

Affine extractors over prime fields

Amir Yehudayoff*

Abstract

An affine extractor is a map that is balanced on every affine subspace of large enough dimension. We construct explicit affine extractors over \mathbb{F}^n , \mathbb{F} a prime field, that is balanced on every affine subspace of dimension at least δn , for any constant $0 < \delta \leq 1$ (the dependency on δ is better than was previously known).

1 Introduction

An *affine extractor* is a map EXT from the vector space \mathbb{F}^n to the field \mathbb{F} that is balanced on every affine subspace of \mathbb{F}^n of large enough dimension. Formally, for every linear subspace X of \mathbb{F}^n with dimension at least δn , $0 < \delta \leq 1$, and for every fixed vector $\xi \in \mathbb{F}^n$, the statistical distance between the distribution $\text{EXT}(x + \xi)$, for x chosen uniformly at random from X , and the uniform distribution on \mathbb{F} is at most ε .

Affine extractors have been studied in several papers before. The first to consider affine extractors were Gabizon and Raz [4]. They constructed affine extractors over large fields, for arbitrary δ and polynomially small ε . Later, Bourgain [1] constructed an affine extractor over \mathbb{F}_2 , the field with 2 elements, that works for an arbitrary constant δ and exponentially small ε . Rao constructed low-weight affine extractors, that are guaranteed to work as long as the subspace X has a basis of low-weight vectors [5]. Rao's construction works for δ as small as $\text{polylog}(n)/n$.

*School of Mathematics, Institute for Advanced Study, Princeton NJ.
Email: amir.yehudayoff@gmail.com. Research partially supported by NSF grants CCF-0832797 and DMS-0835373.

We construct an explicit affine extractor that works for an arbitrary constant δ and exponentially small ε (see Theorem 1 below for exact parameters). This affine extractor works over arbitrary prime fields (a slight generalization of Bourgain’s construction [1] seems to work for arbitrary prime fields as well).

The construction here is similar to that of Bourgain [1]. As in [1], we think of the input x as consisting of $r = r(\delta)$ blocks, each of size n/r , that is, $x = (x_1, \dots, x_r)$ with x_i a vector of length n/r , $1 \leq i \leq r$. The extractor is defined to be a carefully chosen polynomials in x_1, \dots, x_r . Our choice of a polynomial is different than the choice of [1], and it has a few advantages. One advantage is that the proof is simpler, for example, we use simpler structural properties of linear subspaces. A second advantage is that it enables us to get a better dependency on the parameters: our estimates are trivial for δ roughly $1/\sqrt{\log \log n}$, whereas previous estimates are trivial for δ roughly $1/\sqrt{\log \log \log n}$. This improvement is possible, since our choice of a polynomial allows us to carry the main part of the proof, which is basically an exponential sum estimate, in a slightly different way.

For simplicity, we focus on extracting one almost uniform field element. Standard arguments (see, e.g., Section 10 in [6]) show that Theorem 1 below implies that we can, in fact, extract a linear number of almost uniform field elements.

2 An affine extractor

Denote $[n] = \{1, \dots, n\}$. For two primes $p, m \in \mathbb{N}$, denote by \mathbb{F}_{p^m} the field with p^m elements. Denote $\mathbb{F} = \mathbb{F}_p$, the field with p elements. Think of the elements of \mathbb{F}^m both as vectors in the vector space \mathbb{F}^m and as field elements in the field \mathbb{F}_{p^m} . Let $n = rm$, for $r \in \mathbb{N}$. By X we denote a linear subspace of \mathbb{F}^n of dimension δn , for some $0 < \delta \leq 1$. Think of every x in \mathbb{F}^n also as in $(\mathbb{F}_{p^m})^r$, that is, $x = (x_1, \dots, x_r)$ with $x_1, \dots, x_r \in \mathbb{F}_{p^m}$.

2.1 The extractor

By $\Lambda \subset \mathbb{N}$ we denote the set of $\lambda \in \mathbb{N}$ such that $\lambda = \sum_{i \in \mathbb{N}} a_i p^i$ with $a_i \in \{0, 1\}$, for every $i \in \mathbb{N}$. We abuse notation and denote $\langle \lambda \rangle = \{i \in \mathbb{N} : a_i = 1\}$ and $|\lambda| = |\langle \lambda \rangle|$.

Let \mathcal{I}_s be the family of all subsets of $[r]$ of size s . For $I = \{i_1 < i_2 < \dots < i_s\} \in \mathcal{I}_s$, define f_I as the following monomial in the variables x_{i_1}, \dots, x_{i_s} :

$$f_I = x_{i_1}^{\lambda_{I,1}} x_{i_2}^{\lambda_{I,2}} \dots x_{i_s}^{\lambda_{I,s}}$$

where

$$\lambda_{I,j} = \sum_{1 \leq \ell \leq s^{2(r-i_j)+j}} p^\ell \in \Lambda,$$

for every $j \in [s]$. Define

$$F = \sum_{I \in \mathcal{I}_s} f_I.$$

The following theorem implies that the first coordinate of F , for example, is an affine extractor with an arbitrary constant δ and exponentially small ε (for more information about this implication see, e.g., Section 10 in [6]).

Theorem 1. *There exists a constant $C > 1$ such that the following holds. Let $0 < \delta \leq 1$, and let $n \in \mathbb{N}$ be large enough. Let $\xi \in \mathbb{F}^n$, let X be a linear subspace of \mathbb{F}^n with dimension at least δn , and let ψ be a nontrivial additive character of \mathbb{F}_{p^m} . Then,*

$$|\mathbb{E}_{x \in X} \psi[F(x + \xi)]| < Cp^{-2^{-\delta-C/\delta^2} n}, \quad (2.1)$$

where $\mathbb{E}_{x \in X}$ means expectation with respect to a uniformly chosen element x in X .

Before proving Theorem 1, we need a few preliminaries.

2.2 The structure of subspaces

Let A be a $\delta n \times n$ matrix whose rows form a basis for X . Assume that A has row echelon form, that is, the pivot of each row is strictly to the right of the pivot of the row above it (the pivot of a vector is the position of the leftmost nonzero entry in it). We can assume that A has this form, since otherwise we can transform A to such a form using Gaussian elimination without changing X .

For every $i \in [r]$, define W_i as the span of all the rows of A for which the pivot is between $(i-1)m+1$ and im . Every $x = (x_1, \dots, x_r)$ in X can be expressed uniquely

as $x = w_1 + w_2 + \cdots + w_r$ with $w_i \in W_i$, for every $i \in [r]$. Note that x_1, \dots, x_r are in \mathbb{F}^m , whereas w_1, \dots, w_r are in \mathbb{F}^n . Every x can be also expressed as

$$\begin{aligned} x_1 &= u_1 + y_1, \\ x_2 &= u_2 + y_2, \\ &\dots \\ x_r &= u_r + y_r, \end{aligned} \tag{2.2}$$

where $u_i = (w_1)_i + \cdots + (w_{i-1})_i$ and $y_i = (w_i)_i$, for every $i \in [r]$ (so $u_1 = 0$). The vector w_i can be linearly reconstructed from y_i , and so u_i is a linear function of y_1, \dots, y_{i-1} . Denote

$$Y_i = \{(w_i)_i \in \mathbb{F}^m : w_i \in W_i\}.$$

The dimension of Y_i is the same as the dimension of W_i .

The simple representation of X described above appears in [1] as well. This representation suffices for the proof of Theorem 1 presented here, whereas in [1] it is the starting point of a more elaborated structure.

2.3 Van der Corput differencing

As in [1], a key ingredient in proving Theorem 1 is Van der Corput differencing. For a map $f : \mathbb{F}_{p^m} \rightarrow \mathbb{F}_{p^m}$ and $z \in \mathbb{F}_{p^m}$, define the map $D_z[f] : \mathbb{F}_{p^m} \rightarrow \mathbb{F}_{p^m}$ as

$$D_z[f](x) = f(x + z) - f(x),$$

for every $x \in \mathbb{F}_{p^m}$. The map $D_z[f]$ is the derivative of f with respect to x in the direction z . For $z_1, \dots, z_t \in \mathbb{F}_{p^m}$, define

$$D_{z_1, \dots, z_t}[f] = D_{z_1}[D_{z_2}[\cdots [D_{z_t}[f]] \cdots]].$$

If f is linear in x , then $D_z[f]$ does not depend on x , for every z . For x^λ with $\lambda \in \Lambda$,

$$D_z[x^\lambda] = \prod_{j \in \langle \lambda \rangle} (x + z)^{p^j} - \prod_{j \in \langle \lambda \rangle} x^{p^j} = \prod_{j \in \langle \lambda \rangle} (x^{p^j} + z^{p^j}) - \prod_{j \in \langle \lambda \rangle} x^{p^j} = \sum_{\lambda' \in \Lambda} z^{\lambda - \lambda'} x^{\lambda'},$$

where the sum is over λ' such that $\langle \lambda' \rangle \subset \langle \lambda \rangle$ and $|\lambda'| < |\lambda|$. Iterating for $t = |\lambda| - 1$ times,

$$D_{z_1, \dots, z_t} [x^\lambda] = \sum_{\ell \in \langle \lambda \rangle} P_\ell(z_1, \dots, z_t) x^{p^\ell}, \quad (2.3)$$

for nonzero polynomials $P_\ell \in \mathbb{F}_{p^m}[z_1, \dots, z_t]$, $\ell \in \langle \lambda \rangle$. The total degree of P_ℓ is at most λ , for every $\ell \in \langle \lambda \rangle$.

We say that the *linear-degree* of a variable x in a map f is t , if f is a product of t linear maps in x . For example, the linear-degree of x in x^λ is $|\lambda|$. If f has linear-degree less than t in x , then $D_{z_1, \dots, z_t}[f](x) = 0$.

Our goal is to upper bound,

$$|\mathbb{E}_{x \in X} \psi[F(x + \xi)]|. \quad (2.4)$$

Recall the representation of x in X given by (2.2). Differencing once with respect to y_1 gives (using the Cauchy-Schwarz inequality)

$$\begin{aligned} |(2.4)|^2 &= |\mathbb{E}_{y_i \in Y_i: i > 1} \mathbb{E}_{y_1 \in Y_1} \psi[F(x + \xi)]|^2 \\ &\leq \mathbb{E}_{y_i \in Y_i: i > 1} |\mathbb{E}_{y_1 \in Y_1} \psi[F(x + \xi)]|^2 \\ &= \mathbb{E}_{y_i \in Y_i: i > 1} \mathbb{E}_{y_1, z \in Y_1} |\psi[D_z[F](x + \xi)]| \\ &= \mathbb{E}_{z \in Y_1} \mathbb{E}_{y_i \in Y_i} \psi[D_z[F](x + \xi)]. \end{aligned}$$

Differencing t times gives

$$\begin{aligned} |(2.4)|^{2^t} &\leq \mathbb{E}_{z_1, \dots, z_t \in Y_1} \mathbb{E}_{y_i \in Y_i} \psi[D_{z_1, \dots, z_t}[F](x + \xi)] \\ &= \mathbb{E}_{1, t} \mathbb{E}_{y_i \in Y_i} \psi[D_{1, t}[F](x + \xi)]. \end{aligned}$$

The last equality defines notation: $\mathbb{E}_{1, t}$ means expectation over t independent copies of a uniformly distributed element of Y_1 , and $D_{1, t}$ means differencing t times with respect to y_1 (using the t variables implicitly given by $\mathbb{E}_{1, t}$).

2.4 Reducing the exponential sum

In this section we reduce the estimation of (2.4) to an estimation of a more structured exponential sum. We use induction to prove this reduction, so we need to slightly generalize the definitions given in Section 2.1.

Let $1 \leq s \leq r \leq b$ and a in \mathbb{N} be such that $s \leq a$, and let \mathcal{I}_s be the family of all subsets of $[r]$ of size s . For $I = \{i_1 < \dots < i_s\} \in \mathcal{I}_s$, define f_I as

$$f_I = x_{i_1}^{\lambda_{I,1}} x_{i_2}^{\lambda_{I,2}} \dots x_{i_s}^{\lambda_{I,s}}$$

with

$$\lambda_{I,j} = \sum_{1 \leq \ell \leq a^{2(b-i_j)+j}} p^\ell \in \Lambda,$$

for every $j \in [s]$ (this definition of $\lambda_{I,j}$ does not depend on r and s). Define

$$F = \sum_{I \in \mathcal{I}_s} f_I \tag{2.5}$$

(when $s = a$ and $r = b$, this is the same definition of F as in Section 2.1).

The following proposition shows that the definition of F allows us to assume that the entropy of the subspace appears in s independent blocks.

Proposition 2. *Let $1 \leq s \leq r \leq b$ and a in \mathbb{N} be such that $s \leq a$ and $a \geq 3$, and let F be as defined in (2.5) above. Let $\{r_1 < \dots < r_s\}$ be a subset of $[r]$ of size s , and denote*

$$t_j = a^{2(b-r_j)+j} - 1, \quad j \in [s].$$

Then, there exist nonzero polynomials $P_\ell^{(j)}$, $j \in [s]$ and $1 \leq \ell \leq t_j + 1$, such that the following holds. For every linear space $X \subset \mathbb{F}^n$ (recall (2.2)), for every $\xi \in \mathbb{F}^n$, and for every additive character ψ of \mathbb{F}_{p^m} ,

$$|\mathbb{E}_{y_i \in Y_i} \psi[F(\mathbf{x})]|^{2^{t_1+\dots+t_s+s}} \leq \mathbb{E}_{r_1, t_1} \dots \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i} \psi[Z_1 \dots Z_s]$$

with $\mathbf{x} = x + \xi$, where for every $j \in [s]$,

$$Z_j = \sum_{1 \leq \ell \leq t_j+1} P_\ell^{(j)} y_{r_j}^\ell,$$

Each of the polynomials $P_\ell^{(j)}$ is in the variables $z_1^{(j)}, \dots, z_{t_j}^{(j)}$ that are implicitly given by \mathbb{E}_{r_j, t_j} . The total degree of $P_\ell^{(j)}$ is at most p^{t_j+2} .

The random variables Z_1, \dots, Z_s are independent. In the proof of Theorem 1 below, we show that we can choose $\{r_1 < \dots < r_s\}$ so that each of Z_1, \dots, Z_s is uniformly distributed over a subspace of large dimension. So, in order to establish Theorem 1, we can use an exponential sum estimate of Bourgain [2] and of Bourgain, Glibichuk and Konyagin [3].

Proof. The proof of the proposition is by induction of s . We start with the induction step. Assume that the proposition holds for $s - 1 \geq 1$. The induction step is done by removing every f_I with $\max I \neq r_s$ from the sum. This is done by differencing t_s times with respect to y_{r_s} .

Let $I = \{i_1 < \dots < i_s\} \in \mathcal{I}_s$ be such that $\max I \neq r_s$. Think of f_I as a polynomial in y_{r_s} . Let $0 \leq k \leq s$ be the number of elements in I that are at least r_s . If $k = 0$, then, according to (2.2), f_I does not depend on y_{r_s} . If $k = 1$, since $\max I \neq r_s$, the linear-degree of y_{r_s} in f_I is

$$|\lambda_{I,s}| = a^{2(b-i_s)+s} < a^{2(b-r_s)+s} - 1 = t_s$$

(recall $a \geq 3$). If $k \geq 2$, the linear-degree of y_{r_s} in f_I is

$$\sum_{\ell=0}^{k-1} |\lambda_{I,s-(k-1)+\ell}| \leq \sum_{\ell=0}^{k-1} a^{2(b-(r_s+\ell))+s-(k-1)+\ell} \leq a^{2(b-r_s)+s} \sum_{\ell=0}^{k-1} a^{-\ell-1} < t_s$$

To conclude, if $\max I \neq r_s$, then the linear-degree of y_{r_s} in f_I is less than t_s , which implies that

$$D_{r_s,t_s}[f_I] = 0.$$

On the other hand, if $\max I = r_s$, then

$$f_I(\mathbf{x}) = \mathbf{x}_{i_1}^{\lambda_{I,1}} \dots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} \mathbf{x}_{r_s}^{\sum_{1 \leq \ell \leq t_s+1} p^\ell},$$

and by (2.3),

$$D_{r_s,t_s}[f_I](\mathbf{x}) = \mathbf{x}_{i_1}^{\lambda_{I,1}} \dots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} \sum_{1 \leq \ell \leq t_s+1} P_\ell^{(s)} \mathbf{x}_{r_s}^{p^\ell},$$

where $P_\ell^{(s)}$ is a nonzero polynomial of total degree at most $\sum_{1 \leq \ell \leq t_s+1} p^\ell \leq p^{t_s+2}$ in the variables that are implicitly given by D_{r_s,t_s} , for every $1 \leq \ell \leq t_s + 1$. By (2.2),

$$D_{r_s,t_s}[f_I](\mathbf{x}) = Q_I(\mathbf{x}) + \mathbf{x}_{i_1}^{\lambda_{I,1}} \dots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} Z_s,$$

where

$$Q_I(\mathbf{x}) = \mathbf{x}_{i_1}^{\lambda_{I,1}} \dots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} \sum_{1 \leq \ell \leq t_s+1} P_\ell^{(s)} (u_{r_s} + \xi_{r_s})^{p^\ell}.$$

Since Q_I does not depend on y_{r_s} ,

$$\begin{aligned} |(2.4)|^{2^{t_s}} &\leq \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i} \psi \left[D_{r_s, t_s}[F] \right] \\ &= \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i: i \neq r_s} \psi \left[\sum_{I \in \mathcal{I}_s: \max I = r_s} Q_I \right] \\ &\quad \mathbb{E}_{y_{r_s} \in Y_{r_s}} \psi \left[Z_s \sum_{I \in \mathcal{I}_s: \max I = r_s} \mathbf{x}_{i_1}^{\lambda_{I,1}} \cdots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} \right], \end{aligned}$$

and using the Cauchy-Schwarz inequality,

$$\begin{aligned} |(2.4)|^{2^{t_s+1}} &\leq \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i: i \neq r_s} \left| \mathbb{E}_{y_{r_s} \in Y_{r_s}} \psi \left[Z_s \sum_{I \in \mathcal{I}_s: \max I = r_s} \mathbf{x}_{i_1}^{\lambda_{I,1}} \cdots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}} \right] \right|^2 \\ &= \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i} \psi \left[Z_s \sum_{I \in \mathcal{I}_{s-1}} f_I^{(s-1)} \right], \end{aligned} \quad (2.6)$$

where \mathcal{I}_{s-1} is the family of all subsets of $[r_s - 1]$ of size $s - 1$, and

$$f_I^{(s-1)}(\mathbf{x}) = \mathbf{x}_{i_1}^{\lambda_{I,1}} \cdots \mathbf{x}_{i_{s-1}}^{\lambda_{I,s-1}}$$

(we abuse notation and denote by $\lambda_{I,j}$ what was previously denoted $\lambda_{I \cup \{r_s\}, j}$).

Fix $z_1^{(s)}, \dots, z_{t_s}^{(s)} \in Y_{r_s}$ and $y_i \in Y_i$, $i \geq r_s$, and consider

$$\mathbb{E}_{y_i \in Y_i: i < r_s} \psi \left[Z_s \sum_{I \in \mathcal{I}_{s-1}} f_I^{(s-1)} \right]. \quad (2.7)$$

By the induction hypothesis with $s-1 \leq r_s-1 \leq b$, with $s-1 \leq a$, with $\{r_1 < \cdots < r_{s-1}\}$, and with the additive character $\psi'(\alpha) = \psi(Z_s \alpha)$, $\alpha \in \mathbb{F}_{p^m}$,

$$\begin{aligned} |(2.7)|^{2^{t_1+\cdots+t_{s-1}+s-1}} &\leq \mathbb{E}_{r_1, t_1} \cdots \mathbb{E}_{r_{s-1}, t_{s-1}} \mathbb{E}_{y_i \in Y_i: i < r_s} \psi' [Z_1 \cdots Z_{s-1}] \\ &= \mathbb{E}_{r_1, t_1} \cdots \mathbb{E}_{r_{s-1}, t_{s-1}} \mathbb{E}_{y_i \in Y_i: i < r_s} \psi [Z_1 \cdots Z_s]. \end{aligned}$$

By (2.6),

$$\begin{aligned} \left| |(2.4)|^{2^{t_s+1}} \right|^{2^{t_1+\cdots+t_{s-1}+s-1}} &\leq \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i: i \geq r_s} \left| |(2.7)|^{2^{t_1+\cdots+t_{s-1}+s-1}} \right| \\ &\leq \mathbb{E}_{r_1, t_1} \cdots \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i} \psi [Z_1 \cdots Z_s]. \end{aligned}$$

This proves the induction step. It remains to prove the induction base. The argument above shows that, in the case $s = 1$ (see (2.6)),

$$|(2.4)|^{2^{t_s+1}} \leq \mathbb{E}_{r_s, t_s} \mathbb{E}_{y_i \in Y_i} \psi [Z_s],$$

as needed. \square

2.5 Estimating the exponential sum

Before proving Theorem 1, we state two known results.

The following lemma is known as the Schwartz-Zippel lemma.

Lemma 3. *Let $P \in \mathbb{F}_{p^m}[z_1, \dots, z_t]$ be a nonzero polynomial with total degree at most d . Then, for every set $S \subset \mathbb{F}_{p^m}$,*

$$|S|^{-t} |\{\sigma \in S^t : P(\sigma) = 0\}| \leq d |S|^{-1}.$$

The following exponential sum estimate is due to Bourgain [2] (see also [3]).

Theorem 4. *There exists a constant $C > 1$ such that the following holds. Let $0 < \delta < 1$ and $s \in \mathbb{N}$ with $s > C/\delta$. Let p be a prime, let m be a sufficiently large prime, and let $A_1, \dots, A_s \in \mathbb{F}_{p^m}$ satisfy $|A_i| > p^{\delta m}$, for every $i \in [s]$. Then,*

$$\left| \mathbb{E}_{a_1 \in A_1, \dots, a_s \in A_s} \psi(a_1 \cdots a_s) \right| < p^{-\delta' m}$$

with $\delta' > C^{-s}$.

We are now ready for the proof of our main theorem.

Proof of Theorem 1. We start by finding the coordinates of X that have high dimension. For $i \in [r]$, let $0 \leq \delta_i \leq 1$ be such that $\dim Y_i = \delta_i m$. Thus, $\delta_1 + \dots + \delta_r = \delta r$. Let $\mathcal{R}' = \{i \in [r] : \delta_i \geq \delta/2\}$, and let $s = \delta r/2$ (assume that $s \geq 3$). Since $|\mathcal{R}'| > s$, denote

$$\mathcal{R} = \{r_1 < \dots < r_s\} \subseteq \mathcal{R}',$$

the first s elements of \mathcal{R}' .

Let $P_\ell^{(j)}$, $j \in [s]$ and $1 \leq \ell \leq t_j + 1$, be the polynomials of total degree at most p^{t_j+2} given by Proposition 2 with $s = a$, $r = b$ and \mathcal{R} . Recall

$$Z_j = \sum_{1 \leq \ell \leq t_j+1} P_\ell^{(j)} y_{r_j}^{p^\ell},$$

$j \in [s]$, defined in Proposition 2.

Proposition 2 enables us to use Theorem 4 above. In order to use Theorem 4, we show that Z_1, \dots, Z_s have a lot of entropy, with high probability. This consists of two parts. The first part is to show that, for every $j \in [s]$, if $P_1^{(j)}$ is nonzero, then Z_j is uniformly distributed over a large subspace. In the second part we use Lemma 3 to argue that the probability that there exists $j \in [s]$ such that $P_1^{(j)} = 0$ is small.

Part one. Let $j \in [s]$, and fix $z_1^{(j)}, \dots, z_{t_j}^{(j)}$ in Y_{r_j} so that the field element $P_1^{(j)} \in \mathbb{F}_{p^m}$ is nonzero. Think of Z_j as a linear map of y_{r_j} . The dimension of $Z_j(Y_{r_j})$ is at least the dimension of Y_{r_j} minus the dimension of the kernel of Z_j . Lemma 3 with $t = 1$ implies that the size of the kernel of Z_j is at most p^{t_j+1} , and so the dimension of the kernel of Z_j is at most $t_j + 1 \leq s^{2r}$. The dimension of Y_{r_j} is $\delta_{r_j} m \geq \delta m/2$. Thus, as long as

$$s^{2r} \leq \delta m/4, \quad (2.8)$$

the dimension of $Z_j(Y_{r_j})$ is at least $\delta m/4$.

Theorem 4 implies that there exists a constant $C_2 > 1$ such that if $P_1^{(j)} \in \mathbb{F}_{p^m}$ is nonzero, for every $j \in [s]$, then

$$\mathbb{E}_{y_{r_i} \in Y_{r_i}; i \in [s]} \psi[Z_1 \cdots Z_s] \leq p^{-C_2^{-C_2/\delta} m},$$

as long as $s > C_2/\delta$.

Part two. For every $j \in [s]$, Lemma 3 with $t = t_j$ implies that

$$\mathbb{E}_{r_j, t_j} \mathbf{1}_{\{P_1^{(j)}(z_1^{(j)}, \dots, z_{t_j}^{(j)})=0\}} \leq p^{t_j+2} |Y_{r_j}|^{-1} \leq p^{s^{2r}} p^{-\delta m/2}.$$

Concluding. As long as

$$sp^{s^{2r}} \leq p^{\delta m/4}, \quad (2.9)$$

Proposition 2 together with part one and two above (using the union bound) imply that

$$|(2.4)|^{2^{t_1+\dots+t_s+s}} \leq sp^{s^{2r}} p^{-\delta m/2} + p^{-C_2^{-C_2/\delta} m} \leq C_3 p^{-C_3^{-C_3/\delta} m},$$

for some constant $C_3 > 1$. Thus,

$$|(2.4)| \leq \left(C_3 p^{-C_3^{-C_3/\delta} m} \right)^{2^{-t_1-\dots-t_s-s}} \leq C p^{-2^{-\delta-C/\delta^2} n},$$

for a constant $C > 1$. The theorem follows, since if conditions (2.8) and (2.9) do not hold, δ is small enough so that the statement of the theorem is trivial. \square

Acknowledgement. I would like to thank Ariel Gabizon for useful comments on an earlier version of this work.

References

- [1] J. Bourgain. On the construction of affine extractors. *Geometric And Functional Analysis* 17 (1), pages 33–57, 2007.
- [2] J. Bourgain. Multilinear exponential sum bounds with optimal entropy assignments. *Geometric And Functional Analysis* (to appear).
- [3] 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. *Journal of the London Mathematical Society* 73(2), pages 380–398, 2006.
- [4] A. Gabizon and R. Raz. Deterministic extractors for affine sources over large fields. In *proceedings of the 46th FOCS*, pages 407–418, 2005.
- [5] A. Rao. Extractors for low-weight affine sources. Manuscript, 2007.
- [6] R. Raz and A. Yehudayoff. Multilinear formulas, maximal-partition discrepancy and mixed-sources extractors. In the *proceeding of 49th FOCS*, 2008.