m/2}}\right\}$. In the next section we comment on the role that the oddness of m, and more generally, the absence of subfields, plays in our proofs.

On subspaces and polynomials Our analysis makes use of a class of polynomials called *subspace polynomials*. These polynomials were first systematically studied by Ore in the 1930's [Ore, 1933, 1934]. They have numerous applications in the study of finite fields and in the theory of error correcting codes (See Berlekamp [1968, Chapter 11] and Lidl and Niederreiter [1997, Chapter 3, Section 4]). More recently, they have been used within computational complexity to construct short PCPs [Ben-Sasson et al., 2004, Ben-Sasson and Sudan, 2005, Ben-Sasson et al., 2005] and to study limitations on the list-decodability of the Reed-Solomon code [Ben-Sasson et al., 2006].

The polynomials studied in this last line of works are what we call the *kernel-subspace*[3] polynomial associated to a linear subspace $L \subseteq \mathbb{F}_{p^m}$, which is a polynomial whose set of roots equals L. In this work we analyze our dispersers using elementary properties of the *image-subspace polynomial* of a linear subspace L. These polynomials have the property that their image, i.e., the set of values they take over $\mathbb{F}_{p^m}$, equals L. Our proofs begin by first reformulating the property of being an affine disperser in terms of these polynomials. We then use a simple-to-prove, yet extremely powerful, structural lemma about these polynomials, to get our main results.

Pseudorandomness from the absence of subfields It was recently realized, starting with the work of Barak et al. [2004] and further developed in [Zuckerman, 2006, Kamp et al., 2006, Barak et al., 2006, Bourgain, 2007], that finite fields without large subfields are the source of many

[2]Explicitly, ϕ is a bijection between the vector space $\mathbb{F}_p^m$ and the finite field $\mathbb{F}_{p^m}$ and π is a nonzero linear map from $\mathbb{F}_{p^m}$ to $\mathbb{F}_p$. Both mappings are $\mathbb{F}_p$-linear, i.e., they respect addition and multiplication by scalars in $\mathbb{F}_p$.

[3]The terms "kernel-" and "image-subspace polynomials" were suggested by Prahladh Harsha and we thank him for introducing this nomenclature.

pseudorandom phenomena, and that this can be put to good use in the construction of extractors and dispersers. The above mentioned works all harnessed this pseudorandomness via recent results from additive combinatorics such as the sum-product theorem of Bourgain, Katz, and Tao [2004] and the related multilinear exponential sum estimates of Bourgain, Glibichuk, and Konyagin [2006].

In our work, we offer a different algebraic incarnation of this phenomenon. Specifically, we show that the absence of large subfields directly affects the structure of the image-subspace polynomials of the field. Image-subspace polynomials are *linearized*, which means that they are of the form $\sum_{i=0}^{m-1} a_i X^{p^i}$. Roughly speaking, our main structural lemma (Lemma 3.10) says that the image-subspace polynomial of a subspace $\mathcal{A}$ of dimension d cannot have d consecutive coefficients a_i that are all zero. Moreover, and this is the crucial part, if $\mathcal{A}$ is not contained in a constant multiple of a subfield of $\mathbb{F}_{p^n}$, then the polynomial cannot have even $d-1$ consecutive coefficients that are all zero. This lemma has a short proof (appearing in Section 3.3), yet is extremely powerful. Surprisingly, reducing the maximal length of a sequence of zero-coefficients by 1 (from d to $d-1$) for spaces that are not contained in subfields is all it takes for the underlying pseudorandomness to get exposed.

1.2 Proof strategy for affine dispersers

We now give a brief description of the basic proof strategy that we use to prove that a function is an affine disperser. We demonstrate the steps involved in the special case of the function f defined in (1) for the case of $p = 2$ and $\pi(y) = \mathrm{Tr}(y)$ (where $\mathrm{Tr} : \mathbb{F}_{2^m} \to \mathbb{F}_2$ is the Trace map). We will show in Theorem 2.6 that if m is odd (so that $\mathbb{F}_{2^m}$ has no proper subfields of size $2^{m/2}$), then for any affine space $\mathcal{A} \subseteq \mathbb{F}_{2^m}$ of dimension $\geq 2m/5 + \Omega(1)$, we have $\left\{\mathrm{Tr}(a^7) \;\middle|\; a \in \mathcal{A}\right\} = \mathbb{F}_2$.

1. **Reduce to showing that a certain polynomial h is not a constant polynomial:** We first parameterize the affine space $\mathcal{A}$ using subspace polynomials. Let $Q(X)$ be the image-subspace of $\mathcal{A}$, so that $\mathcal{A} = \{Q(x) : x \in \mathbb{F}_{2^m}\}$. In terms of the polynomial $Q(X)$, we want to show that the composed map $f \circ Q : \mathbb{F}_{2^m} \to \mathbb{F}_2$ is non-constant. Let $h(X)$ be the polynomial $\mathrm{Tr}(Q(X)^7) \mod \langle X^{2^m} - X\rangle$, so that $h(x) = f(Q(x))$ for each $x \in \mathbb{F}_{2^m}$ (cf. Proposition 3.1). Thus to show that h is a nonconstant map, it suffices to show that $h(X)$ is a nonconstant polynomial. We do this in the next two steps of our proof strategy, by finding a monomial of positive degree that appears in $h(X)$ with a nonzero coefficient.

2. **Express the coefficients of h in terms of the coefficients of the subspace polynomials:** To show that h has a monomial of positive degree with a nonzero coefficient, it will be convenient to get an explicit expression for the coefficients themselves. Such an explicit expression can be obtained by direct substitution. In all the cases we consider, there is a good deal of structure in the resulting formulae. For example, for the polynomial we obtained while studying $f(x) = \mathrm{Tr}(x^7)$, we have the following lemma.

   Lemma 1.1 *Let $Q(X) = \sum_{i=0}^{m-1} a_i X^{2^i}$. Let $h(X) = \mathrm{Tr}(Q(X)^7) \mod \langle X^{2^m} - X\rangle$. Then for distinct j, k, l, the coefficient of $X^{2^j+2^k+2^l}$ in $h(X)$ is given by the expression:*

$$\sum_{r=0}^{m-1} \mathrm{Perm}\begin{pmatrix} a_{j-r} & a_{k-r} & a_{l-r} \\ a_{j-r-1}^2 & a_{k-r-1}^2 & a_{l-r-1}^2 \\ a_{j-r-2}^4 & a_{k-r-2}^4 & a_{l-r-2}^4 \end{pmatrix}^{2^r}, \tag{3}$$

where Perm *is the matrix permanent, and the subscripts of the a's are taken mod m.*

3. **Argue combinatorially that some coefficient of h must be nonzero:** Finally, we show that some positive degree monomial of h has a nonzero coefficient. Using the regular form of the coefficients of the polynomial h, for example as given in Lemma 1.1, and the structural results about the coefficients of subspace polynomials, this part of the argument reduces to the combinatorics of cyclic shifts on $\mathbb{Z}_m$. More to the point, we use our main structural lemma (Lemma 3.10) to prove that there is a choice of j, k, l such that *(i)* the matrix appearing in the first summand (corresponding to $r = 0$) in equation (3) is lower triangular with nonzero entries on its diagonal, hence its permanent is nonzero, whereas *(ii)* the matrices appearing in all other summands in equation (3) (corresponding to $r = 1, \ldots, m-1$) contain a zero column, hence have a zero permanent.

1.3 From affine dispersers to extractors

We believe that all constructions provided in this paper are affine extractors, not merely dispersers. We can prove this only for our simplest construction, that described in (1). This proof goes via a general theorem saying that *every* degree-3 function that is an affine disperser for dimension d, is also an affine extractor for dimension $d + d'$, with the bias decreasing as the dimension-redundancy d' increases. Here, a degree-3 function is one that, when viewed as an m-variate polynomial over $\mathbb{F}_p$, is cubic, i.e., has degree at most 3. Indeed, the function described in (1) is of this form (cf. Proposition 3.2) whereas that appearing in (2) is already of degree 4 and the other dispersers analyzed here have even higher degree.

For cubic functions we show (in Theorem 2.9) that having large bias on a certain subspace implies being constant on a slightly smaller subspace contained in it. Thus, assuming we have proved that a cubic function is a disperser for dimension d we can immediately deduce that it is an extractor for dimension $d + d'$. The bias in this case is bounded by $1/\mathrm{poly}(d')$ (recently Haramaty and Shpilka [2009] showed that the bias is in fact bounded by $\exp(-d'^{\epsilon})$ for some $\epsilon > 0$). We conjecture that for the special cubic function appearing in (1) the actual bound on bias should be $\exp(-O(d'))$ (see Conjecture 2.8).

The method by which we convert cubic affine dispersers to affine extractors differs significantly from the rest of this paper. It involves tools from additive combinatorics, most notably the method introduced by Bogdanov and Viola [2007] for approximating a biased function by a majority of its (randomly chosen) directional derivatives, and a so-called "energy increment" argument that is in the spirit of the proof of Roth's theorem over $\mathbb{F}_3^n$ due to Meshulam [1995] (see Section 5 for more details).

2 Main results

In this section, we state our main results. We start by formally defining affine dispersers and extractors.

Definition 2.1 ($\mathbb{F}_p$-affine dispersers and extractors) *A function $f : \mathbb{F}_p^m \to \mathbb{F}_p$ is an* **$\mathbb{F}_p$-affine disperser** *for dimension d if for every affine subspace $S \subseteq \mathbb{F}_p^m$ of dimension at least d, we have $|f(A)| > 1$.*

A function $f : \mathbb{F}_p^m \to \mathbb{F}_p$ is an **$\mathbb{F}_p$-affine ϵ-extractor** *for dimension d if for every affine subspace $S \subseteq \mathbb{F}_p^m$, if x is picked uniformly at random from S, the statistical distance of $f(x)$ from the uniform distribution on $\mathbb{F}_p$ is at most ϵ.*

We briefly indicate the relation between this definition and the more general setting. Following the derandomization literature, we will refer to a distribution over a domain $\mathcal{D}$ as a "source". A function $f : \mathcal{D} \to \mathcal{R}$ is said to be an **ϵ-extractor** for a set of sources $\mathcal{S}$ if, for every $S \in \mathcal{S}$, if x is picked according S, then the statistical distance of $f(x)$ from the uniform distribution on $\mathcal{R}$ is at most ϵ (ϵ is called the error-parameter of the extractor). The function f is a **disperser** for $\mathcal{S}$ if it is an ϵ-extractor for some $\epsilon < 1$. (This is equivalent to saying that f is nonconstant on the support of S for each source $S \in \mathcal{S}$).

A **d-dimensional affine source in $\mathbb{F}_p^m$** is the uniform distribution over some d-dimensional affine space. In this language, we see that a function $f : \mathbb{F}_p^m \to \mathbb{F}_p$ is an *$\mathbb{F}_p$-affine disperser (ϵ-extractor, respectively) for dimension d* if and only if it is a disperser (ϵ-extractor, respectively) for the set of d-dimensional affine sources in $\mathbb{F}_p^m$.

A more standard definition of a disperser, as appearing in, say, [Shaltiel, 2002], requires that for every d-dimensional affine source S, $f(\mathrm{supp}(S))$ equals the full range $\mathbb{F}_p$. Notice that for the case of $p = 2$ the two definitions match. All our constructions give $\mathbb{F}_p$-affine dispersers according to Definition 2.1. When p is clear from the context, we simply refer to them as *affine dispersers*.

2.1 Disperser for affine spaces of sublinear dimension

We begin by describing the function $f : \mathbb{F}_p^m \to \mathbb{F}_p$ which will prove to be a disperser over $\mathbb{F}_p$ for affine sources of sublinear dimension. The integers n, r and t are parameters of the construction to be specified later. As in [Bourgain, 2007], we partition the m coordinates of an input x into r blocks $(x_1, \ldots, x_r)$ of n coordinates each (we assume n divides m by discarding a few field-elements, if necessary). We will pick n to be prime, so that $\mathbb{F}_{p^n}$ has no nontrivial subfields. Each block x_i is interpreted as an element of $\mathbb{F}_{p^n}$ by using an $\mathbb{F}_p$-linear isomorphism from $\mathbb{F}_p^n$ to $\mathbb{F}_{p^n}$. We then raise each x_i to a suitable distinct power and let y_i denote the result of this powering. Next, we apply the t^{th} symmetric polynomial to $y_1, \ldots, y_r$, and get $z \in \mathbb{F}_{p^n}$, where this polynomial is defined by

$$\mathsf{Sym}_r^t(Y_1, \ldots, Y_r) = \sum_{I \subseteq [r], |I| = t} \prod_{i \in I} Y_i.$$

Finally, we take a nontrivial $\mathbb{F}_p$-linear map $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$, and output $\pi(z)$. We now formally state our main result.

Theorem 2.2 (Affine dispersers for sublinear dimension) *Given integer m fix parameters n, r, t as follows. Let n be the smallest prime bigger than $2 \cdot m^{3/5}$. Let $r = \lceil m/n \rceil$ and let $t = \lceil \sqrt{r} \rceil$. (We have $n \approx m^{3/5}, r \approx m^{2/5}$ and $t \approx m^{1/5}$.) Let $\phi : \mathbb{F}_p^m \to (\mathbb{F}_{p^n})^r$ be an injective $\mathbb{F}_p$-linear map,*

where $\phi(y) = (\phi_1(y), \ldots, \phi_r(y))$ *and* $\phi_i(y) \in \mathbb{F}_{p^n}$. *Let* $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ *be a nontrivial* $\mathbb{F}_p$*-linear map. Then the function* $f : \mathbb{F}_p^m \to \mathbb{F}_p$ *defined by*

$$f(x) = \pi\left(\mathsf{Sym}_r^t\left((\phi_1(x))^{1+p}, (\phi_2(x))^{1+p+p^2}, \ldots, (\phi_r(x))^{1+p+p^2+\ldots+p^r}\right)\right) \tag{4}$$

is an affine disperser for dimension greater than $6m^{4/5}$*, i.e., for all affine* $\mathcal{A} \subseteq \mathbb{F}_p^m$ *with* $\dim(\mathcal{A}) > 6m^{4/5}$ *we have* $|f(\mathcal{A})| > 1$.

Notice f can be computed in polynomial time in p and m because Sym_r^t can be computed efficiently in the said time (using the Newton-Girard identities). From a computational viewpoint our construction is more efficient than that of [Bourgain, 2007] which for spaces of dimension ϵm required a running time of $m^{2^{\Omega(1/\epsilon)}}$.

The method by which we prove Theorem 2.2 is quite general and in the following subsections we show that a few natural variants of it can also be shown to be good affine dispersers and extractors in various settings.

2.2 Disperser for independent affine sources

Informally, we say a function $f : (\mathbb{F}_p^n)^t \to \mathbb{F}_p$ is a *disperser for independent affine sources* if on every set of affine spaces $\mathcal{A}_1, \ldots, \mathcal{A}_t \subseteq \mathbb{F}_p^n$ of sufficiently large dimensions, we have $|f(\mathcal{A}_1 \times \cdots \times \mathcal{A}_t)| > 1$. The following theorem presents an affine disperser for independent sources that works as long as the sum of dimensions is greater than n. The analysis of this independent source affine disperser turns out to play a crucial role in our proof Theorem 2.2.

In what follows, a *proper subfield* of $\mathbb{F}_{p^n}$ is a subfield $\mathbb{K}$ of size $< p^n$ and an *affine shift* of $\mathbb{K}$ is a set of the form $\{a \cdot s + b \mid s \in \mathbb{K}\}$ for some fixed $a, b \in \mathbb{F}_{p^n}$. (Notice that every one-dimensional $\mathbb{F}_p$-affine subspace of $\mathbb{F}_{p^n}, n > 1$ is an affine shift of the proper subfield $\mathbb{F}_p$.)

Theorem 2.3 (Disperser for independent affine sources) *Let* $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ *be a nontrivial* $\mathbb{F}_p$*-linear map. Consider the function* $f : \mathbb{F}_{p^n}^t \to \mathbb{F}_p$ *given by*

$$f(x_1, \ldots, x_t) = \pi\left(\prod_{i=1}^{t} x_i^{1+p}\right). \tag{5}$$

Let $\mathcal{A}_1, \cdots, \mathcal{A}_t \subseteq \mathbb{F}_{p^n}$ *be* $\mathbb{F}_p$*-affine spaces of dimensions* $d_1, \ldots, d_t$ *respectively, where* $\sum_{i=1}^{t}(d_i - 2) > n$. *Suppose furthermore that no* $\mathcal{A}_i$ *is contained in an affine shift of a proper subfield of* $\mathbb{F}_{p^n}$. *Then* $|f(\mathcal{A}_1 \times \cdots \times \mathcal{A}_t)| > 1$.

Remark 2.4 *The assumption that* $\mathcal{A}_i$ *is not contained in an affine shift of a proper subfield is necessary. Without it we could set* $\mathcal{A}_i = \mathbb{K}$ *for a proper subfield* $\mathbb{K}$*, and select some nontrivial* π *such that the resulting function* f *is constant on* $\mathcal{A}_1 \times \ldots \times \mathcal{A}_t$.

Remark 2.5 *A result of Hou et al. [2002] implies the following statement (cf. DeVos and Gabizon [2009]). Let $\mathcal{A}_1, \ldots, \mathcal{A}_t \subseteq \mathbb{F}_{p^n}$ be affine spaces of dimensions $d_1, \ldots, d_t$ respectively and none are contained in affine shifts of proper subfields. Then*

$$\dim\left(\operatorname{span}\left\{\prod_i x_i \;\middle|\; x_i \in \mathcal{A}_i\right\}\right) \geq \min\left\{n, \sum_i (d_i - 1)\right\}.$$

So if $\sum(d_i - 1) \geq n$ then $\pi\left(\prod_{i=1}^t x_i\right)$ is nonconstant on $\mathcal{A}_1 \times \cdots \times \mathcal{A}_t$. The proof technique of Hou et al. [2002] differs significantly from ours and it is not clear how to derive one result from the other.

2.3 Univariate dispersers

Our next set of results is a pair of constructions based on univariate polynomials. We treat our input $x \in \mathbb{F}_p^n$ as a single element of the field $\mathbb{F}_{p^n}$ by using any $\mathbb{F}_p$-linear isomorphism between $\mathbb{F}_p^n$ and $\mathbb{F}_{p^n}$. We raise x to a suitable power and map the result to $\mathbb{F}_p$ using any nontrivial $\mathbb{F}_p$-linear map. The first construction will be shown in the next subsection to be an extractor for dimension greater than $2n/5$ and the second works for lower dimension $(n/3)$ but we cannot show that it is an extractor (cf. Conjecture 2.8). We call the next construction "cubic", and the one that follows "quartic", because the relevant functions f, when viewed as having domain $(\mathbb{F}_p)^n$, are computed by polynomials of degree 3 and 4 respectively (cf. the first bullet of Proposition 3.2).

Theorem 2.6 (Univariate cubic affine disperser) *Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear map. The function $f : \mathbb{F}_{p^n} \to \mathbb{F}_p$ given by*

$$f(x) = \pi\left(x^{1+p+p^2}\right)$$

is a disperser for the set of affine spaces of dimension greater than $\frac{2n}{5} + 10$ that are not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

In particular, if n is odd, then f is an affine disperser for dimension $\frac{2n}{5} + 10$.

Theorem 2.7 (Univariate quartic affine disperser) *Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear homomorphism. The function $f : \mathbb{F}_{p^n} \to \mathbb{F}_p$ given by*

$$f(x) = \pi\left(x^{1+p+p^2+p^3}\right)$$

is a disperser for the set of affine spaces of dimension greater than $\frac{n}{3} + 10$ that are not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

In particular, if n is odd, then f is an affine disperser for dimension $\frac{n}{3} + 10$.

We believe that the dimension bound in the above pair of theorems is not tight. In particular, we think the cubic construction of Theorem 2.6 should be a disperser for dimension $> n/3$ and the quartic construction of Theorem 2.7 should work for dimension $> n/4$. In fact, we believe in the stronger conjecture stated next.

Conjecture 2.8 (Univariate extractors) *For every prime p and integer k there exists an integer $c = c(p,k)$ and constant $\epsilon = \epsilon(p,k) > 0$ such that the following holds for all sufficiently large n. Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear map. Let $s_k = \sum_{i=0}^{k} p^i$. The function $f_k : \mathbb{F}_{p^n} \to \mathbb{F}_p$ given by*

$$f_k(x) = \pi\left(x^{s_k}\right)$$

is an $\exp(-\epsilon d)$-extractor for the set of affine spaces of dimension greater than $\left(\frac{n}{k} + c\right) + d$ that are not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

2.4 A cubic affine disperser is an affine extractor

Our final set of results shows that any *cubic* function that is a disperser for dimension d, is an $\epsilon(d')$-extractor for dimension $d + d'$, where $\epsilon(d')$ goes to 0 as d' increases.

Theorem 2.9 (Cubic affine dispersers are affine extractors) *There exists a universal constant $\epsilon > 0$ such that the following holds. Let $f : \mathbb{F}_p^m \to \mathbb{F}_p$ be computed by a cubic polynomial. If f is an affine disperser for dimension d_0 then f is an affine $O(d^{\epsilon})$-extractor for dimension $d_0 + \hat{d}$.*

Using the cubic construction from Theorem 2.6, the previous theorem implies the following affine extractor.

Corollary 2.10 (Univariate cubic affine extractor) *There exists a universal constant $\epsilon > 0$ such that the affine disperser f defined in Theorem 2.6 is an affine $O(d^{-\epsilon})$-extractor for dimension $\left(\frac{2n}{5} + 10\right) + d$.*

Remark 2.11 *Recent work of Haramaty and Shpilka [2009] gives a better bound on the error-parameter of f stated in 2.9. They show a bound of $\exp(-d^{\epsilon})$ on the error-parameter for some universal constant $\epsilon > 0$.*

Unfortunately, *quartic* affine dispersers are not necessarily affine extractors for comparable dimension (see Section 5.3 for a counter-example). So, although we believe the quartic construction of Theorem 2.7 is an affine extractor (cf. Conjecture 2.8), a proof of this conjecture will have to rely on the particular algebraic structure of this quartic function.

Returning to *cubic* affine extractors, counting arguments show that there exist cubic functions that are dispersers for affine spaces of dimension as small as $\sqrt{n}$ (see Lemma 5.10). Given Theorem 2.9, this implies that one way to get affine extractors for sublinear dimension is to find an explicit cubic affine disperser that works for the same dimension bound.

Organization of the rest of the paper The next section gives a brief introduction to the theory of subspace polynomials and establishes the properties we need for our analysis. Of particular importance are *(i)* the Main Structural Lemma 3.10 which connects the fact that a subspace is not an affine shift of a subfield to the zero-nonzero pattern of the coefficients of its corresponding

subspace polynomial, and *(ii)* Lemma 3.13 which is used to show that our constructions, when restricted to a subspace of sufficiently large dimension, are polynomials of positive degree.

The proofs of our main results go in increasing order of complexity. In Section 4 we discuss our univariate constructions, proving Theorems 2.6 and 2.7. In Section 5 we prove that cubic affine dispersers are extractors. In Section 6 we analyze the disperser for independent sources and prove Theorem 2.3. In Section 7 we analyze our construction for sublinear dimension and prove Theorem 2.2. Together, Sections 3, 6, and 7 contain a complete proof of Theorem 2.2.

3 Properties of subspace polynomials

In this section we build up some preliminaries on polynomials and subspace polynomials. This section is organized as follows. In Section 3.1 we discuss basic properties of polynomials, and the relationship between polynomials over $\mathbb{F}_{p^n}$ and polynomials over $\mathbb{F}_p^n$. In Section 3.2 we define linearized and subspace polynomials and mention some of their properties. We follow this in Section 3.3 with the key structural Lemma 3.10 regarding coefficients of subspace polynomials. Finally, in Section 3.4 we state and prove Lemma 3.13 which discusses the coefficient structure of products of linearized polynomials and will be employed in all subsequent proofs of our affine disperser results.

3.1 Preliminaries

Throughout this paper capital letters such as X_i are used for formal variables, and small letters such as x_i are used for field-elements. For a polynomial $h(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$, abusing notation we define

$$h(X_1, \ldots, X_r) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle$$

to be the unique polynomial congruent to $h(X_1, \ldots, X_n) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle$ of degree $< p^n$ in each variable. Equivalently, h' is the polynomial obtained by starting with h and repeatedly replacing, for each i, every occurrence of $X_i^{p^n}$ by X_i. The following proposition, stated without proof, will be used repeatedly in our arguments.

Proposition 3.1 *Let $h(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$. Let*

$$h'(X_1, \ldots, X_r) = h(X_1, \ldots, X_r) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle.$$

Then for any $x \in \mathbb{F}_{p^n}^r$ we have $h(x) = h'(x)$.

Consequently, $|h(\mathbb{F}_{p^n}^r)| > 1$ if and only if $h'(X_1, \ldots, X_r)$ is a polynomial of degree greater than 0.

For a nonnegative integer i, let $\mathrm{wt}_p(i)$ denote the sum of the digits of i in the base-p representation. If $m(X_1, \ldots, X_t) \in \mathbb{F}_{p^n}[X_1, \ldots, X_t]$ is the monomial $\prod_{i=1}^t X_i^{\beta_i}$, we define the $\mathbb{F}_p$-*degree* of m in the variable X_i to be $\mathrm{wt}_p(\beta_i)$. We define the $\mathbb{F}_p$-*degree* of the monomial $\mathcal{M}$ to be the sum of the $\mathbb{F}_p$-degrees of $\mathcal{M}$ in each variable X_i. We then define the $\mathbb{F}_p$-degree of a polynomial to be the maximum $\mathbb{F}_p$-degree of any of its monomials.

Proposition 3.2 *Let $P(X_1, \ldots, X_t), Q(X_1, \ldots, X_t)$ be polynomials in $\mathbb{F}_{p^n}[X_1, \ldots, X_t]$ with $\mathbb{F}_p$-degrees $d_1, d_2 < n$ respectively. Let $\phi = (\phi_1, \ldots, \phi_t) : \mathbb{F}_p^{n \cdot t} \to \mathbb{F}_{p^n}^t$ be an $\mathbb{F}_p$-linear isomorphism. and $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be an $\mathbb{F}_p$-linear map. Then*

- *Let $f = (f_1, \ldots, f_n) : \mathbb{F}_p^{n \cdot t} \to \mathbb{F}_p$ be given by $f(x) = \pi\left(P(\phi_1(x), \ldots, \phi_t(x))\right)$. Then f is computed by a polynomial $P' \in \mathbb{F}_p[Y_1, \ldots, Y_{n \cdot t}]$ of total degree at most d_1.*
- *The $\mathbb{F}_p$-degree of $P(X_1, \ldots, X_t) \cdot Q(X_1, \ldots, X_t)$ is at most $d_1 + d_2$.*
- *The $\mathbb{F}_p$-degree of $P(X_1, \ldots, X_t)^{p^r} \mod \langle (X_i^{p^n} - X_i)_{i \in [t]} \rangle$ equals d_1.*

We recall one final basic fact about finite field extensions — that $\mathbb{F}_p$-linear maps from $\mathbb{F}_{p^n}$ to $\mathbb{F}_p$ are computed by trace maps (cf. Lidl and Niederreiter [1997]).

Proposition 3.3 *Let $\mathrm{Tr}(Y) = \sum_{i=0}^{n-1} Y^{p^i}$ be the* trace map *from $\mathbb{F}_{p^n}$ to $\mathbb{F}_p$. For every $\mathbb{F}_p$-linear map $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ there exists $\mu = \mu_\pi \in \mathbb{F}_{p^n}$ such that for all $x \in \mathbb{F}_{p^n}$ we have*

$$\pi(x) = \mathrm{Tr}(\mu \cdot x).$$

Furthermore, π is trivial if and only if $\mu = 0$.

3.2 Introduction to the theory of subspace polynomials

The information in this subsection was first described in the work of Ore [1933, 1934]. We state the minimal set of definitions and claims that will be needed to analyze our constructions and we refer the reader interested in a more thorough introduction to the subject to [Lidl and Niederreiter, 1997, Chapter 4] and to [Berlekamp, 1968, Chapter 11].

A polynomial $P \in \mathbb{F}_{p^n}[X]$ is said to be $\mathbb{F}_p$-*linearized* if it is of the form:

$$P(X) = \sum_{i=0}^{n-1} a_i X^{p^i}, a_i \in \mathbb{F}_{p^n}$$

(when p is clear from context, we will simply refer to them as linearized polynomials). P being linearized is equivalent to having $P(\beta b + \gamma c) = \beta P(b) + \gamma P(c)$ for all $b, c \in \mathbb{F}_{p^n}$ and $\beta, \gamma \in \mathbb{F}_p$. By extension, a polynomial is said to be *affine linearized* if $P(X) = \hat{P}(X) + \hat{a}$ where $\hat{P}$ is linearized and $\hat{a} \in \mathbb{F}_{p^n}$. The affine linearized polynomials over $\mathbb{F}_{p^n}$ are precisely the polynomials of $\mathbb{F}_p$-degree at most 1.

Lemma 3.4 *Let $\phi : \mathbb{F}_p^n \to \mathbb{F}_{p^n}$ be an $\mathbb{F}_p$-linear isomorphism. There is a one-to-one correspondence between affine transformations from $\mathbb{F}_p^n$ to $\mathbb{F}_p^n$ and affine linearized polynomials in $\mathbb{F}_{p^n}[X]$, i.e., for every affine transformation $T : \mathbb{F}_p^n \to \mathbb{F}_p^n$ there exists a unique affine linearized polynomial P_T satisfying $P_T(\phi(b)) = T(\phi(b))$ for all $b \in \mathbb{F}_p^n$.*

We shall take particular interest in a special class of linearized polynomials that split completely in $\mathbb{F}_{p^n}$ to a set of roots that forms a $\mathbb{F}_p$-affine subspace of $\mathbb{F}_{p^n}$.

Definition 3.5 (Kernel-subspace polynomial) *Let $L \subseteq \mathbb{F}_{p^n}$ be an affine subspace of dimension d. Define $P_L(X) \in \mathbb{F}_{p^n}[X]$, the* kernel-subspace polynomial *of L, to be*

$$P_L(X) = \prod_{\alpha \in L} (X - \alpha).$$

Lemma 3.6 (Kernel-subspace polynomials are affine) *If $L \subseteq \mathbb{F}_{p^n}$ is an affine subspace of dimension d then $P_L(X)$ is a monic affine linearized polynomial of degree p^d. Furthermore, P_L is linearized iff L is a linear space.*

Every kernel-subspace polynomial P_L corresponds to an affine transformation whose kernel is L, so by linearity $P(\mathbb{F}_{p^n})$ is an affine subspace of $\mathbb{F}_{p^n}$ of dimension $n - \dim(L)$. Surprisingly, the images of all d-dimensional subspace polynomials are precisely all the $n - d$ dimensional subspaces of $\mathbb{F}_{p^n}$. These *image-subspace* polynomials will be the starting point of our analysis of affine dispersers.

Lemma 3.7 (Existence of an image-subspace polynomial) *If $L \subseteq \mathbb{F}_{p^n}$ is an affine subspace of dimension d then there exists a monic affine linearized polynomial $Q_L(X)$ with $\deg(Q_L) = p^{n-d}$, called the* image-subspace polynomial *of L, such that*

$$L = Q_L(\mathbb{F}_{p^n}) \triangleq \{Q_L(c) \mid c \in \mathbb{F}_{p^n}\}.$$

Moreover, if $P_L(X)$ is the subspace polynomial of L then

$$P_L(Q_L(X)) \equiv Q_L(P_L(X)) \equiv X^{p^n} - X. \tag{6}$$

Thus the kernel of $Q_L : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ is the image of $P_L : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$. In particular $Q_L(X)$ has p^{n-d} roots in $\mathbb{F}_{p^n}$, and is thus also a kernel subspace polynomial of some $(n-d)$-dimensional subspace.

For the sake of completeness we include the (short) proof of this lemma from Berlekamp [1968].

Proof Let $L' = P_L(\mathbb{F}_{p^n})$ be the image of $P_L(X)$. Define $Q_L(X)$ to be $P_{L'}(X)$, the kernel-subspace polynomial of L'.

Notice that $Q_L(P_L(X))$ is a monic polynomial of degree p^n that vanishes on $\mathbb{F}_{p^n}$, hence $Q_L(P_L(X)) = X^{p^n} - X$. Thus $P_L(Q_L(P_L(X)))) = P_L(X^{p^n} - X) = P_L(X^{p^n}) - P_L(X) = P_L(X)^{p^n} - P_L(X)$. Letting $g(Y)$ be the polynomial $P_L(Q_L(Y)) - (Y^{p^n} - Y)$, we have just proved that $g(P_L(X)) = 0$. This implies $g(Y) = 0$, since $\deg(g(P_L(X))) = (\deg g(Y)) \cdot (\deg(P_L(X)))$.

So $P_L(Q_L(y)) = 0$ for each $y \in \mathbb{F}_{p^n}$. In particular, we see that the image of Q_L is contained in L, and by dimension counting, the image of Q_L equals L. □

3.3 Coefficients of subspace polynomials

The following claims will be needed to prove our main structural lemma. In what follows, let $\overline{\mathbb{F}}_p$ denote the algebraic closure of $\mathbb{F}_p$.

Claim 3.8 *Let $k > 1$, and suppose $a, c \in \mathbb{F}_{p^n}$ are such that $a^{p^k} - ca = 0$. Then, letting b be any $(p^k - 1)$-th root of c in $\overline{\mathbb{F}}_p$, we have $a \in b \cdot \mathbb{F}_{p^k}$.*

Proof If $a = 0$ then the claim is trivial. Otherwise, we have $a^{p^k} = ca$, and hence $a^{p^k-1} = c$. Thus $(a/b)^{p^k-1} = 1$, which implies that $a/b \in \mathbb{F}_{p^k}$. □

Claim 3.9 *For linearized polynomial $Q(X) = \sum_{j=0}^{n-1} a_j X^{p^j} + \hat{a} \in \mathbb{F}_{p^n}[X]$ and integer t, we have*

$$(Q(X))^{p^t} \pmod{X^{p^n} - X} \equiv \sum_{j=0}^{n-1} \left(a_{((j-t) \mod n)}\right)^{p^t} X^{p^j} + \hat{a}^{p^t}.$$

The proof follows by direct expansion, using the $\mathbb{F}_p$-linearity of the map $Z \longmapsto Z^{p^t}$.

We now state and prove our main structural lemma about the zero/nonzero pattern of consecutive coefficients of subspace polynomials.

Lemma 3.10 (Main structural lemma for subspace polynomials) *Let L be a d-dimensional linear subspace in $\mathbb{F}_{p^n}$. Let $Q_L(X) = \sum_{j=0}^{n-1} a_j X^{p^j}$ be the image-subspace polynomial of L.*

1. *For any integer r and set $J = \{(r+j) \mod n \mid j = 0, \ldots, d-1\}$ of d consecutive indices in $\mathbb{Z}_n$, there is some $j \in J$ with $a_j \neq 0$. In particular, a_0 and a_{n-d} are nonzero.*
2. *Suppose that L is not contained in any constant multiple of a proper subfield of $\mathbb{F}_{p^n}$, i.e. $L \not\subseteq \beta \cdot \mathbb{F}_{p^k}$ for any $\beta \in \mathbb{F}_{p^n}$ and any $\mathbb{F}_{p^k} \subset \mathbb{F}_{p^n}$. Then for any integer $r \neq n-d+1$ and set $J = \{(r+j) \mod n \mid j = 0, \ldots, d-2\}$ of $d-1$ consecutive indices in $\mathbb{Z}_n$, there is some $j \in J$ with $a_j \neq 0$.*

Proof For the first part, suppose $a_j = 0$ for all $j \in J$. Note that by Lemma 3.7, Q_L has p^{n-d} distinct roots in $\mathbb{F}_{p^n}$. Let $Q'(X) := Q_L(X)^{p^{n-(r+d)}} \mod X^{p^n} - X$. Then, by Claim 3.9 we conclude $Q'(X) = \sum_{j=0}^{n-1} a_{j+r+d-n}^{p^{n-(r+d)}} X^{p^j}$. Now for any $j \in [n-d, n-1]$, we have $a_{j+r+d-n} = 0$ by assumption, and thus $Q'(X)$ is of degree at most p^{n-d-1}. In addition, by Proposition 3.1, $Q'(\alpha) = Q_L(\alpha)^{p^{n-(r+d)}} = 0$ for every $\alpha \in \mathbb{F}_{p^n}$ satisfying $Q_L(\alpha) = 0$, and hence Q' has at least p^{n-d} roots. This is a contradiction.

In particular, since by definition $a_{n-d+1}, \ldots, a_{n-1}$ forms a sequence of $d-1$ consecutive coefficients that are all zero, we conclude both a_{n-d} and a_0 must be nonzero.

For the second part, suppose $a_j = 0$ for all $j \in J$. Again, by Lemma 3.7, Q_L has p^{n-d} distinct roots in $\mathbb{F}_{p^n}$. Let $k = n - (r+d) + 1$ (note that $0 < k < n$). Then as above the polynomial $Q'(X) := Q_L(X)^{p^k} \mod X^{p^n} - X$ is nonzero of degree at most p^{n-d}. In addition, $Q'(\alpha) = Q_L(\alpha)^{p^k} = 0$ for every $\alpha \in \mathbb{F}_{p^n}$ for which $Q_L(\alpha) = 0$. As Q' and Q_L are of the same degree p^{n-d}, there is a constant $c \in \mathbb{F}_{p^n}$ such that $Q'(X) - cQ_L(X)$ is of degree at most p^{n-d-1} and vanishes on the p^{n-d} roots of $Q_L(X)$. Thus the polynomial $Q'(X) - cQ_L(X)$ is identically zero. Recalling the definition of $Q'(X)$, have just showed that $Q_L(X)^{p^k} - cQ_L(X) = 0 \mod X^{p^n} - X$. Thus for each $\alpha \in \mathbb{F}_{p^n}$, we have $Q_L(\alpha)^{p^k} - cQ_L(\alpha) = 0$. Now, since the image of Q_L is L, by Claim 3.8 we conclude that $L \subseteq b \cdot \mathbb{F}_{p^k}$ (where $b \in \overline{\mathbb{F}}_p$ is a $p^k - 1$-th root of c). This almost gives the desired contradiction, but for the possibility that $b \notin \mathbb{F}_{p^n}$, and that $\mathbb{F}_{p^k}$ may not be a subfield of $\mathbb{F}_{p^n}$.

Let $\beta \in L \setminus \{0\}$. For any $\alpha \in L$, we have $\alpha/\beta \in (b \cdot \mathbb{F}_{p^k})/(b \cdot \mathbb{F}_{p^k})$, and hence $\alpha/\beta \in \mathbb{F}_{p^k}$. Thus $\beta^{-1} \cdot L \subseteq \mathbb{F}_{p^k} \cap \mathbb{F}_{p^n} = \mathbb{F}_{p^{(k,n)}}$, where $(k,n) = \gcd(k,n)$. Thus $L \subseteq \beta \cdot \mathbb{F}_{p^{(k,n)}}$, contradicting the hypothesis on L. □

3.4 Coefficients of products of subspace polynomials

In our subsequent arguments, we will need time and again to prove that a certain polynomial P, which is the trace of products of linearized polynomials reduced $\mod \langle (X_i^{p^n} - X_i)_{i\in[r]}\rangle$, is not a constant. In this subsection we describe a lemma that will allow us to argue such statements by showing that a well-chosen monomial of P has a nonzero coefficient. We start with a definition.

Definition 3.11 (Associated matrix and its zero-one indicator matrix) *For a linearized polynomial $Q(X) = \sum_{i=0}^{n-1} a_i X^{p^i}$ over $\mathbb{F}_{p^n}$, we define its* associated matrix $M_Q \in \mathbb{F}_{p^n}^{\{0,\ldots,n-1\}\times\{0,\ldots,n-1\}}$ *by setting the (i,j)-entry of M_Q to be $(a_{j-i})^{p^i}$, where both rows and columns are indexed by $\{0,1,\ldots,n-1\}$ and index arithmetic, as well as powers of p are computed modulo n. Explicitly, M_Q is the following matrix*

$$\begin{pmatrix} a_0 & a_1 & a_2 & \ldots & \ldots & a_{n-1} \\ (a_{n-1})^p & (a_0)^p & (a_1)^p & \ldots & \ldots & (a_{n-2})^p \\ (a_{n-2})^{p^2} & (a_{n-1})^{p^2} & (a_0)^{p^2} & \ldots & \ldots & (a_{n-3})^{p^2} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ (a_1)^{p^{n-1}} & (a_2)^{p^{n-1}} & (a_3)^{p^{n-1}} & \ldots & \ldots & (a_0)^{p^{n-1}} \end{pmatrix}.$$

For $a_i \in \mathbb{F}_{p^n}$ let a_i' indicate whether a_i is zero, i.e., $a_i' = 0$ if $a_i = 0$ and otherwise $a_i' = 1$. Similarly, let $M' = M_Q'$ denote the zero-one indicator matrix of M_Q. The (i,j)-entry of this matrix is a_{j-i}', or, in other words, the (i,j)-entry of M' indicates whether the (i,j)-entry of M is nonzero.

The use of the associated matrix is captured by the following claim. The proof of the claim (which is omitted) follows immediately from Claim 3.9.

Claim 3.12 *The (i,j)-entry of M_Q is the coefficient of X^{p^j} in the linearized polynomial $(Q(X))^{p^i} \mod X^{p^n} - X$.*

To state the main lemma of this subsection we need the following notation. For A, B nonempty subsets of $\{0,\ldots,n-1\}$ let $M[A,B]$ be the minor corresponding to rows A and columns B. For an integer r, let $B + r = \{s + r \mod n \mid s \in B\}$.

Lemma 3.13 *Let $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$. Let $A_1,\ldots,A_t, B_1,\ldots,B_t \subseteq \{0,\ldots,n-1\}$ satisfy $|A_i| = |B_i| > 0$ for $i = 1,\ldots,t$. Let $\alpha_i = \sum_{j\in A_i} p^j, \beta_i = \sum_{k\in B_i} p^k$. Let $Q_1(X_1),\ldots,Q_t(X_t)$ be linearized polynomials with associated matrices $M_1,\ldots,M_t$ and zero-one indicator matrices $M_1',\ldots,M_t'$ respectively.*

The coefficient $c_{\mathcal{M}}$ of the monomial $\mathcal{M} = \prod_{i=1}^t X_i^{\beta_i}$ in

$$R(X_1,\ldots,X_t) = \mathrm{Tr}\left(\mu \prod_{i=1}^{t} (Q_i(X_i))^{\alpha_i}\right) \mod \langle (X_i^{p^n} - X_i)_{i\in[t]}\rangle \tag{7}$$

is given by the expression

$$c_{\mathcal{M}} = \sum_{r=0}^{n-1} \mu^{p^r} \cdot \prod_{i=1}^{t} \mathrm{Perm}\left(M_i[A_i + r, B_i]\right) = \sum_{r=0}^{n-1} \mu^{p^r} \cdot \prod_{i=1}^{t} \mathrm{Perm}\left(M_i[A_i, B_i - r]\right)^{p^r}. \tag{8}$$

Proof Notice

$$\mathrm{Tr}\left(\mu \prod_{i=1}^{t} (Q_i(X_i))^{\alpha_i}\right) = \sum_{r=0}^{n-1} \mu^{p^r} \cdot \prod_{i=1}^{t} (Q_i(X_i))^{\alpha_i \cdot p^r}.$$

Thus, $c_{\mathcal{M}}$ is a sum of n elements, where the rth element, denoted $c_{\mathcal{M}}^{(r)}$, is the coefficient of m in the rth summand in the right hand side above. We can break $c_{\mathcal{M}}^{(r)}$ further into μ^{p^r} times a product of t terms, where the ith term is the coefficient of $X_i^{\beta_i}$ in $(Q_i(X_i))^{\alpha_i \cdot p^r}$. So to prove the lemma it suffices to show that the coefficient of $X_i^{\beta_i}$ in $(Q_i(X_i))^{\alpha_i \cdot p^r}$ is $\mathrm{Perm}\,(M_i[A_i + r, B_i])$.

Expand $(Q_i(X_i))^{\alpha_i \cdot p^r}$ as

$$\prod_{j \in A_i} (Q_i(X_i))^{p^{j+r}} = \prod_{j \in A_i + r} (Q_i(X_i))^{p^j}.$$

By assumption $|A_i| = |B_i|$ and expanding $X_i^{\beta_i}$ as $\prod_{k \in B_i} X_i^{p^k}$ we see that for every one-to-one mapping $h : B_i \to A_i$ we get a contribution to the coefficient of $X_i^{\beta_i}$ by picking $X_i^{p^k}$ from $(Q_i(X_i))^{p^{h(k)+r}}$, i.e., the coefficient of $X_i^{\beta_i}$ is (using Claim 3.12):

$$\sum_{h: B_i \to A_i, h \text{ one-to-one}} \prod_{k \in B_i} a_{i,k-(h(k)+r)}^{p^{h(k)+r}} = \mathrm{Perm}(M_i[A_i + r, B_i]).$$

This completes the proof of the lemma. □

The above lemma gives us an explicit formula for the coefficients of a certain polynomials. The following remark describes the exact way in which this lemma gets used to show that such a polynomial is nonzero.

Remark 3.14 *Keep the notation of the previous lemma. Suppose that the following two conditions hold:*

1. *$M_1'[A_1, B_1], \ldots, M_t'[A_t, B_t]$ are each, up to reordering of rows and columns, upper triangular with every diagonal entry nonzero.*
2. *For every $r \in \{1, \ldots, n-1\}$ there exists $i_r \in \{1, \ldots, t\}$ such that $M_{i_r}'[A_{i_r}, B_{i_r} - r]$ contains an all-zero column.*

Then the coefficient $c_{\mathcal{M}}$ of the monomial $\mathcal{M}$ in $R(X_1, \ldots, X_t)$ is nonzero.

Indeed, assumption 1 implies that the first summand on the right hand side of (8) is nonzero, because it is a product of permanents of upper triangular matrices with nonzero diagonal. Assumption 2 implies that all other summands are zero, because one matrix in the product has a zero permanent on account of its all-zero column.

4 Univariate constructions

In this section we prove our results about univariate dispersers. We start with the cubic affine disperser (in the next section, we will show that it is even an affine extractor).

4.1 Cubic affine disperser

Theorem 2.6 (Univariate cubic affine disperser, restated) *Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear map. The function $f : \mathbb{F}_{p^n} \to \mathbb{F}_p$ given by*

$$f(x) = \pi\left(x^{1+p+p^2}\right)$$

is a disperser for the set of affine spaces of dimension greater than $\frac{2n}{5} + 10$ that are not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

In particular, if n is odd, then f is an affine disperser for dimension $\frac{2n}{5} + 10$.

Proof We assume without loss of generality that $\dim(\mathcal{A}) = d = \lceil \frac{2n}{5} \rceil + 10$ (by replacing $\mathcal{A}$ with an arbitrary subspace of $\mathcal{A}$ of this dimension). By Proposition 3.3, we know that $\pi(x)$ is of the form $\mathrm{Tr}(\mu x)$ for some $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$. Let $Q(X)$ be the image-subspace polynomial of $\mathcal{A}$, so that $\mathcal{A} = Q(\mathbb{F}_{p^n})$. Let

$$R(X) = \mathrm{Tr}(\mu \cdot Q(X)^{1+p+p^2}) \mod \langle X^{p^n} - X \rangle,$$

so that by Proposition 3.1, $R(x) = f(Q(x))$ for each $x \in \mathbb{F}_{p^n}$ and hence $R(\mathbb{F}_{p^n}) = f(\mathcal{A})$. The same proposition implies that to prove Theorem 2.7, it suffices to show that $R(X)$ has a monomial of positive degree, and this is what we shall do.

To find the desired monomial we start by invoking Lemma 3.13. Applying this lemma to our case we have $t = 1$ and we get a single linearized polynomial $Q_1(X_1) = Q(X)$. The set $A = A_i$ is $\{0, 1, 2\}$, which corresponds to the exponent $\alpha = \alpha_1 = p^0 + p^1 + p^2$. Thus, Lemma 3.13 reads in our case as follows.

Claim 4.1 *For $B = \{i, j, k\} \subseteq \{0, \ldots, n-1\}$ let $\beta = p^i + p^j + p^k$. The coefficient $c_{\mathcal{M}}$ of the monomial $\mathcal{M} = X^\beta$ in*

$$\mathrm{Tr}\left(\mu \cdot Q(X)^{p^0+p^1+p^2}\right) \pmod{X^{p^n} - X}$$

is given by

$$c_{\mathcal{M}} = \sum_{r=0}^{n-1} \mu^r \mathrm{Perm}\left(M[A, B-r]\right)^{p^r}. \tag{9}$$

By Remark 3.14, the above claim implies that in order to show that $R(X)$ is nonconstant, letting $M' = M'_Q$ be the zero-one indicator matrix of M_Q as defined in Definition 3.11, it suffices to find a $B \subseteq \{0, \ldots, n-1\}$ with $|B| = 3$, such that:

1. The matrix $M'[\{0, 1, 2\}, B]$ is, up to reordering of rows and columns, upper triangular with each diagonal entry nonzero.
2. For every $r \in \{1, \ldots, n-1\}$ the matrix $M'[\{0, 1, 2\}, B - r]$ contains an all-zero column.

We proceed to find such a B. Thus all the action is in the first 3 rows of the matrix M'.

To this end, we state a few useful properties of the coefficients of Q that all follow immediately from Lemma 3.10 and will be used later on in the proof. Notice that *(iv)* below follows via the second part of Lemma 3.10 from our assumption that $\mathcal{A}$ is not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

Claim 4.2 *Let $Q(X) = \sum_{i=0}^{n-1} a_i X^{p^i} + \hat{a}$ be the image-subspace polynomial of $\mathcal{A}$. Letting $d = \dim(\mathcal{A})$ we have (i) $d \geq \frac{2n}{5} + 10$, (ii) $a_0, a_{n-d} \neq 0$, (iii) $a_{n-d+1} = \ldots = a_{n-1} = 0$ and (iv) for every $0 \leq j \leq n-d$ there is at least one nonzero coefficient amongst $a_j, a_{j+1}, \ldots, a_{j+d-2}$.*

To further simplify notation, for $r_1 < r_2$ let $[r_1, r_2]$ denote the set of integers in the interval $[r_1, r_2]$. Let $I_0 = \{i \in [0, n-1] : a_i = 0\}$ denote the set of indices of the zero coefficients of Q and let $I_1 = [0, n-1] \setminus I_0$ be the set of indices of nonzero ones.

We show the existence of a set B satisfying properties 1 and 2 and break the proof into three cases according to the structure of I_0, I_1.

Case I — $I_1 \cap [n/5 - 15, 2n/5 + 7] \neq \emptyset$**:** Let $j \in I_1 \cap [n/5 - 15, 2n/5 + 7]$. We claim the set $B = \{0, j+1, n-d+2\}$ satisfies our pair of properties. Property 1 holds because

$$M'[\{0,1,2\}, \{0, j+1, n-d+2\}] = \begin{pmatrix} a'_0 & a'_{j+1} & a'_{n-d+2} \\ a'_{n-1} & a'_j & a'_{n-d+1} \\ a'_{n-2} & a'_{j-1} & a'_{n-d} \end{pmatrix} = \begin{pmatrix} a'_0 & a'_{j+1} & 0 \\ 0 & a'_j & 0 \\ 0 & a'_{j-1} & a'_{n-d} \end{pmatrix} = \begin{pmatrix} 1 & * & 0 \\ 0 & 1 & 0 \\ 0 & * & 1 \end{pmatrix}.$$

The second equality holds because of Claim 4.2 *(i)* , *(ii)* .

We now argue property 2. We have

$$M'[A, B-r] = M'[\{0,1,2\}, \{n-r, j+1-r, n-d+2-r\}] = \begin{pmatrix} a'_{n-r} & a'_{j+1-r} & a'_{n-d+2-r} \\ a'_{n-r-1} & a'_{j-r} & a'_{n-d+1-r} \\ a'_{n-r-2} & a'_{j-1-r} & a'_{n-d-r} \end{pmatrix}$$

For $r \in [1, d-3]$ the first column of $M'[A, B-r]$ is seen to be zero because the set of indices appearing there is

$$\{n-r, n-r-1, n-r-2\} \subseteq [n-d+1, n-1] \subseteq I_0,$$

(the last inclusion follows from Claim 4.2 *(ii)*). Similarly, for $r \in [n-(d-3), n-1]$ the last column of $M'[A, B-r]$ is zero, since the set of indices appearing there,

$$\{n-d+2-r, n-d+1-r, n-d-r\} \subseteq [n-d+1, n-1] \subseteq I_0.$$

Finally, for the remaining $r \in [d-2, n-(d-2)] \subseteq [2n/5 + 8, 3n/5 - 8]$, using the fact that $j \in [n/5 - 15, 2n/5 + 7]$, we see that the middle column of $M'[A, B-r]$ is zero, since the set of indices appearing there,

$$\begin{aligned} \{j+1-r, j-r, j-1-r\} &\subseteq \left[\left(\frac{n}{5} - 15\right) - 1 - \left(\frac{3n}{5} - 8\right), \left(\frac{2n}{5} + 7\right) + 1 - \left(\frac{2n}{5} + 8\right)\right] \\ &\subseteq \left[\frac{3n}{5} - 8, n-1\right] \subseteq I_0, \end{aligned}$$

where the last inclusion uses the bound on d which implies $3n/5 - 8 > n - d$. We conclude that property 2 also holds and the proof of the first case is complete.

Case II — $I_1 \cap [n/5-15, 2n/5+7] = \emptyset$**:** Let j_1 be the largest element in $[0, n/5-15] \cap I_1$ and let j_2 be the minimal element in $[2n/5+7, n-d] \cap I_1$. By Claim 4.2 *(iii)* we cannot have both $j_1 = 0$ and $j_2 = n-d$. Consider the following four intervals: $[0, j_1]$ (whose end points are in I_1), $[j_1+1, j_2-1]$ (which is contained int I_0), $[j_2, n-d]$ (whose end points are in I_1), and $[n-d+1, n-1]$ (which is contained in I_0). Denote the length of these intervals by $\alpha_1, \ldots, \alpha_4$ respectively. Notice the length of each of the zero intervals (α_2, α_4) is strictly greater than the length of the other two "nonzero" intervals. Moreover, by the assumption that $\mathcal{A}$ is not contained in an affine shift of a proper subfield, part 2 of Lemma 3.10 implies that $\alpha_4 > \alpha_2$. By assumption $\alpha_1, \alpha_3 < n/5-10$. We summarize this for future reference by

$$d-2 = \alpha_4 > \alpha_2 > \max\{\alpha_1, \alpha_3\} + 10. \tag{10}$$

There are two subcases,

Case II.a — $\alpha_1 \neq \alpha_3$**:** Assume without loss of generality $\alpha_1 > \alpha_3$. We claim that $B = \{0, n-d+1, j_1+2\}$ satisfies our pair of properties. (The case of $\alpha_1 < \alpha_3$ can be seen to be identical by using the argument below to show $B = \{j_2, j_1+1, n-d+2\}$ satisfies the said pair of properties.) Property 1 holds because

$$M'[\{0,1,2\}, \{0, n-d+1, j_1+2\}] = \begin{pmatrix} a'_0 & a'_{n-d+1} & a'_{j_1+2} \\ a'_{n-1} & a'_{n-d} & a'_{j_1+1} \\ a'_{n-2} & a'_{n-d-1} & a'_{j_1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & * & 1 \end{pmatrix}.$$

Regarding property 2, the key observation is that any nonzero shift will force either the first or the last column to be all zero. Indeed, the difference between the top left and bottom right indices of $M'[A, B-r = \{n-r, n-d+1-r, j_1+2-r\}]$ is $j_1 = \alpha_1 - 1$, and the difference between any other pair of indices chosen one from each of the first and third columns is between j_1 and j_1+4. Thus, by (10) the only value of r such that both these columns are not entirely zero is $r = 0$. We conclude property 2 holds and the proof of this case is complete.

Case II.b — $\alpha_1 = \alpha_3$**:** In this case we claim that $B = \{0, j_2+1, j_1+2\}$ satisfies our pair of properties. Property 1 can be verified by inspection as in the previous two cases. Furthermore, since the first and last column in this case are identical to the first and last column in the previous case, the same argument as there shows that the only nonzero shift that has both these columns nonzero must be $r = n - j_2$. We get the following matrix

$$M'[\{0,1,2\} + (n-j_2), B = \{0, j_2+1, j_1+2\}] = \begin{pmatrix} a'_{j_2} & a'_{2j_2+1} & a'_{n-d+2} \\ a'_{j_2-1} & a'_{2j_2} & a'_{n-d+1} \\ a'_{j_2-2} & a'_{2j_2-1} & a'_{n-d} \end{pmatrix} = \begin{pmatrix} 1 & * & 0 \\ 0 & a'_{2j_2} & 0 \\ 0 & * & 1 \end{pmatrix}.$$

The only way the matrix above can be nonzero is if $a'_{2j_2} \neq 0$. Since $j_2 > 2n/5$ the only way this can happen is to have $j_2 \geq n/2$. But in this case we get $\alpha_1 + \alpha_2 \geq \alpha_3 + \alpha_4$ which contradicts (10). We conclude the above matrix has permanent 0 and property 2 holds. This completes the proof for the final case and Theorem 2.6 follows. □

4.2 Quartic affine disperser

Theorem 2.7 (Univariate quartic affine disperser, restated) *Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear homomorphism. The function $f : \mathbb{F}_{p^n} \to \mathbb{F}_p$ given by*

$$f(x) = \pi\left(x^{1+p+p^2+p^3}\right)$$

is a disperser for the set of affine spaces of dimension greater than $\frac{n}{3} + 10$ that are not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

In particular, if n is odd, then f is an affine disperser for dimension $\frac{n}{3} + 10$.

Proof The proof is similar to that of Theorem 2.6. Let μ be as before. Let $Q(X)$ be the image-subspace polynomial of $\mathcal{A}$, so that $\mathcal{A} = Q(\mathbb{F}_{p^n})$. Let

$$R(X) = \mathrm{Tr}(\mu \cdot Q(X)^{1+p+p^2+p^3}) \mod \langle X^{p^n} - X \rangle.$$

As in the case of the previous proof, it is sufficient to prove the existence of a quadruple $B \subseteq \{0, \ldots, n-1\}$ which satisfies the conditions of Remark 3.14. We get started by adapting Lemma 3.13 to our present situation.

Claim 4.3 *For $B = \{i_1, \ldots, i_4\} \subseteq \{0, \ldots, n-1\}$ let $\beta = p^{i_1} + p^{i_2} + p^{i_3} + p^{i_4}$. The coefficient $c_{\mathcal{M}}$ of the monomial $\mathcal{M} = X^\beta$ in*

$$\mathrm{Tr}\left(\mu \cdot Q(X)^{p^0+p^1+p^2+p^3}\right) \pmod{X^{p^n} - X}$$

is given by

$$c_{\mathcal{M}} = \sum_{r=0}^{n-1} \mu^{p^r} \mathrm{Perm}\left(M[A, B-r]\right)^{p^r}. \tag{11}$$

As in Remark 3.14, letting $M' = M'_Q$ (as defined in Definition 3.11), we seek $B \subseteq \{0, \ldots, n-1\}$ with $|B| = 4$ such that:

1. The matrix $M'[\{0,1,2,3\}, B]$ is, up to reordering of rows and columns, upper triangular with a nonzero diagonal.
2. For every $r \in \{1, \ldots, n-1\}$ the matrix $M'[\{0,1,2,3\}, B-r]$ contains an all-zero column.

Having found such a B, the above claim lets us conclude that the polynomial $R(X)$ defined above is nonconstant.

As in the analysis of the cubic affine disperser, we begin by stating a few useful properties of the coefficients of Q that all follow immediately from Lemma 3.10 and will be used later on in the proof. Notice that *(iv)* below follows via the second part of Lemma 3.10 from our assumption that $\mathcal{A}$ is not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$.

Claim 4.4 *Let $Q(X) = \sum_{i=0}^{n-1} a_i X^{p^i} + \hat{a}$ be the image-subspace polynomial of $\mathcal{A}$. Letting $d = \dim(\mathcal{A})$ we have (i) $d \geq \frac{n}{3} + 10$, (ii) $a_0, a_{n-d} \neq 0$, (iii) $a_{n-d+1} = \ldots = a_{n-1} = 0$ and (iv) for every $0 \leq j \leq n-d$ there is at least one nonzero coefficient amongst $a_j, a_{j+1}, \ldots, a_{j+d-1}$.*

We use the notation introduced in the proof of Theorem 2.6 in the previous subsection. Recalling the definition of I_0, I_1, notice that *(ii)* implies $\{0, n-d\} \in I_1$, *(iii)* implies $[n-d+1, n-1] \subseteq I_0$ and *(iv)* implies $I_1 \cap [j, j+d-1] \neq \emptyset$.

As stated earlier, we show that a set B satisfying part 2 of Claim 4.3 exists, thereby proving Theorem 2.7. Our proof is divided into three cases according to the structure of I_0, I_1.

Case I — $I_1 \cap [n/3-14, n/3+4] \neq \emptyset$**:** Let $j \in I_1 \cap [n/3-14, n/3+4]$. We claim $B = \{0, 1, j+2, n-d+3\}$ satisfies the two properties. Property 1 holds because

$$M'[A,B] = \begin{pmatrix} a'_0 & a'_1 & a'_{j+2} & a'_{n-d+3} \\ a'_{n-1} & a'_0 & a'_{j+1} & a'_{n-d+2} \\ a'_{n-2} & a'_{n-1} & a'_j & a'_{n-d+1} \\ a'_{n-3} & a'_{n-2} & a'_{j-1} & a'_{n-d} \end{pmatrix} = \begin{pmatrix} 1 & * & * & 0 \\ 0 & 1 & * & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & * & 1 \end{pmatrix}, \quad \text{where } a'_i = \begin{cases} 0 & a_i = 0 \\ 1 & a_i \neq 0 \end{cases}.$$

Regarding property 2, consider

$$M'[A, B-r] = \begin{pmatrix} a'_{n-r} & a'_{n+1-r} & a'_{j+2-r} & a'_{n-d+3-r} \\ a'_{n-1-r} & a'_{n-r} & a'_{j+1-r} & a'_{n-d+2-r} \\ a'_{n-2-r} & a'_{n-1-r} & a'_{j-r} & a'_{n-d+1-r} \\ a'_{n-3-r} & a'_{n-2-r} & a'_{j-1-r} & a'_{n-d-r} \end{pmatrix}.$$

For $r \in [1, d-4]$, the first column of $M'[A, B-r]$ is zero, since the set of indices appearing there

$$[n-3-r, n-r] \subseteq [n-d+1, n-1] \subseteq I_0$$

The last inclusion follows from Claim 4.4 *(iii)* . Similarly, for $r \in [n-(d-4), n-1]$ the last column of $M'[A, B-r]$ is zero, since

$$[n-d+3-r, n-d-r] \subseteq [n-d+1, n-1] \subseteq I_0.$$

Finally, for the remaining $r \in [d-3, n-(d-3)] \subseteq [n/3+7, 2n/3-7]$ the third column of $M'[A, B-r]$ is zero by selection of $j \in [n/3-14, n/3+4]$, since

$$[j-1-r, j+2-r] \subseteq [(n/3-15)-(2n/3-7), (n/3+6)-(n/3+7)] \subseteq [2n/3-8, n-1] \subseteq I_0,$$

where the last inclusion uses Claim 4.4 *(i)* . We conclude property 2 also holds and the proof of the first case is complete.

Case II — $I_1 \cap [n/3-14, n/3+4] = \emptyset$**:** As in the proof of Theorem 2.6, let j_1 be the largest element in $[0, n/3-15] \cap I_1$ and let j_2 be the minimal element in $[n/3+5, n-d] \cap I_1$. By Claim 4.4 *(iv)* we cannot have both $j_1 = 0$ and $j_2 = n-d$. Consider the following four intervals: $[0, j_1]$ (whose end points are in I_1), $[j_1+1, j_2-1]$ (which is contained int I_0), $[j_2, n-d]$ (whose end points are in I_1), and $[n-d+1, n-1]$ (which is contained in I_0). Denote the length of these intervals by $\alpha_1, \ldots, \alpha_4$ respectively. Notice $\alpha_4 = d-2 > \alpha_1 + 10, \alpha_3 + 10$. There are two subcases.

Case II.a — $\alpha_1 \neq \alpha_3$**:** Assume without loss of generality $\alpha_1 > \alpha_3$. We claim that the set $B = \{0, j_2+1, n-d+2, j_1+3\}$ satisfies both properties. (The case of $\alpha_1 < \alpha_3$ can be seen to be identical by using the argument below to show that $B = \{j_2, 1, j_1+2, n-d+3\}$ satisfies our properties.) Property 1 holds because

$$M'[A,B] = \begin{pmatrix} a'_0 & a'_{j_2+1} & a'_{n-d+2} & a'_{j_1+3} \\ a'_{n-1} & a'_{j_2} & a'_{n-d+1} & a'_{j_1+2} \\ a'_{n-2} & a'_{j_2-1} & a'_{n-d} & a'_{j_1+1} \\ a'_{n-3} & a'_{j_2-2} & a'_{n-d-1} & a'_{j_1} \end{pmatrix} = \begin{pmatrix} 1 & * & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & * & 1 \end{pmatrix}.$$

Regarding property 2, the key observation is that any the difference between the top left and bottom right indices of $M'[A, B-r]$ — and this difference is independent of r — equals $j_1 = \alpha_1 - 1$, and similarly the difference between any other pair of indices chosen one from each of the first and third columns is between j_1 and j_1+6. Since $\alpha_4 - 10 > \alpha_1 > \alpha_3$ and $\alpha_2 \geq 20$ the only shifts r that make both the first and the last columns nonzero must satisfy $r \in [n-j_1, n-10]$. This implies that the third column is zero (since this choice of r puts the indices of its entries in I_0), hence property 2 holds and the proof of this case is complete.

Case II.b — $\alpha_1 = \alpha_3$**:** We claim that $B = \{0, j_2+1, n-d+2, j_1+3\}$ satisfies both our properties. Indeed, property 1 holds by the reasoning of case II.a. By the same reasoning as in that case, the only nonzero shift r for which both the first and last columns are nonzero is the shift $r = n - j_2$ which gives

$$M'[A, B-(n-j_2)] = \begin{pmatrix} a'_{j_2} & a'_{2j_2+1} & a'_{n-d+j_2+2} & a'_{j_2+j_1+3} \\ a'_{j_2-1} & a'_{2j_2} & a'_{n-d+j_2+1} & a'_{j_2+j_1+2} \\ a'_{j_2-2} & a'_{2j_2-1} & a'_{n-d+j_2} & a'_{j_2+j_1+1} \\ a'_{j_2-3} & a'_{2j_2-2} & a'_{n-d+j_2-1} & a'_{j_2+j_1} \end{pmatrix} = \begin{pmatrix} 1 & * & * & 0 \\ 0 & a'_{2j_2} & a'_{n-d+j_2+1} & 0 \\ 0 & a'_{2j_2-1} & a'_{n-d+j_2} & 0 \\ 0 & * & * & 1 \end{pmatrix}.$$

The last column is calculated using $\alpha_1 = \alpha_3$ which implies $j_2 + j_1 = n - d$. Consider the middle 2×2 matrix on the right hand side above. The difference between the upper left and bottom right indices is $n - d + j_2 - 2j_2 = n - d - j_2 = j_1$ and that between the bottom left and upper right is $j_1 + 2$. Thus, we conclude $a'_{n-d+j_2+1} = a'_{2j_2-1} = 0$. Claim 4.4 *(iv)* , which relies on the fact that $\mathcal{A}$ is not contained in an affine shift of a proper subfield, implies that $\alpha_2 < \alpha_4$. Together with the assumption $\alpha_1 = \alpha_3$ we conclude $\alpha_1 + \alpha_2 < \alpha_3 + \alpha_4$. This implies $j_2 < n/2$ which, together with the assumption $j_2 > n/3$ gives us $n - d < 2j_2 < n$. This implies, via part *(i)* of Claim 4.4, that $a_{2j_2} = 0$ and the third column is all zero. This shows property 2.

Summing up, in each of the three cases above, we have shown the existence of a set B that satisfies both properties of part 2 of Claim 4.3. This implies $R(X)$ is nonzero and Theorem 2.7 follows. □

5 Cubic affine dispersers are affine extractors

In this section we show that cubic polynomials that are dispersers for dimension d are also extractors for dimension $d' \gg d$ and the bias of these extractors decreases as d' grows larger. The method-of-proof of Theorem 2.9, restated next, is very different from what we use in the rest of this paper. It

relies on an *energy-increment* argument regarding random *directional derivatives* and does not use subspace polynomials. (In particular, to follow the proof one does not need Sections 3–4).

Theorem 2.9 (Cubic affine dispersers are affine extractors, restated) *There exists a universal constant $\epsilon > 0$ such that the following holds. Let $f : \mathbb{F}_p^m \to \mathbb{F}_p$ be computed by a cubic polynomial. If f is a disperser for degree d_0 then f is an extractor for dimension $d_0 + \hat{d}$ with bias $\leq \hat{d}^\epsilon$.*

The rest of this section is devoted to a proof of the theorem for the binary case of $p = 2$, stated below. This proof can be readily extended to hold for all prime p. Since a tighter bound is provided in the recent work of Haramaty and Shpilka [2009] we prefer simplicity of proofs to generality of statement.

Theorem 5.1 (Cubic affine dispersers are affine extractors — binary case) *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2, \deg(f) = 3, \mathrm{bias}(f) \geq \beta > 0$. Then f is constant on a $\mathbb{F}_2$-affine space $S \subseteq \mathbb{F}_2^n$ of co-dimension at most $O\left(\frac{\log 1/\beta}{\beta^2}\right)$.*

Roughly speaking, the proof of this theorem goes by way of contradiction as follows. Assume $\mathcal{A}$ is an affine space of very large dimension on which f is very biased. Using an *energy-increment* argument (Claim 5.7) we restrict our view to a subspace of $\mathcal{A}$ that is also of very large dimension, and on which f is very biased and all its directional derivatives are very small. In this case f is very-well approximated by a function of a constant number of its directional derivatives, and these derivatives are quadratic functions (this idea of approximating a function by its directional derivatives is from Bogdanov and Viola [2007].) Using our understanding of biased quadratic polynomials we conclude that f is in fact *constant* on a subspace of $\mathcal{A}$ of very large dimension, contradicting the assumption that f is an affine disperser.

5.1 Bias of random directional derivatives

We start by introducing a formal notation for the bias of a function on a subset of inputs.

Definition 5.2 (Bias) *For $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $S \subset \mathbb{F}_2^n$ let*

$$\beta_S(f) = \mathbb{E}_x[(-1)^{f(x)}]$$

and let $\mathrm{bias}_S(f) = |\beta_S(f)|$. For simplicity of notation let $\beta(f)$ and $\mathrm{bias}(f)$ denote $\beta_{\mathbb{F}_2^n}(f)$ and $\mathrm{bias}_{\mathbb{F}_2^n}(f)$, respectively.

The following lemma, an immediate consequence of Dickson's Theorem (cf. [Lidl and Niederreiter, 1997, Chapter 6]), characterizes the set of quadratic functions with large bias.

Lemma 5.3 (Biased quadratics) *If $\deg(f) \leq 2$ and $\mathrm{bias}(f) = \beta > 0$ then $\mathbb{F}_2^n$ can be partitioned into $2/\beta$ affine spaces of co-dimension $1 + \log 1/\beta$ such that f is constant on each of them.*

Proof By Dickson's Theorem, up to a linear renaming of variables, f is of the form

$$f(x_1,\ldots,x_n)=\sum_{i=1}^{r}x_{2i-1}x_{2i}+ax_{2r+1}+b \quad \text{for some } a,b\in\mathbb{F}_2.$$

Since $\beta>0$ we must have $a=0$. Clearly $\mathrm{bias}(f)=2^{-r}$, or, equivalently, $r=\log\frac{1}{\beta}$. Let $I=\{2,4,\ldots,2r\}$. For each $\rho\in\mathbb{F}_2^I$ notice $\deg(f|_\rho)\le 1$. Thus, $f|_\rho:\mathbb{F}_2^{[n]\setminus I}\to\mathbb{F}_2$ is either constant on $\mathbb{F}_2^{[n]\setminus I}$ or else, being an affine-linear function, can be partitioned into two affine spaces such that $f|_\rho$ is constant on each. □

Let us recall the notion of a random directional derivative.

Definition 5.4 (Bias of random directional derivatives) *For $a\in\mathbb{F}_2^n$ let f_a be the* directional derivative of f in direction a, *defined by $f_a(x)=f(x+a)-f(x)$. Let Z denote the random variable $\beta(f_a)$ over random direction a.*

We now bound the first and second moment of the random variable Z. The expectation in both claims is over a uniformly random direction a.

Claim 5.5 (First moment) $\mathbb{E}_a[Z]=\beta^2(f)$.

Proof We have

$$\begin{aligned}\mathbb{E}_a[Z] &= \mathbb{E}_a\left[\mathbb{E}_x\left[(-1)^{f(a+x)-f(x)}\right]\right]=\mathbb{E}_{a,x}\left[(-1)^{f(a+x)-f(x)}\right]=\mathbb{E}_x\left[(-1)^{f(x)}\mathbb{E}_a\left[(-1)^{f(a+x)}\right]\right]\\ &= \beta(f)\cdot\mathbb{E}_x\left[(-1)^{f(x)}\right]=\beta^2(f).\end{aligned}$$

□

For $\alpha\in\mathbb{F}_2^n$ let $\hat{f}_\alpha$ be the α-Fourier coefficient of f defined by

$$\hat{f}_\alpha=\mathbb{E}_x[(-1)^{f(x)}\cdot\chi_\alpha(x)]$$

where $\chi_\alpha:\mathbb{F}_2^n\to\{-1,1\}$ is the α-character of $\mathbb{F}_2^n$ defined by $\chi_\alpha(x)=\prod_{i=1}^n(-1)^{\alpha_i\cdot x_i}$. Notice $\beta(f)=\hat{f}_0$.

Claim 5.6 (Second moment) $\mathrm{Var}[Z]=\sum_{\alpha\neq 0}\hat{f}_\alpha^4$.

Proof We have

$$\begin{aligned}\mathbb{E}[Z^2] &= \mathbb{E}_a\left[\mathbb{E}_x\left[(-1)^{f(a+x)-f(x)}\right]\cdot\mathbb{E}_{x'}\left[(-1)^{f(a+x')-f(x')}\right]\right]\\ &= \mathbb{E}_{a,x,x'}\left[(-1)^{f(a+x)+f(x)+f(a+x')+f(x')}\right]\\ &= \mathbb{E}_{a,x,x'}\left[\sum_{\alpha_1,\ldots,\alpha_4}\hat{f}_{\alpha_1}\cdot\hat{f}_{\alpha_2}\cdot\hat{f}_{\alpha_3}\cdot\hat{f}_{\alpha_4}\cdot\chi_{\alpha_1}(a+x)\cdot\chi_{\alpha_2}(x)\cdot\chi_{\alpha_3}(a+x')\cdot\chi_{\alpha_4}(x')\right]\\ &= \sum_{\alpha_1,\ldots,\alpha_4}\hat{f}_{\alpha_1}\cdot\hat{f}_{\alpha_2}\cdot\hat{f}_{\alpha_3}\cdot\hat{f}_{\alpha_4}\cdot\mathbb{E}_a\left[\chi_{\alpha_1+\alpha_3}(a)\right]\cdot\mathbb{E}_x\left[\chi_{\alpha_1+\alpha_2}(x)\right]\cdot\mathbb{E}_{x'}\left[\chi_{\alpha_3+\alpha_4}(x')\right]\\ &= \sum_{\alpha}\hat{f}_\alpha^4.\end{aligned}$$

Thus,

$$\mathrm{Var}[Z] = \left|\mathbb{E}_a[Z^2] - \mathbb{E}_a^2[Z]\right| = \sum_\alpha \hat{f}_\alpha^4 - \hat{f}_0^4.$$

The last equality follows from Claim 5.5 and $\beta(f) = \hat{f}_0$ and this completes the proof. □

The next claim says that if f has a large Fourier coefficient, then f has an even larger bias on a subspace of codimension 1. Such a claim is known as an *energy increment* argument and appears in many proofs in additive combinatorics, a notable example is the proof of Roth's Theorem over $\mathbb{F}_3^n$ of Meshulam [1995] (cf. [Tao and Vu, 2006, Chapter 10]).

Claim 5.7 (Energy increment) *For $\alpha \neq 0$ let $L_0 = \{x \in \mathbb{F}_2^n \mid \langle x, \alpha\rangle = 0\}$ and $L_1 = \mathbb{F}_2^n \setminus L_0$. Then for every function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and $b \in \{0,1\}$ we have*

$$\beta_{L_b}(f) = \beta(f) + (-1)^b \cdot \hat{f}_\alpha.$$

Proof Assume $b = 0$ (the proof for the case of $b = 1$ is identical). Let $\beta = \beta(f)$. By definition we have

$$\beta = \Pr_x[f(x) = 0] - \Pr_x[f(x) = 1].$$

Thus,

$$\Pr_x[f(x) = 0] = \frac{1+\beta}{2}.$$

Similarly, letting , $\beta' = \beta_{L_0}(f)$ we get

$$\Pr_{x \in L_0}[f(x) = 0] = \frac{1+\beta'}{2}.$$

Finally, letting $\gamma = \hat{f}_\alpha$ we have

$$\frac{1}{2}\left(\Pr_{x \in L_0}[f(x) = 0] + \Pr_{x \in L_1}[f(x) = 1]\right) = \frac{1+\gamma}{2}.$$

The second summand on the left hand side in the previous equation can be rewritten as

$$\Pr_{x \in L_1}[f(x) = 1] = 2\left(\Pr_x[f(x) = 1] - \frac{1}{2}\Pr_{x \in L_0}[f(x) = 1]\right)$$

Plugging in $\Pr_x[f(x) = 1] = \frac{1-\beta}{2}$ and $\Pr_{x \in L_0}[f(x) = 1] = \frac{1-\beta'}{2}$ we get

$$\frac{1+\gamma}{2} = \frac{1}{2}\left(\frac{1+\beta'}{2} + (1-\beta) - \frac{1-\beta'}{2}\right).$$

Rearranging terms we get $1 + \gamma = 1 + \beta' - \beta$ which gives $\beta' = \beta + \gamma$ as claimed. □

5.2 Proof of Theorem 5.1

To prove Theorem 5.1 we find, using Claim 5.7, a subspace on which f is biased and all its nonzero Fourier coefficients are small. On this subspace, by Claim 5.6 the variance of f is very small. So for most directions the derivative f_a, which has degree 2, is also biased. The work of Bogdanov and Viola [2007] showed that in this case f can be very well approximated by a majority of (random) derivatives. Since our random derivatives are biased we can use Lemma 5.3 to argue that they jointly partition the space into a small number of affine spaces on which f is constant. But since f agrees with the majority of derivatives almost everywhere there must be a subspace on which the two agree completely. Details follow.

Proof of Theorem 5.1: We use positive constants c_1, c_2 whose value will be fixed later. First, using the energy increment Claim 5.7 we inductively construct a sequence of affine spaces $A_0 = \mathbb{F}_2^n \supset A_1 \supset A_2 \supset \ldots$ and sequence of functions $f_i : A_i \to \mathbb{F}_2$ of increasing biases $\beta_0 = \beta < \beta_1 < \ldots$ as follows. Let $f_0 = f$. For $i \geq 0$ if f_i has a Fourier coefficient that is greater in magnitude than β^2/c_1 then by Claim 5.7 there exists a subspace $A' \subset A_i$ such that $\mathrm{bias}_{A'}(f_i) \geq \mathrm{bias}(f_i) = \beta_i + \beta^2/c_1$ so set $A_{i+1} = A'$ and let f_{i+1} be the restriction of f to A_{i+1}. By induction we have $\beta_i \geq \beta(1+\beta^2/c_1)^i$ we conclude that this process must stop after $t \leq c_1/\beta^2$ steps. Let $f' = f_t$ and notice that $f' : \mathbb{F}_2^m \to \mathbb{F}_2$ for $m = n - t \geq n - c_1/\beta^2$ satisfies $\mathrm{bias}(f') \geq \beta$ and for all $\alpha \neq 0$ we have $|\hat{f'}_\alpha| \leq \beta^2/c_1$.

Next, set $s = c_2/\beta^2$ and pick s random directions $a_1, \ldots, a_s \in \mathbb{F}_2^m$. Let B_i be the "bad" event that $\mathrm{bias}(f'_{a_i}) < \beta^2/2$. Let C be the "bad" event that $\Pr_{x \in \mathbb{F}_2^m}[f(x) \neq \mathrm{majority}(f_{a_1}(x), \ldots, f_{a_s}(x))] > 2^{-5}$. Theorem 5.1 follows from the following claim, whose proof appears below.

Claim 5.8

$$\Pr_{a_1,\ldots,a_s \in \mathbb{F}_2^m}[B_1 \vee \ldots \vee B_s \vee C] < 1.$$

Let us complete the proof of the theorem. Fix $a_1, \ldots, a_s$ for which none of $B_1, \ldots, B_s, C$ occur, Claim 5.8 proves these a_i's exist. We have $\mathrm{bias}(f'_{a_i}) \geq \beta^2/2$ for $i = 1, \ldots, s$. By Lemma 5.3 each f'_{a_i} partitions $\mathbb{F}_2^m$ into spaces of codimension $\leq 2 + 2\log 1/\beta$ each, such that f'_{a_i} is constant on each such space. Thus, we can partition F_2^m into affine spaces of codimension $\leq s \cdot (2 + 2\log 1/\beta) \leq O(\frac{\log 1/\beta}{\beta^2})$ each such that for each such space A we have that f'_{a_i} is constant on A for $i = 1, \ldots, s$.

Furthermore, for all but a 2^{-5} fraction of inputs, we have that f' agrees with $g = \mathrm{majority}(f'_{a_1}, \ldots, f'_{a_s})$. Consider one of the spaces A. If f' does not agree with g completely on A then the two must disagree on a 2^{-3} fraction of points of A, because g is constant on A so the degree of the function $f - g$, when restricted to A, is at most 3. We conclude that there must be some space on which f' and g agree completely. But g is fixed on A which is an affine space of codimension $O(\frac{\log 1/\beta}{\beta})$. Adding the codimension of $n - m \leq O(1/\beta^2)$ which we incurred when moving from f to f' we complete the proof of Theorem 2.2. □

Proof of Claim 5.8: First consider B_i. Using the random variable Z defined as $\mathrm{bias}(f'_a)$ we have from Claim 5.5 that $\mathbb{E}[Z] \geq \beta^2$ because $\mathrm{bias}(f') \geq \beta$ and from Claim 5.6 we get $\mathrm{Var}[Z] \leq \beta^4/c_1^2$. Using Chebychev's inequality we get

$$\Pr[\mathrm{bias}(f'_{a_i}) \leq \beta^2/2] \leq \frac{4\beta^2}{c_1^2}.$$

Next consider C. Fix $x \in \mathbb{F}_2^m$ and consider the "bad" event C_x which is "$f(x) \neq \text{majority}(f'_{a_1}(x), \ldots, f'_{a_s}(x))$". We know that $f(x) = f'_{a_i}(x)$ if and only if $f(a+x) = 0$ and this happens with probability $\frac{1+\beta}{2}$. Using the independence of a_i and Chernoff's bound we conclude

$$\Pr[C_x] \leq \exp(-\beta^2 s/16).$$

By linearity of expectation the expected fraction of inputs on which $f(x) \neq \text{majority}(f'_{a_1}(x), \ldots, f'_{a_s}(x))$ is at most $\exp(-\beta^2 s/16)$. By Markov's inequality the probability that the fraction of errors is more than $2\exp(-\beta^2 s/16)$ is at most half. We now fix c_1, c_2 so that

$$2\exp(-\beta^2 s/16) = 2\exp(-c_2/16) \leq 2^{-5}$$

and additionally

$$s \cdot \frac{4\beta^2}{c_1^2} = \frac{4c_2}{c_1^2} < 1/2.$$

We conclude using a union bound that the probability that none of C or $B_1, \ldots, B_s$ occur is nonzero and this completes the proof of Claim 5.8. □

5.3 Quartic affine dispersers are not necessarily affine extractors

The next lemma shows that for $f : \mathbb{F}_2^n \to \mathbb{F}_2, \deg(f) > 3$, any connection between $\text{bias}(f)$ and the dimension of the largest subspace on which f is constant, must depend on n. In particular, to show that the quartic affine disperser given in Theorem 2.7 is an affine extractor, one cannot simply rely on the dispersing nature of this function and a deeper algebraic understanding of it is required.

Lemma 5.9 *The function* $f : \left(\mathbb{F}_2^{n/4}\right)^4 \to \mathbb{F}_2$ *defined by*

$$f(x,y,z,w) = \left(\sum_{i=1}^{n/4} x_i y_i\right) \cdot \left(\sum_{i=1}^{n/4} z_i w_i\right)$$

is an affine disperser for dimension $d > 3n/4$ *and has* $\text{bias}(f) = \frac{1}{2} - o(1)$.

Before giving the proof, notice f can be computed by the 4-variate polynomial over $\mathbb{F}_{2^{n/4}}$ given by

$$f(X,Y,Z,W) = \text{Tr}(\mu XY) \cdot \text{Tr}(\mu ZW)$$

where Tr is the trace map from $\mathbb{F}_{2^{n/4}}$ to $\mathbb{F}_2$ and μ is some element of $\mathbb{F}_{2^{n/4}}$ (cf. Proposition 3.3).

Proof The bias of f on $\left(\mathbb{F}_2^{n/4}\right)^4$ can be readily seen to be $\frac{1}{2} - o(1)$ as it is the product of two independent random coins that are each 0 with probability $\frac{1}{2} - o(1)$. Next we argue that f is a disperser. Fix a space $A \subseteq \mathbb{F}_2^{[n]}, \dim(A) > 3n/4$. By Bourgain's decomposition (Lemma 7.1) there exist affine spaces

$$A_0 \subseteq \mathbb{F}_2^{\{1,\ldots,n/2\}}, A_1 \subseteq \mathbb{F}_2^{\{n/2+1,\ldots,n\}}, \dim(A_0) + \dim(A_1) = \dim(A)$$

and an affine transformation

$$T : \mathbb{F}_2^{\{1,\ldots,n/2\}} \to \mathbb{F}_2^{\{n/2+1,\ldots,n\}}$$

such that

$$A = \{a_0 + T(a_0) + a_1 \mid a_0 \in A_0, a_1 \in A_1\}.$$

Notice both $\dim(A_0), \dim(A_1) > n/4$. It was shown in Ben-Sasson et al. [2001] that the function $\sum_{i=1}^{n/4} x_i y_i$ is an affine disperser for dimension $> n/4$ (in fact, it is even an affine extractor). Hence, there exist $(x', y'), (x'', y'') \in A_0$ such that $\sum_{i=1}^{n/4} x'_i y'_i = 0$ and $\sum_{i=1}^{n/4} x''_i y''_i = 1$. In the first case f obtains the value 0. In the second case there exists $(z', w') \in A_1$ such that f evaluates on $(x'', y'') + T(x'', y'') + (w', z')$ to 1. This holds because $T(x'', y'') + A_1$ is an affine space of dimension $> n/4$ and the function $\sum_{i=1}^{n/4} z_i w_i$ is an affine disperser for dimension $> n/4$. We conclude f is not constant on A, thereby completing the proof. □

We end our discussion of cubic extractors by showing that some of them reach dimension as low as $\sqrt{n}$. This offers a possible way to get affine extractors for polynomially small dimension (something out of reach of current techniques) — by constructing a cubic affine disperser for polynomially small dimension.

Lemma 5.10 (Cubic affine dispersers for small dimension) *There exists a cubic function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ that is a disperser for affine spaces of dimension $\Omega(\sqrt{n})$.*

Proof Let f be a *uniformly random* degree 3 polynomial in n variables over $\mathbb{F}_p$. Notice that for any fixed affine space $\mathcal{A}$ of dimension d, the restriction of f to $\mathcal{A}$ is again a uniformly random degree 3 polynomial (but of course, these events are not independent across different $\mathcal{A}$). Letting $\binom{d}{\leq k} = \sum_{j \leq k} \binom{d}{j}$, the probability that the restriction of f to $\mathcal{A}$ is zero equals $p^{-\binom{d}{\leq 3}} \leq p^{-d^3/6}$. Taking a union bound over all subspaces $\mathcal{A}$ of dimension $3\sqrt{n}$, and there are less than $(p^n)^{3\sqrt{n}} = p^{3n^{3/2}}$ of them, we see that the probability that the restriction of f to every dimension $3\sqrt{n}$ subspace is nonzero is at least

$$1 - p^{3n^{3/2}} \cdot p^{-\frac{27}{6} n^{3/2}} > 0,$$

which completes the proof of the lemma. □

6 Disperser for independent affine sources

In this section we prove Theorem 2.3, restated below. Although the analysis is simpler than what is involved in the proof of our main disperser for sublinear dimension (Theorem 2.2), the proof of the following theorem lies at the heart of the more complicated case which is discussed in the next section.

Theorem 2.3 (Disperser for independent affine sources, restated) *Let $\pi : \mathbb{F}_{p^n} \to \mathbb{F}_p$ be a nontrivial $\mathbb{F}_p$-linear map. Consider the function $f : \mathbb{F}_{p^n}^t \to \mathbb{F}_p$ given by*

$$f(x_1, \ldots, x_t) = \pi\left(\prod_{i=1}^{t} x_i^{1+p}\right). \tag{12}$$

Let $\mathcal{A}_1, \ldots, \mathcal{A}_t \subseteq \mathbb{F}_{p^n}$ be $\mathbb{F}_p$-affine spaces of dimensions $d_1, \ldots, d_t$ respectively, where each $\mathcal{A}_i$ is not contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$. If $\sum_{i=1}^{t}(d_i - 2) > n$, then $|f(\mathcal{A}_1 \times \cdots \times \mathcal{A}_t)| > 1$.

Proof We follow the steps outlined in our strategy described in Section 1.2. First, we notice that

$$f(\mathcal{A}_1, \ldots, \mathcal{A}_t) = f(Q_1(\mathbb{F}_{p^n}), \ldots, Q_t(\mathbb{F}_{p^n}))$$

where $Q_i(X_i)$ is the image-subspace polynomial of $\mathcal{A}_i$. By Propositions 3.1, 3.3, in order to show $|f(\mathcal{A}_1 \times \cdots \times \mathcal{A}_t)| > 1$ it suffices to show that for any $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$ the polynomial

$$R(X_1, \ldots, X_t) \stackrel{\text{def}}{=} \mathrm{Tr}\left(\mu \prod_{i=1}^{t} (Q_i(X_i))^{\alpha_i}\right) \mod \langle (X_i^{p^n} - X_i)_{i \in [t]} \rangle$$

contains a monomial of positive degree with nonzero coefficient. We use Lemma 3.13 to prove the existence of such a monomial and in the proof we rely on the structural properties of image-subspace polynomials given in Lemma 3.10.

The key step in our proof is given by the following theorem. We state a somewhat more general form than needed for the proof of Theorem 2.3. The added generality will be useful in the proof of Theorem 2.2. (The general form we refer to deals with large powers α_i whereas for Theorem 2.3 setting all α_i to $1 + p$ would be sufficient.)

Theorem 6.1 (Disperser for independent affine sources — Algebraic version) *Assume that $\mathcal{A}_1, \ldots, \mathcal{A}_t \subseteq \mathbb{F}_{p^n}$ are affine subspaces of dimensions $d_1, \ldots, d_t > 1$, none of which are contained in an affine shift of a proper subfield of $\mathbb{F}_{p^n}$. Let $Q_i(X_i) \in \mathbb{F}_{p^n}[X_i]$ be the image-subspace polynomial of $\mathcal{A}_i$. Let $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$. Let $e_1, \ldots, e_t$ satisfy $1 \leq e_i < d_i - 1$ and let $\alpha_i = \sum_{j=0}^{e_t} p^j$. Let*

$$R(X_1, \ldots, X_t) \stackrel{\text{def}}{=} \mathrm{Tr}\left(\mu \prod_{i=1}^{t} (Q_i(X_i))^{\alpha_i}\right) \mod \langle (X_i^{p^n} - X_i)_{i \in [t]} \rangle \tag{13}$$

If $\sum_{i=1}^{t}(d_i - (e_i + 1)) > n - \max d_i + 1$, then $R(X_1, \ldots, X_t)$ has a monomial $\prod_{i=1}^{t} X_i^{\beta_i}$ with $\mathrm{wt}_p(\beta_i) = e_i + 1$, which has a nonzero coefficient. In particular, $|R(\mathbb{F}_{p^n}^t)| > 1$.

Before giving the proof of Theorem 6.1, let us first show how to use it to complete the proof of Theorem 2.3. We may assume without loss of generality that $d_i > 2$ by fixing nonzero elements of those spaces that have dimension 2. Next, in Theorem 6.1 we set $\mu = 1$ and $e_1 = \ldots = e_t = 1$, which gives $\alpha_1 = \ldots = \alpha_t = 1 + p$. Using Proposition 3.1, the polynomial R defined in (13) satisfies $R(\mathbb{F}_{p^n}^t) = f(\mathcal{A}_1 \times \cdots \times \mathcal{A}_t)$. Since $\sum(d_i - 2) = \sum(d_i - (e_i + 1)) > n$ we conclude from Theorem 6.1 that $|f(\mathcal{A}_1, \ldots, \mathcal{A}_t)| > 1$ and this completes the proof of Theorem 2.3. □

Proof of Theorem 6.1: Let $A_i = \{0, \ldots, e_i\}$. By the first part of Lemma 3.13, if $B_1, \ldots, B_t \subset \{0, \ldots, n-1\}, |B_i| = e_i + 1$ and $\beta_i = \sum_{k \in B_i} p^i$, then the coefficient of $\mathcal{M} = \prod_{i=1}^{t} X_i^{\beta_i}$ in R, which is denoted henceforth by $c_{\mathcal{M}}$, equals

$$\sum_{r=0}^{n-1} \mu^{p^r} \cdot \prod_{i=1}^{t} \mathrm{Perm}\left(M_i[A_i, B_i - r]\right)^{p^r}, \tag{14}$$

where M_i is the matrix associated with Q_i (cf. Definition 3.11). We will find suitable powers β_i with $\mathrm{wt}_p(\beta_i) = e_i + 1$ such that $c_{\mathcal{M}} \neq 0$. We define β_i by specifying B_i with $|B_i| = e_i + 1$ and setting $\beta_i = \sum_{k \in B_i} p^k$.

Assume without loss of generality $d_1 = \max d_i$. To define B_i let $\ell_1 = 0$ and for $1 \leq i < t$ let $\ell_{i+1} = \ell_i + d_i - (e_i+1) \mod n$. In other words, $\ell_i = \sum_{i'<i}(d_i-(e_i+1)) \pmod n$ where $\ell_1 = 0$. Let $Q_i(X_i) = \sum_{j=0}^{n-1} a_{i,j} X^{p^j} + \hat{a}_i$. Our definition of B_i splits into two cases, depending on whether a_{i,ℓ_i} is nonzero or zero. In the first case we set B_i to be the set $\{\ell_i, n-d_i+1, n-d_i+2, \ldots, n-d_i+e_i\}$. In the second case let j_i be the smallest index j' greater than ℓ_i such that $a_{i,j'}$ is nonzero. Similarly, let $\widehat{j}_i$ be the largest index $\widehat{j'}$ smaller than ℓ_i such that $a_{i,\widehat{j'}}$ is nonzero. Let $q_i = j_i - \widehat{j}_i - 1$ be the length of the interval of zero-coefficients of Q_i between indices $\widehat{j}_i$ and j_i. Let $s_i = \min\{q_i, e_i\}$. We set B_i to be the set

$$\{j_i\} \cup \left\{\widehat{j}_i + 1, \ldots, \widehat{j}_i + s_i\right\} \cup \{n - d_i + s_i + 1, \ldots, n - d_i + e_i\}.$$

The last set might be empty in case $s_i = e_i$.

Our proof again employs the strategy of Remark 3.14 via the next two claims, proved below. We point out that the noncontainment of $\mathcal{A}_i$ in a proper subfield and the implication this has on the structure of coefficients of Q_i (cf. Theorem 3.10) will be crucially used in the proof of Claim 6.3 below. Let M'_i denote the zero-one indicator matrix of M_i as given in Definition 3.11.

Claim 6.2 *For all $i = 1, \ldots, t$ the matrix $M'_i[A_i, B_i]$ is lower triangular with nonzero diagonal entries.*

Claim 6.3 *For all $r \in \{1, \ldots, n-1\}$ there exists $i \in \{1, \ldots, t\}$ such that $M'_i[A_i, B_i - r]$ contains an all-zero column.*

Assuming these two claims, Lemma 3.13 and Remark 3.14 imply Theorem 6.1. □

Proof of Claim 6.2: Notice that, by definition, $M'_i[A_i, B_i]$ is a $(e_i + 1) \times (e_i + 1)$ matrix constructed by taking the minor corresponding to the first $e_i + 1$ rows of M'_i and the columns indexed by B_i. To see that $M'_i[A_i, B_i]$ is lower diagonal with nonzero diagonal entries, consider B_i. To simplify notation in this proof let $a_j = a_{i,j}$ be the coefficient of X^{p^j} in $Q(X_i)$ and let a'_j be its zero-one indicator (cf. Definition 3.11). There are two cases.

$\mathbf{a'_{\ell_i} = 1}$: We have $B_i = \{\ell_i, n-d_i+1, \ldots, n-d_i+e_i\}$. Consider the indices j of the coefficients a'_j residing in the various entries of $M'_i[A_i, B_i]$. By assumption $e_i < d_i$ so the entries above the diagonal of $M'_i[A_i, B_i]$ have indices belonging to

$$\{n - d_i + 1, \ldots, n - d_i + e_i\} \subseteq \{n - d_i + 1, \ldots, n - 1\}$$

and this proves $M'_i[A_i, B_i]$ is lower triangular. Regarding the diagonal, at the topmost left entry we have $a'_{\ell_i} = 1$ and in all subsequent positions we have a'_{n-d_i}, which is nonzero by Lemma 3.10. This completes the proof of this case.

$\mathbf{a'_{\ell_i} = 0}$: In this case we have $B_i = \{j_i\} \cup \left\{\widehat{j}_i + 1, \ldots, \widehat{j}_i + s_i\right\} \cup \{n - d_i + s_i + 1, \ldots, n - d_i + e_i\}$ where

$$j_i = \min\{j > \ell_i \mid a_{i,j} \neq 0\} \text{ and } \widehat{j}_i = \max\{j < \ell_i \mid a_{i,j} \neq 0\}.$$

The uppermost left $(s_i+1)\times(s_i+1)$ submatrix of $M'_i[A_i,B_i]$ in this case is

$$\begin{pmatrix} a'_{j_i} & a'_{\widehat{j_i}+1} & \dots & a'_{\widehat{j_i}+s_i} \\ a'_{j_i-1} & a'_{\widehat{j_i}} & \dots & a'_{\widehat{j_i}+s_i-1} \\ \vdots & \vdots & \vdots & \vdots \\ a'_{j_i-s_i} & a'_{(\widehat{j_i}+1)-s_i} & \dots & a'_{\widehat{j_i}} \end{pmatrix}$$

which is lower triangular because the $a'_{\widehat{j_i}+1} = \ldots = a_{\widehat{j_i}+s_i} = 0$, and the diagonal entries of this submatrix are nonzero because $a'_{j_i}, a'_{\widehat{j_i}}$ are nonzero. The last $e_i - s_i$ columns of the matrix — if they exist — are identical to the same last columns of the previous case and this shows that $M'_i[A_i,B_i]$ is lower triangular with nonzero diagonal entries. This completes the proof of Claim 6.2. □

Proof of Claim 6.3: In what follows we denote for $c<d$ by $[c,d]$ the set of integers in the interval $[c,d]$ and by $[c,d] \bmod n$ the set $\{i \bmod n \mid i \in [c,d]\}$. We start by observing that

$$M'_i[A_i,\{k\}-r] = \begin{pmatrix} a'_{i,k-r} \\ a'_{i,k-(r+1)} \\ \vdots \\ a'_{i,k-(r+e_i)} \end{pmatrix}.$$

Thus for any $k \in B_i$, if r is such that

$$[k-(r+e_i), k-r] \bmod n \subseteq [n-d_i+1, n-1], \tag{15}$$

then the matrix $M'_i[A_i, B_i - r]$ contains a zero column. So we get the following proposition.

Proposition 6.4 *Whenever $k \in B_i$ and*

$$r \in [k+1, k+d_i-(e_i+1)] \bmod n$$

then $M'_i[A_i, B_i - r]$ contains an all-zero column.

Thus, to prove the claim it suffices to show

$$[1,n-1] \subseteq \bigcup_{i=1}^{t} \cup_{k\in B_i} [k+1, k+(d_i-e_i)-1]. \tag{16}$$

(Notice that Claim 6.2 implies the containment in the previous equation is in fact an equality.)

Indeed, since $\ell_1 = 0$ we have $B_1 = \{0\} \cup [n-d_1+1, n-d_1+e_1]$, which implies by Proposition 6.4 that $M'_1[A_1, B_1 - r]$ contains a zero column for r belonging to

$$[1, d_1-(e_1+1)] \cup [n-d_1+2, n-1] = [\ell_1+1, \ell_2] \cup [n-d_1+2, n-1]. \tag{17}$$

Let t' be the minimal i such that $\sum_{i'\le i}(d_{i'}-(e_{i'}+1)) \ge n-d_1+1$, noticing such t' exists by assumption. In this case we have $\sum_{i'\le t'}(d_{i'}-(e_{i'}+1)) < n$ and so $\ell_{t'+1} = \sum_{i'\le t'}(d_{i'}-(e_{i'}+1))$.

We claim that for $1 < i \leq t'$ we have

$$\bigcup_{k \in B_i} [k+1, k+d_i-(e_i+1)] \supseteq [\ell_i+1, \ell_{i+1}]. \tag{18}$$

which, together with (17), proves (16) and completes the proof of our claim. There are two cases to consider when proving (18).

$\mathbf{a_{i,\ell_i} \neq 0}$: In this case $\ell_i \in B_i$ so the claim follows from Proposition 6.4 by recalling that $\ell_{i+1} = \ell_i + d_i - (e_i+1)$.

$\mathbf{a_{i,\ell_i} = 0}$: There are two subcases to consider.

Case 1: $q_i < e_i$**.** In this case

$$B_i = \{j_i\} \cup [\widehat{j}_i + 1, \widehat{j}_i + q_i] \cup [n - d_i + q_i + 1, n - d_i + e_i].$$

Substituting $\widehat{j}_i + (q_i+1)$ for j_i and reordering elements of B_i we get

$$B_i = [\widehat{j}_i + 1, j_i = \widehat{j}_i + q_i + 1] \cup [n - d_i + q_i + 1, n - d_i + e_i].$$

We conclude $\ell_i \in B_i$ so by Proposition 6.4 our proof is complete, as in the case of $a_{i,\ell_i} \neq 0$ above.

Case 2: $q_i \geq e_i$**.** In this case we have

$$B_i = \{j_i\} \cup [\widehat{j}_i + 1, \widehat{j}_i + e_i].$$

Substituting $j_i = \widehat{j}_i + q_i + 1$ we get

$$B_i = [\widehat{j}_i + 1, \widehat{j}_i + e_i] \cup \left\{\widehat{j}_i + q_i + 1\right\}$$

Now we use the fact that $\mathcal{A}_i$ is not contained in an affine shift of a proper subfield. We notice that since $i \leq t'$ we have by maximality of d_1 that

$$\widehat{j}_i < \ell_i \leq n - d_1 \leq n - d_i$$

which implies (using the maximality of d_1 again) that $\widehat{j}_i + 1 \neq n - d_i + 1$. As $\mathcal{A}_i$ is not contained in an affine shift of a proper subfield and $\widehat{j}_i + 1 \neq n - d_i + 1$, our Structural Lemma 3.10 implies that $j_i - \widehat{j}_i \leq d_i - 1$, or, equivalently, $j_i \leq \widehat{j}_i + d_i - 1$.

Taking all but the last element of B_i in the previous equation notice

$$\bigcup_{k \in [\widehat{j}_i+1, \widehat{j}_i+e_i]} [k+1, k+d_i-(e_i+1)] \supseteq [\widehat{j}_i + 2, \widehat{j}_i + d_i - 1],$$

which contains j_i. Now, since $\widehat{j}_i < \ell_i < j_i$ when we reinsert j_i into B_i we conclude

$$\begin{aligned} \cup_{k \in B_i} [k+1, k+d_i-(e_i+1)] &\supseteq [\widehat{j}_1 + 2, j_i + d_i - (e_i+1)] \\ &\supseteq [\ell_i + 1, \ell_{i+1}]. \end{aligned}$$

This completes the last case and with it the proof of Claim 6.3 $\square$

7 Disperser for affine spaces of sublinear dimension

In this section we prove Theorem 2.2. We start by examining what happens to $\mathcal{A} \subset \mathbb{F}_p^{nr}$ when it is partitioned into r blocks of size n. Then we prove the main theorem, by essentially reducing it to the case of independent affine sources described in Theorem 6.1.

7.1 Preparatory lemmata

Our first lemma, already used by Bourgain [2007] in his construction of affine extractors, gives a certain kind of direct sum decomposition of $\mathbb{F}_p$-affine subspaces of $\mathbb{F}_{p^n}^r$.

Lemma 7.1 (Bourgain's decomposition) *Let $\mathcal{A} \subseteq (\mathbb{F}_p^n)^r$ be an $\mathbb{F}_p$-affine subspace. Let $\gamma \in \mathcal{A}$. Then there exist linear spaces $Y_1, \ldots, Y_r \subseteq \mathbb{F}_p^n$ and linear maps $\sigma_{ij} : Y_j \to \mathbb{F}_p^n$ such that:*

$$\mathcal{A} = \{(x_1, \ldots, x_r) \mid \exists y_i \in Y_i \text{ such that } x_i = \gamma_i + y_i + \sum_{j<i} \sigma_{ij}(y_j)\}$$

and $\dim \mathcal{A} = \sum_{i \in [r]} \dim Y_i$.

This lemma amounts to taking the echelon-form of a matrix whose rows form a basis for the linear subspace underlying the affine subspace $\mathcal{A}$.

The next lemma should be thought of as a complement to Theorem 6.1. It expands the class of sources on which the function R given in that theorem is nonconstant. This expanded class is what we will use in the proof of our main theorem.

Lemma 7.2 *For each $i \in [r]$, let $P_i(X_i) \in \mathbb{F}_{p^n}[X_i]$ be a linearized polynomial. For each $j < i$, let $P_{ij}(X_j) \in \mathbb{F}_{p^n}[X_j]$ be a linearized polynomial. Let $\gamma \in \mathbb{F}_{p^n}^r$. Let $I_0 \subseteq [r]$ with $I_0 = \{i_1 < i_2 < \ldots < i_t\}$. Let $e_{i_1}, \ldots, e_{i_t} > 1$ be integers and let $\alpha_i = \sum_{k=0}^{e_i} p^i$, for $i \in I_0$. Let $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$. Let $g(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$ be the polynomial*

$$\mathrm{Tr}\left(\mu \prod_{i \in I_0} \left(P_i(X_i) + \sum_{j<i} P_{ij}(X_j) + \gamma_i\right)^{\alpha_i}\right) \bmod \langle X_i^q - X_i \rangle_{i \in [r]}.$$

Let $g'(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$ be the polynomial

$$\mathrm{Tr}\left(\mu \prod_{i \in I_0} P_i(X_i)^{\alpha_i}\right) \bmod \langle X_i^q - X_i \rangle_{i \in [r]}.$$

Then for any $(\beta_{i_1}, \ldots, \beta_{i_t})$ where $\mathrm{wt}_p(\beta_{i_k}) = e_{i_k} + 1$, the coefficients of the monomial $\prod_{i \in I_0} X_i^{\beta_i}$ in g and in g' are equal.

Proof We want to show that the coefficient of $\prod_{i\in I_0} X_i^{\beta_i}$ in $g(X_1,\ldots,X_r)$ is the same as in $g'(X_1,\ldots,X_r)$. We do this by expanding out the expressions for $g(X_1,\ldots,X_r)$ and $g'(X_1,\ldots,X_r)$ and keeping track of the monomials.

Let $X_{<i}$ denote the tuple of variables $(X_1,\ldots,X_{i-1})$. Let $\hat{P}_i(X_{<i})$ be the polynomial $\sum_{j<i} P_{ij}(X_j)+\gamma_i$.

Expanding $g(X_1,\ldots,X_r)$ we get

$$g(X_1,\ldots,X_r) = \sum_{r=0}^{n-1}\left(\mu\prod_{i\in I_0}\left(P_i(X_i)+\hat{P}_i(X_{<i})\right)^{\alpha_i}\right)^{p^r} \mod \langle (X_i^{p^n}-X_i)_{i\in[r]}\rangle \quad (19)$$

$$= \sum_{r=0}^{n-1}\mu^{p^r}\left(\prod_{i\in I_0}\left(P_i(X_i)+\hat{P}_i(X_{<i})\right)^{\sum_{l=0}^{e_i}p^{r+l}}\right) \mod \langle (X_i^{p^n}-X_i)_{i\in[r]}\rangle. \quad (20)$$

We now expand the term $\left(P_i(X_i)+\hat{P}_i(X_{<i})\right)^{\sum_{l=0}^{e_i}p^{r+l}}$ to obtain

$$\prod_{l=0}^{e_i}\left(P_i(X_i)+\hat{P}_i(X_{<i})\right)^{p^{r+l}} \quad (21)$$

$$=\prod_{l=0}^{e_i}\left(P_i(X_i)^{p^{r+l}}+\hat{P}_i(X_{<i})^{p^{r+l}}\right) \quad (22)$$

$$=\prod_{l=0}^{e_i}(P_i(X_i))^{p^{r+l}}+\sum_{\substack{L\subseteq\{0,1,\ldots,e_i\}\\ L\neq\emptyset}}\left(\prod_{l\notin L}P_i(X_i)^{p^{r+l}}\right)\left(\prod_{l'\in L}\hat{P}_i(X_{<i})^{p^{r+l'}}\right) \quad (23)$$

Now the first term has $\mathbb{F}_p$-degree in X_i equal to e_i+1, while all the other terms have $\mathbb{F}_p$-degree in X_i strictly less than e_i+1. The reason for this is that the polynomial $\hat{P}_i(X_{<i})$ does not mention the variable X_i, and each $P_i(X_i))^{p^{r+l}}$ and $\hat{P}_i(X_{<i})^{p^{r+l'}}$ are linearized polynomials (and hence of $\mathbb{F}_p$-degree 1). Let us summarize this by writing (23) as

$$(P_i(X_i))^{\sum_{l=0}^{e_i}p^{r+l}}+G_i(X_{\le i})$$

and noting that the $\mathbb{F}_p$-degree of G_i in X_i is at most e_i.

Now let us go back to (20) and consider the rth summand within the parenthesis.

$$\begin{aligned}\prod_{i\in I_0}\left(P_i(X_i)+\hat{P}_i(X_{<i})\right)^{\sum_{l=0}^{e_i}p^{r+l}} &= \prod_{i\in I_0}\left((P_i(X_i))^{\sum_{l=0}^{e_i}p^{r+l}}+G_i(X_{\le i})\right)\\ &= \left((P_{i_t}(X_{i_t}))^{\sum_{l=0}^{e_{i_t}}p^{r+l}}+G_{i_t}(X_{\le i_t})\right)\cdot\prod_{i\in I_0, i<i_t}\left((P_i(X_i))^{\sum_{l=0}^{e_i}p^{r+l}}+G_i(X_{\le i})\right)\\ &= \left((P_{i_t}(X_{i_t}))^{\sum_{l=0}^{e_{i_t}}p^{r+l}}+G_{i_t}(X_{\le i_t})\right)\cdot H_t(X_{\le i_{t-1}})\end{aligned}$$

The rightmost term above, denoted H_t, does not mention X_{i_t}. Furthermore, as stated above G_t has $\mathbb{F}_p$-degree at most e_{i_t} in X_{i_t}. But m_ξ has $\mathbb{F}_p$-degree $e_{i_t}+1$ in X_{i_t}, so to contribute to the coefficient of m_ξ we must select terms *only* from $(P_{i_t}(X_t))^{\sum_{l=0}^{e_{i_t}} p^{r+l}}$ and multiply them by the appropriate terms in H_t. Next, consider the terms inside H_t,

$$H_t(X_{\leq i_{t-1}}) = \left(\left(P_{i_{t-1}}(X_{i_{t-1}})\right)^{\sum_{l=0}^{e_{i_{t-1}}} p^{r+l}} + G_{i_{t-1}}(X_{\leq i_{t-1}}) \right) \cdot H_{t-1}(X_{\leq i_{t-2}})$$

where,

$$H_{t-1}(X_{\leq i_{t-2}}) = \prod_{i \in I_0, i < i_{t-1}} \left((P_i(X_i))^{\sum_{l=0}^{e_i} p^{r+l}} + G_i(X_{\leq i}) \right)$$

As before, we notice that $H_{t-1}(X_{\leq i_{t-2}})$ does not mention $X_{i_{t-1}}$ and H_{t-1} has $\mathbb{F}_p$-degree $e_{i_{t-1}}$ in $X_{i_{t-1}}$. But m_ξ has $\mathbb{F}_p$-degree $e_{i_{t-1}}+1$ in $X_{i_{t-1}}$, implying that we must select terms *only* from $\left(P_{i_{t-1}}(X_{i_{t-1}})\right)^{\sum_{l=0}^{e_{i_{t-1}}} p^{r+l}}$. Continuing in this manner for $i = i_{t-2}, \ldots, i_1$ we conclude that the only contributions to the coefficient of m_ξ come from $\left(\prod_{i \in I_0} (P_i(X_i))^{\sum_{l=0}^{e_i} p^l}\right)^{p^r}$. Summing up over all r, the lemma follows. □

7.2 Proof of Theorem 2.2

We can now analyze our main affine disperser construction. Theorem 2.2 will follow by setting the proper parameters into the following theorem.

Theorem 7.3 (Affine disperser — non-parameterized version) *Let $t < r$ be integers. Let n be prime with $n \geq r(r+1)/t$. For each $i \in [r]$, let $e_i = r+1-i$ and let $\alpha_i = \sum_{k=0}^{e_i} p^k$. Let $\mu \in \mathbb{F}_{p^n} \setminus \{0\}$. Let $f : \mathbb{F}_{p^n}^r \to \mathbb{F}_p$ be given by*

$$f(x_1, \ldots, x_r) = \mathrm{Tr}\left(\mu \sum_{I \subseteq [r], |I|=t} \prod_{i \in I} x_i^{\alpha_i} \right).$$

Let $\mathcal{A} \subseteq \mathbb{F}_{p^n}^r$ be any $\mathbb{F}_p$-affine space with $\dim(\mathcal{A}) > \frac{nr}{t} + nt + r(r+1)$. Then $|f(\mathcal{A})| > 1$.

Before proving this theorem let us show how it implies Theorem 2.2.

Proof of Theorem 2.2: For our selection of parameters n, t, r we notice the assumptions of Theorem 7.3 hold. Indeed, by Bertrand's postulate we can bound n from above by $4m^{3/5}$, hence $r \geq m^{2/5}/4$. Notice that for our setting of parameters

$$r(r+1)/t \leq \sqrt{r}(r+1) \leq m^{3/5} < n$$

and if $d > 6m^{4/5}$ then we have

$$\frac{nr}{t} + nt + r(r+1) \leq \frac{4}{\sqrt{2}} m^{4/5} + \frac{4}{\sqrt{2}} m^{4/5} + \frac{1}{4} m^{4/5} + o(m^{4/5}) < d.$$

Thus, the function f in Theorem 7.3 has the property that for any $\mathcal{A}$ with $\dim(\mathcal{A}) > 6m^{4/5}$, we have $|f(\mathcal{A})| > 1$. Finally notice that Proposition 3.3 implies that f as defined in Theorem 7.3 is identical to f defined in Theorem 2.2, up to renaming of the variables x_i. This completes the proof. □

Proof of Theorem 7.3: Our proof strategy is again as outlined in the introduction. Our first goal is to find a polynomial mapping $H : \mathbb{F}_{p^n}^r \to \mathbb{F}_{p^n}^r$ such that $H(\mathbb{F}_{p^n}^r) = \mathcal{A}$. We will then show that the composed function $f \circ H$ is a non-constant map, by showing that in its representation as a polynomial, there is a positive degree monomial with a nonzero coefficient.

To define the mapping H, we first decompose the affine space $\mathcal{A}$ using Lemma 7.1. Let $\gamma \in \mathcal{A}$. Then by that lemma, we may find a collection of $\mathbb{F}_p$-linear subspaces $Y_1, \ldots, Y_r \subseteq \mathbb{F}_{p^n}$ and linear maps $\sigma_{ij} : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ for $i, j \in [r]$ with $i < j$ such that:

$$\mathcal{A} = \{(x_1, \ldots, x_r) \mid \exists y_i \in Y_i \text{ such that } x_i = \gamma_i + y_i + \sum_{j<i} \sigma_{ij}(y_j)\}$$

and $\dim \mathcal{A} = \sum_{i \in [r]} \dim Y_i$.

Let $Q_i(X) \in \mathbb{F}_{p^n}[X]$ be the image-subspace polynomial of Y_i. Let $Q_{ij}(X)$ be the linearized polynomial (guaranteed to exist by Lemma 3.7) such that $Q_{ij}(x) = \sigma_{ij}(Q_i(x))$ for each $x \in \mathbb{F}_p^n$. Let $R_i(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$ be the polynomial $Q_i(X_i) + \sum_{j<i} Q_{ij}(X_j) + \gamma_i$. Then by the above comments, the image of the function H mapping $x = (x_1, \ldots, x_r) \in \mathbb{F}_{p^n}^r$ to $(R_1(x), \ldots, R_r(x))$ is precisely $\mathcal{A}$.

Now let $h(X_1, \ldots, X_r) \in \mathbb{F}_{p^n}[X_1, \ldots, X_r]$ be the polynomial representing $f \circ H$, namely

$$h(X_1, \ldots, X_r) = f(R_1(X_1, \ldots, X_r), \ldots, R_r(X_1, \ldots, X_r)) \mod \langle X_i^q - X_i \rangle_{i \in [r]} \tag{24}$$

$$= \sum_{I \subseteq [r], |I| = t} \mathrm{Tr}\left(\mu \prod_{i \in I} R_i(X_1, \ldots, X_r)^{\alpha_i}\right) \mod \langle X_i^q - X_i \rangle_{i \in [r]} \tag{25}$$

By Proposition 3.1, we have $h(\mathbb{F}_{p^n}^r) = f(\mathcal{A})$.

Therefore, to show that $|f(\mathcal{A})| > 1$, it suffices to show that $|h(\mathbb{F}_{p^n}^r)| > 1$. We do this by showing that $h(X_1, \ldots, X_r)$ has a monomial of positive degree with a nonzero coefficient and invoking Proposition 3.1.

To find this monomial, we consider the representation (25) of the polynomial $h(X_1, \ldots, X_n)$. We will first find a set $I_0 \subseteq [r]$, with $|I_0| = t$, of "blocks with high entropy". Then via Theorem 6.1, we will argue that the summand in (25) corresponding to I_0 is a nonzero polynomial, with certain monomial $\mathcal{M}$ having a nonzero coefficient. We will then show that no other summand in the sum (25) can have the monomial $\mathcal{M}$ with a nonzero coefficient, thus establishing that $\mathcal{M}$ appears in h with a nonzero coefficient, as desired.

We proceed with implementing this plan. Let $d_i = \dim(Y_i)$ and let $d = \dim(\mathcal{A}) > \frac{nr}{t} + nt + r(r+1)$. We have $\sum_i d_i = d$. Let $S = \{i \in [r] \mid d_i > r + 1\}$. Then we get

- $|S| \geq t$ (since each $d_i \leq n$ and $\sum d_i > nt + r(r+1)$).
- $\sum_{i \in S} d_i \geq \sum_{i \in [r]} (d_i - r - 1) = d - r(r+1) \geq nr/t + nt$.

Thus there exists $I_0 \subseteq S$ (and hence each $i \in I_0$ has $d_i > r+1$) with $|I_0| = t$ such that

$$\sum_{i \in I_0} (d_i - (r+1)) \geq \left(\sum_{i \in S} d_i \right) \frac{t}{r} - (r+1)t \geq n + nt^2/r - (r+1)t \geq n, \tag{26}$$

where the last inequality used the hypothesis that $n \geq r(r+1)/t$.

Let us focus on the term

$$g(X_1, \ldots, X_r) = \mathrm{Tr}\left(\mu \prod_{i \in I_0} R_i(X_1, \ldots, X_r)^{\alpha_i} \right) \bmod \langle X_i^q - X_i \rangle_{i \in [r]}$$

in the representation (25) of the polynomial $h(X_1, \ldots, X_r)$.

Putting $g'(X_1, \ldots, X_r) = \mathrm{Tr}\left(\mu \prod_{i \in I_0} Q_i(X_i)^{\alpha_i}\right) \mod \langle X_i^q - X_i \rangle_{i \in [r]}$ and noting that each $e_i + 1 \leq r+1$, Equation (26) and Theorem 6.1 imply that there is a monomial $\mathcal{M} = \prod_{i \in I_0} X_i^{\beta_i}$ with $\mathrm{wt}_p(\beta_i) = \mathrm{wt}_p(\alpha_i) = e_i + 1$, which has a nonzero coefficient in g'. Lemma 7.2 now implies that the coefficient of $\mathcal{M}$ in g is exactly the same as the coefficient of $\mathcal{M}$ in g', and hence nonzero.

We now show that in the representation (25) of the polynomial $h(X_1, \ldots, X_r)$, no summand other than g can have a nonzero coefficient for the monomial $\mathcal{M}$. First notice that each $R_i(X_1, \ldots, X_r)$ is a polynomial only in the variables $X_1, X_2, \ldots, X_i$, and is a sum of monomials of the form $aX_k^{p^b}$ plus possibly a constant term (i.e., monomials of total $\mathbb{F}_p$-degree at most 1).

Let $J \subseteq [r]$ with $|J| = t$, and consider the expression $\mathrm{Tr}(\mu \prod_{j \in J} R_j(X_1, \ldots, X_r)^{\alpha_j}) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle$. By definition, it equals:

$$\mathrm{Tr}\left(\mu \prod_{j \in J} \prod_{l=0}^{e_j} R_j(X_1, \ldots, X_j)^{p^l} \right) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle.$$

Suppose the monomial $\mathcal{M}$ appeared in the above polynomial with a nonzero coefficient. Then, expanding the trace map, there is some $w \in [n-1]$ such that $\mathcal{M}$ appears in

$$\prod_{j \in J} \prod_{l=0}^{e_j} R_j(X_1, \ldots, X_j)^{p^{l+w}} \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle$$

with a nonzero coefficient. Letting $R_{jl} = R_j^{p^{l+w}}$, we may rewrite the last polynomial as

$$\prod_{j \in J} \prod_{l=0}^{r-j+1} R_{jl}(X_1, \ldots, X_j) \mod \langle (X_i^{p^n} - X_i)_{i \in [r]} \rangle,$$

where each R_{jl} is a sum of monomials of total $\mathbb{F}_p$-degree at most 1. Each monomial $\mathcal{M}'$ that appears in this product is obtained by choosing, for each $j \in J$ and $l \in [0, r-j+1]$, a monomial from $R_{jl}(X_1, \ldots, X_j)$, and multiplying all these monomials out. Since we know that $\mathcal{M}$ appears in this product, let us focus on the choices made in order for $\mathcal{M}$ to appear. We set $\lambda_j(l) = i$ if for (j, l) we chose a monomial from $R_{jl}(X_1, \ldots, X_j)$ whose variable is indexed by i (i.e., we chose some aX_i^b).

Observe that the $\mathbb{F}_p$-degree in X_i of $\mathcal{M}$ is at most the number of (j,l) pairs for which $\lambda_j(l) = i$ (which may be compactly written as $\sum_{j\in J} |\lambda_j^{-1}(i)|$). However, we know that for any $i \in I_0$, the $\mathbb{F}_p$-degree of $\mathcal{M}$ in the variable X_i is $e_i + 1$ (which equals $r + 2 - i$). The following combinatorial claim (whose proof appears next) now shows that J must be equal to I_0.

Claim 7.4 *Let $I_0 \subseteq [r]$ with $|I_0| = t$. Suppose $J \subseteq [r]$ with $|J| = t$, and that there exist functions $\lambda_j : \{0, 1, \ldots, r + 1 - j\} \to \{1, \ldots, j\}$ for $j \in J$, with the property that for each $i \in I_0$,*

$$\sum_{j\in J} |\lambda_j^{-1}(i)| \geq r + 2 - i. \tag{27}$$

Then $J = I_0$.

Therefore, we have shown that there is precisely one summand, namely the one corresponding to I_0, in the representation (25) of $h(X_1, \ldots, X_r)$ that has a nonzero coefficient for the monomial $\mathcal{M}$. Thus $\mathcal{M}$ appears in h with a nonzero coefficient, and thus $|h(\mathbb{F}_{p^n}^r)| > 1$, as desired. □

Proof of Claim 7.4: Note that for any $j \in J$ we have

$$\sum_{i\in I_0} |\lambda_j^{-1}(i)| \leq |\{0, 1, \ldots, r + 1 - j\}| = r + 2 - j.$$

Thus

$$\sum_{i\in I_0} (r + 2 - i) \leq \sum_{j\in J} \sum_{i\in I_0} |\lambda_j^{-1}(i)| \leq \sum_{j\in J} (r + 2 - j).$$

As $|I_0| = |J|$, we have $\sum_{i\in I_0} i \geq \sum_{j\in J} j$.

Consider now the expression $\sum_{j\in J} \sum_{l=0}^{r+1-j} (j - \lambda_j(l))$, which is ≥ 0, because $\lambda_j(l) \leq j$. Thus,

$$\sum_{j\in J} (r + 2 - j) \cdot j \geq \sum_{j\in J} \sum_{l=0}^{r+1-j} \lambda_j(l) \geq \sum_{i\in I_0} (r + 2 - i) \cdot i.$$

The last inequality follows from the assumption (27). Rearranging, we get,

$$\sum_{i\in I_0} i^2 - \sum_{j\in J} j^2 \geq (r + 2) \cdot (\sum_{i\in I_0} i - \sum_{j\in J} j) \geq 0.$$

Thus $\sum_{i\in I_0} i^2 \geq \sum_{j\in J} j^2$.

For general k, considering the expression $\sum_{j\in J} \sum_{l=0}^{r+1-j} (j^k - \lambda_j(l)^k)$, which is nonnegative, we get

$$\sum_{i\in I_0} i^{k+1} - \sum_{j\in J} j^{k+1} \geq (r + 2) \cdot (\sum_{i\in I_0} i^k - \sum_{j\in J} j^k),$$

which by induction on k is ≥ 0. Thus for all k,

$$\sum_{i\in I_0} i^k \geq \sum_{j\in J} j^k \tag{28}$$

This implies that $i_1 := \max(I_0) \geq \max(J) =: j_1$. However, $i_1 \leq j_1$, otherwise $\lambda_j^{-1}(i_1) = \emptyset$ for each j. Thus $i_1 = j_1$. This forces $\lambda_{j_1}(l) = i_1$ for each l.

Taking this information back to Equation (28), we now see that the second-largest element i_2 of $I_0 \geq$ the second-largest element j_2 of J. But we must have $i_2 \leq j_2$, otherwise $\lambda_j^{-1}(i_2) = \emptyset$ for all j (recall that $\lambda_{j_1}(l) = i_1$ for each l, and there is no other j for which $i_1 \leq j$). Thus $i_2 = j_2$.

Inducting now on s, and arguing about the s-th largest element of I_0 and J, we get that $I_0 = J$.

□

Acknowledgements This work was initiated while both authors visited Alex Samorodnitsky at the Hebrew University, Jerusalem. He took an active part in the initial stages of this research yet declined to be a co-author. We are grateful to Alex for his hospitality and support and for many enlightening discussions. We thank Jaikumar Radhakrishnan for helpful discussions.

References

Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. In *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, Washington, DC, USA, 2004. IEEE Computer Society. ISBN 0-7695-2228-9. doi: http://dx.doi.org/10.1109/FOCS.2004.29.

Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, ramsey graphs, dispersers, and extractors. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2005.

Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2-source dispersers for sub-polynomial entropy and ramsey graphs beating the frankl-wilson construction. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 671–680, New York, NY, USA, 2006. ACM. ISBN 1-59593-134-1. doi: http://doi.acm.org/10.1145/1132516.1132611.

Eli Ben-Sasson and Madhu Sudan. Simple PCPs with poly-log rate and query complexity. In *STOC '05: Proceedings of the 37th annual ACM Symposium on Theory of Computing*, pages 266–275, New York, NY, USA, 2005. ACM Press. ISBN 1-58113-960-8. doi: http://doi.acm.org/10.1145/1060590.1060631.

Eli Ben-Sasson, S. Hoory, E. Rozenman, and S. Vadhan. Extractors for affine sources, unpublished manuscript. 2001.

Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Robust PCPs of proximity, shorter PCPs and applications to coding. In ACM, editor, *Proceedings of the 36th Annual ACM Symposium on the Theory of Computing: Chicago, Illinois, USA, June 13–15, 2004*, pages 1–10, pub-ACM:adr, 2004. pub-ACM. ISBN 1-58113-852-0.

Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil P. Vadhan. Short PCPs verifiable in polylogarithmic time. In *IEEE Conference on Computational Complexity*, pages 120–134, 2005. URL `http://dx.doi.org/10.1109/CCC.2005.27`.

Eli Ben-Sasson, Swastik Kopparty, and Jaikumar Radhakrishnan. Subspace polynomials and list decoding of reed-solomon codes. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, 2006.

Elwyn R. Berlekamp. *Algebraic Coding Theory*. Mc Graw-Hill, revised 1984 edition, 1968.

Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. In *FOCS*, pages 41–51. IEEE Computer Society, 2007.

J. Bourgain. On the construction of affine extractors. *Geometric and Functional Analysis*, 17(1): 33–57, 2007.

J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004. ISSN 1016-443X.

J. Bourgain, A. A. Glibichuk, and S. V. Konyagin. Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. London Math. Soc. (2)*, 73(2):380–398, 2006. ISSN 0024-6107.

Benny Chor, Oded Goldreich, Johan Håstad, Joel Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem of t-resilient functions (preliminary version). In *FOCS*, pages 396–407. IEEE, 1985.

Matt DeVos and Ariel Gabizon. Simple affine extractors using dimension expansion. *Electronic Colloquium on Computational Complexity (ECCC)*, (63), 2009. URL `http://eccc.hpi-web.de/eccc-reports/2009/TR09-063/index.html`.

Gabizon, Raz, and Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SICOMP: SIAM Journal on Computing*, 36, 2006.

Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. In *FOCS '05: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 407–418, Washington, DC, USA, 2005. IEEE Computer Society.

Elad Haramaty and Amir Shpilka. On the structure of cubic and quartic polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, (80), 2009. URL `http://eccc.hpi-web.de/eccc-reports/2009/TR09-080/index.html`.

Xiang-Dong Hou, Ka Hin Leung, and Qing Xiang. A generalization of an addition theorem of kneser. *Journal of Number Theory*, 97(1):1 – 9, 2002. ISSN 0022-314X. doi: DOI:10.1006/jnth.2002.2793. URL `http://www.sciencedirect.com/science/article/B6WKD-478RYHN-1/2/cba94db63fca2ee1188ab30fb6b102e4`.

Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM J. Comput.*, 36(5):1231–1247, 2007.

Jesse Kamp, Anup Rao, Salil P. Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. In *STOC*, pages 691–700, 2006.

Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1997. ISBN 0-521-39231-4. With a foreword by P. M. Cohn.

Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Comb. Theory, Ser. A*, 71(1):168–172, 1995.

O. Ore. On a special class of polynomials. *Trans. Amer. Math. Soc.*, 35(3):559–584, 1933.

O. Ore. Contributions to the theory of finite fields. *Trans. Amer. Math. Soc.*, 36(2):243–274, 1934.

Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

Terence Tao and Van H. Vu. *Additive Combinatorics*. Number 105 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, UK, first edition, 2006. ISBN 13 978-0-521-85386-6. doi: 10.2277/0521853869.

Luca Trevisan and Salil P. Vadhan. Extracting randomness from samplable distributions. In *FOCS*, pages 32–42, 2000.

David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 681–690, New York, NY, USA, 2006. ACM. ISBN 1-59593-134-1. doi: http://doi.acm.org/10.1145/1132516.1132612.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de