From Affine to Two-Source Extractors via Approximate Duality

Eli Ben-Sasson* Noga Zewi*

May 18, 2011

Abstract

Two-source and affine extractors and dispersers are fundamental objects studied in the context of derandomization. This paper shows how to construct two-source extractors and dispersers for arbitrarily small min-entropy rate in a black-box manner given affine extractors with sufficiently good parameters. Our analysis relies on the study of approximate duality, a concept related to the polynomial Freiman-Ruzsa conjecture (PFR) from additive combinatorics.

Two black-box constructions of two-source extractors from affine ones are presented. Both constructions work for min-entropy rate $\rho < \frac{1}{2}$. One of them can potentially reach arbitrarily small min-entropy rate provided the the affine extractor used to construct it outputs, on affine sources of min-entropy rate $\frac{1}{2}$, a relatively large number of output bits, and has sufficiently small error.

Our results are obtained by first showing that each of our constructions yields a two-source disperser for a certain min-entropy rate $\rho < \frac{1}{2}$ and then using a general extractor-to-disperser reduction that applies to a large family of constructions. This reduction says that any two-source disperser for min-entropy rate ρ coming from this family is also a two-source extractor with constant error for min-entropy rate $\rho + \gamma$ for arbitrarily small $\gamma > 0$. We show that assuming the PFR conjecture, the error of this two-source extractor is exponentially small.

The extractor-to-disperser reduction arises from studying *approximate duality*, a notion related to additive combinatorics. The *duality measure* of two sets $A, B \subseteq \mathbb{F}_2^n$ aims to quantify how "close" these sets are to being dual and is defined as

$$\mu^{\perp}(A,B) = \left| \mathbb{E}_{a \in A, b \in B} \left[(-1)^{\sum_{i=1}^{n} a_i b_i} \right] \right|.$$

Notice that $\mu^{\perp}(A,B) = 1$ implies that A is contained in an affine shift of $B^{\perp}$ — the space dual to the $\mathbb{F}_2$-span of B. We study what can be said of A, B when their duality measure is a large constant strictly smaller than 1 and show that A, B contain subsets A', B' of nontrivial size for which $\mu^{\perp}(A', B') = 1$ and consequently A' is contained in an affine shift of $(B')^{\perp}$. This implies that our constructions are two-source extractors with constant error. Surprisingly, the PFR implies that such A', B' exist when A, B are large, even if the duality measure is exponentially small in n, and this implication leads to two-source extractors with exponentially small error.

*Department of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel. eli,nogaz@cs.technion.ac.il. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 240258.

Contents

1 Introduction

This paper shows a new connection between two objects that have been studied in recent years in the context of derandomization — *two-source* and *affine* extractors. First, we show two constructions that convert in a black-box manner any affine extractor for min-entropy rate below half into a two-source disperser for min-entropy rate below half. One of our constructions can reach arbitrarily small min-entropy rate as long as the affine extractor outputs sufficiently many output bits, and has sufficiently small error. Second, we show that the error of our dispersers is bounded by a constant on sources of slightly larger min-entropy rate using bounds on *approximate duality*, a new notion related to additive combinatorics to which a large portion of this paper is devoted. Third, we propose an approximate duality conjecture (ADC), show it is implied by the polynomial Freiman-Ruzsa conjecture (PFR) and implies a weak version of PFR, and use ADC to obtain exponentially small bounds on the error of our two-source extractors.

1.1 Extractors and dispersers for affine and two independent sources

Two-source extractors, dispersers and bipartite Ramsey graphs *Randomness extractors*, or, simply, extractors, deal with the task of extracting uniformly random bits from weak sources of randomness. (See the survey Shaltiel [2002] for an introduction to this topic.) The gold-standard measure for the randomness of a *source* X, i.e., a distribution over $\{0,1\}^n$, is its *min-entropy* which is defined to be the largest k such that for every $x \in \{0,1\}^n$ the probability assigned to x by X is at most 2^{-k}. (It is useful to think of X as uniformly distributed over an arbitrary subset of $\{0,1\}^n$ of size precisely 2^k.)

A function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$ is said to be a *two-source extractor* for *min-entropy* k with *error* ϵ if for every pair of independent sources X, Y that each have min-entropy at least k, the distance between the uniform distribution on m bits and the distribution $f(X,Y)$ is at most ϵ. A *two-source disperser* is a one-output-bit ($m = 1$) two-source extractor with a nontrivial (but possibly large) bound on the error. In other words, a two-source disperser for min-entropy k is a function f that is non constant on $S \times T$ for every pair of subsets S, T of size at least 2^k. Viewing f as the indicator function of the edge-set of a bipartite graph G_f on vertex sets of size 2^n, the graph G_f is known as a 2^k*-bipartite Ramsey graph* because the subgraph induced by any pair of sets of vertices of size at least 2^k is neither complete, nor empty.

Erdös [1947] inaugurated the use of the probabilistic method in combinatorics and showed among other things that a random function f is with high probability a two-source disperser [1] for min-entropy $\log n + O(1)$. Probabilistic arguments can also be used to show that such a function is with high probability a two-source extractor for the same min-entropy. However, up until a few years ago the best known construction of two-source dispersers (and extractors) required min-entropy at least $n/2$ [Chor and Goldreich, 1988]. This lower bound of half on the *min-entropy rate* — defined as the ratio between the min-entropy k and n — was first broken by Pudlák and Rödl [2004] for the case of two-source dispersers. They constructed two-source dispersers for min-entropy rate $\frac{1}{2} - o(1)$. Later on, following the seminal work of Barak, Impagliazzo, and Wigderson [2006a] which brought tools from additive combinatorics to bear on the construction of extractors, Barak et al. [2005] reduced the min-entropy for dispersers down to δn for any $\delta > 0$. In the meanwhile Bourgain [2005] used more tools from additive combinatorics and constructed a two-source extractor for min-entropy rate $\frac{1}{2} - \epsilon_0$ for some constant $\epsilon_0 > 0$, and this construction remains to this date the best in terms of its min-entropy rate. (If the *sum* of min-entropies of both sources is considered, Raz [2005] showed a construction that requires one source to have min-entropy rate just above half but the other source

[1] The original statement of Erdös [1947] was in terms of non-bipartite Ramsey graphs, but the proof method holds non-the-less for the case of bipartite Ramsey graphs, which are equivalent to two-source dispersers.

can have its min-entropy be as small as $O(\log n)$.) Finally, Barak et al. [2006b] constructed what remains the state of the art for two-source dispersers, achieving min-entropy n^δ for any $\delta > 0$.

Regarding conditional results, ones that depend on unproven conjectures, [Chor and Goldreich, 1988, Corollary 11] and [Zuckerman, 1991, Section 6.3] showed that the Paley Conjecture from number theory implies two-source extractors for very small min-entropy rate and Tauman Kalai et al. [2009] constructed two-source extractors based on cryptographic assumptions.

Affine extractors An *affine extractor* for min-entropy k is a function $g : \mathbb{F}_2^n \to \mathbb{F}_2^m$, where $\mathbb{F}_2$ denotes the two-element finite field, such that for every random variable X distributed uniformly on a k-dimensional affine subspace A of $\mathbb{F}_2^n$, the random variable $g(X)$ is close to being uniformly distributed on $\mathbb{F}_2^m$. A one-output-bit ($m = 1$) function that is nonconstant on every k-dimensional affine subspace is called an *affine disperser* for min-entropy k.

The probabilistic method can be used to show that affine extractors exist for min-entropy as small as $\log n + O(1)$ but up until recently explicit constructions were known only for min-entropy rate above half [Ben-Sasson et al., 2001]. This bound was broken by Barak et al. [2005] for the case of dispersers, they obtained dispersers for min-entropy rate δ for any $\delta > 0$. Bourgain [2007] used new bounds on exponential sums resulting from additive-combinatorics to construct affine extractors for min-entropy δ for any $\delta > 0$ that achieve exponentially small error (cf. Yehudayoff [2009], Li [2010] for improvements and alternative constructions along this line). Finally, Ben-Sasson and Kopparty [2009] showed constructions of affine dispersers for sublinear min-entropy as small as $n^{4/5}$, and some of their constructions turn out to be also affine extractors for constant min-entropy rate with sub-exponential error [Personal Communication].

1.2 Our two-source constructions

In this paper we analyze two families of two-source constructions. In both families we start with a function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ that we assume to be a "good" affine extractor (what we mean by "good" is explained in the next subsection). For the second construction we use $f^{-1}(z)$ to denote the set of preimages of $z \in \mathbb{F}_2^m$ and assume the existence of z with at least 2^{n-m} preimages (such z exists by the pigeonhole principle).

Concatenated two-source construction This construction takes two n-bit inputs x, y and is computed by *(i)* concatenating $f(x)$ to x, *(ii)* concatenating $f(y)$ to y and *(iii)* outputting the binary inner-product of the two (concatenated) strings. The binary inner-product of $z, z' \in \mathbb{F}_2^k$ is denoted by $\langle z, z' \rangle$ and defined by $\sum_{i=1}^k z_i \cdot z'_i$ where all arithmetic operations are in $\mathbb{F}_2$.

Preimage two-source construction Let F be a one-to-one mapping of $\mathbb{F}_2^{n-m}$ to $f^{-1}(z)$. On a pair of $(n-m)$-bit inputs x, y output the inner-product of $F(x)$ and $F(y)$.

Special cases of both constructions, which used specific functions f not necessarily known to be affine extractors, have been studied in the context of two-source dispersers and extractors — Bourgain [2007] used certain concatenated constructions and Pudlák and Rödl [2004] used preimage ones.

Each construction has its advantages. Assuming f is *explicit*, i.e., can be computed in time $n^{O(1)}$, inspection reveals that the concatenated construction is also explicit. The preimage one is not necessarily explicit, because F is not necessarily explicit even if f is. It is nonetheless *semi-explicit* — it can be computed in time $2^n \cdot \mathrm{poly}(n)$, which is quasi-linear[2] time in the size of the truth-table of the function (which is 2^n). Thus the concatenated construction is more efficient from a computational perspective. On

[2]We call a function $t : \mathbb{N} \to \mathbb{N}$ quasi-linear if $t(n) = O(n \cdot \mathrm{poly} \log n)$.

the other hand, using the preimage consturction one can potentially reach smaller min-entropy rate, as we shall see in the next Subsection.

Our proofs and resutls can be broken into two parts. First we show that each of our constructions is a two-source disperser for a certain min-entropy rate ρ that depends only on the parameters of the affine extractor we started with. Then we use approximate duality to bound the error on sources of min-entropy rate slightly larger than ρ. Next we elaborate on these two parts.

1.3 Part 1 - Two-source dispersers

First we explain what we mean by a "good" affine extractor. Suppose that $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is an affine extractor for min-entropy rate δ (reserving ρ for denoting the min-entropy rate of the two-source disperser built from f). This means that for every affine source X of min-entropy δn, the distance between the distribution of the random variable $f(X)$ and the uniform distribution over $\mathbb{F}_2^m$ is small. This distance is called the error of the extractor and is usually measured with the ℓ_1-norm (statistical distance). However, for our purposes it will be more convenient to measure the distance with the ℓ_∞-norm (throughout all the paper the term 'error' refers to the ℓ_∞-error unless stated otherwise, see Remark 2.3 for further explanation). By a "good" affine extractor we mean that its ℓ_∞-error is at most 2^{-m}.

Note that all one-output bit affine extractors satisfy the latter requirement for $m = 1$, and for our purposes we would like m to be as large as possible. In Section 6 we show that all existing affine extractors [Bourgain, 2007, Ben-Sasson and Kopparty, 2009, Li, 2010], which are typically stated as one-output bit extractors, can be easily converted into m-output bit affine extractors with ℓ_∞-error at most 2^{-m}, for a relatively large m. In particular, the construction of Bourgain [2007] (see also Yehudayoff [2009]) can be converted into such, where m is linear in n. Moreover, standard probabilistic methods (the Chernoff bound) can be used to show that a random function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ with $m = \delta n - O(\log n)$ is, with high probability, an affine extractor for min-entropy δn with ℓ_∞-error at most 2^{-m}.

It is convenient to express the min-entropy of our two-source dispersers in terms of the *min-entropy loss rate* $\lambda = 1 - \frac{m}{\delta n}$ of the affine extractor, which measures how much entropy is lost when going from X to $f(X)$. To see that λ does indeed measure entropy loss notice that in the extreme case in which $\lambda = 0$ we have $m = \delta n$ which means that f recovers almost all the entropy of X. In general, we would like λ to be as small as possible. The remarks above imply that a random function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ with $m = \delta n - O(\log n)$ is, with high probability, a "good" affine extractor with min-entropy loss rate $O(\log n/\delta n)$, while the affine extractor of Bourgain [2007] is a "good" affine extractor with min-entropy loss rate which is a constant strictly smaller than 1.

Our first pair of main results is that the given two constructions are two-source dispersers for min-entropy rates that depend only on the parameters of the affine extractors mentioned above. For the concatenated one we can show (in Theorem 2.7) it is a two-source disperser for min-entropy rate $\rho_{\text{concatenated}} < \frac{1}{2}$ as long as $\delta < \frac{1}{2}$, where $\rho_{\text{concatenated}} \to_{\lambda \to 0} \frac{2}{5}$. For the preimage one, assuming $\delta = \frac{1}{2}$, we show (in Theorem 2.9) that it is a two-source extractor for min-entropy rate at least $\rho_{\text{preimage}} = \frac{\lambda}{1+\lambda}$.

Notice that both constructions easily give dispersers for min-entropy rate at least $\rho < \frac{1}{2}$. As mentioned previously, the preimage construction can potentially reach much smaller min-entropy rate than the concatenated construction: when the min-entropy loss rate λ approaches zero so does the min-entropy rate of the preimage construction, but the concatenated construction does not go below min-entropy $\frac{2}{5}$ even if we assume that f has no min-entropy loss ($\lambda = 0$).

Proof overview To prove that our constructions are two-source dispersers consider $S, T \subset \{0,1\}^n, |S|, |T| > 2^{\rho n}$. Note that in both our constructions we first apply a function $h : \mathbb{F}_2^n \to \mathbb{F}_2^r$ to each of x and y separately to obtain x', y' and then apply a full-rank bilinear map (the binary inner-product function) to x', y'. Let $h(S) = \{h(s) \mid s \in S\}$ and define $h(T)$ similarly. Our main observation in this part is that if $\dim(\text{span}\,(h(S))) + \dim(\text{span}\,(h(T))) > r + 1$, then the function $E(x, y) = \langle h(x), h(y)\rangle$ must be non-constant on $S \times T$. Therefore in order to show that $E(x, y)$ is non-constant on $S \times T$ it suffices to show that each of $\dim(\text{span}\,(h(S)))$, $\dim(\text{span}\,(h(T)))$ is greater than $r/2$.

We note that in order to apply the argument above one does not necessarily have to use affine extractors. In particular, an explicit construction of a function $h : \mathbb{F}_2^n \to \mathbb{F}_2^r$, such that $\dim(\text{span}\,(h(S))) > r/2$ for every subset $S \subset \{0,1\}^n, |S| > 2^{\rho n}$ would suffice. Indeed, this part of our proof is inspired by [Pudlák and Rödl, 2004, Bourgain, 2007] (cf. Rao [2007]) which applied similar reasoning to particular functions which does not necessarily come from affine extractors.

1.4 Part 2 - Extractor-to-disperser reduction

As mentioned above, both our constructions are obtained by first applying a function on each input separately, and then applying the binary inner-product function. In fact, this property is not specific to our constructions, and we argue that every binary function $E(x, y) : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ can be constructed in this way. More precisely, for a binary function $E(x, y) : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ define its *rank* to be the smallest integer r, such that there exist functions $h_1, h_2 : \mathbb{F}_2^n \to \mathbb{F}_2^r$, such that $E(x, y) = \langle h_1(x), h_2(y)\rangle$. Alternatively, the rank of the function $E(x, y)$ is the rank over $\mathbb{F}_2$ of the $2^n \times 2^n$ matrix M, whose (x, y) entry equals $E(x, y)$.

The special thing about our constructions is that in both our constructions the rank is linear in the size n of the inputs. We point out this common denominator of the two constructions because our second main result says that any two-source disperser for min-entropy rate ρ which has rank n/ν has bounded error on sources of min-entropy rate $\rho + \gamma$ for any $\gamma > 0$. We give two bounds on the error in this case. The first one bounds the ℓ_∞-error by a non-trivial constant $\gamma' < \frac{1}{2}$ where γ' depends only on γ and ν. This result is stated in Lemma 2.12 and Theorem 2.13. The second bound says that the error is at most $2^{-\zeta n}$ for an absolute constant $\zeta > 0$ which depends on γ, ρ, and ν (see Lemma 2.17 and Theorem 2.18). The latter bound assumes the approximate duality conjecture (discussed next), a natural conjecture which is implied by, and implies a weak (though unproven) version, of the polynomial Freiman Ruzsa conjecture from additive combinatorics.

Proof overview To bound the error of two-source dispersers of linar-rank we study the notion of approximate duality. We define the *duality measure* of A, B in (1) as an estimate of how "close" this pair is to being dual [3].

$$\mu^\perp(A, B) \triangleq \left| \mathbb{E}_{a \in A, b \in B} \left[(-1)^{\langle a,b\rangle} \right] \right|. \tag{1}$$

It can be verified that if $\mu^\perp(A, B) = 1$ then A is contained in an affine shift of $B^\perp$ which is the space dual to the linear span of B. The question we study is what happens when $\mu^\perp(A, B)$ is large though strictly less than 1. We postulate that A, B contain pretty large subsets that have a duality measure 1 and prove (in Lemma 2.10) that this indeed holds when $\mu^\perp(A, B) > 1 - \epsilon$ for sufficiently small $\epsilon > 0$. The approximate duality conjecture (ADC, Conjecture 2.15) says that a similar statement should hold even when $\mu^\perp(A, B)$

[3] The duality measure can be alternatively defined as the discrepancy of the inner product function on the rectangle $A \times B$ (up to a normalization factor of $\frac{2^n}{|A||B|}$). We nevertheless chose to use the term 'duality measure' instead of 'discrepancy' because of the algebraic context in which we use it, and on which we elaborate next.

is exponentially small in n and before we justify our belief in this conjecture by relating it to the PFR let us see how approximate duality comes up in the analysis of the error of two-source extractors.

Suppose that the construction $E(x,y) = \langle h_1(x), h_2(y)\rangle$ is known to be a rank r two-source disperser for min-entropy rate ρ, assuming h_1, h_2 map $\mathbb{F}_2^n$ to $\mathbb{F}_2^r$. To prove that E is an extractor assume by way of contradiction that there exist $S, T \subset \mathbb{F}_2^n, |S|, |T| > 2^{(\rho+\gamma)n}$ on which E is very biased. Letting $\tilde{S} \subseteq \mathbb{F}_2^m$ be the set $\tilde{S} = h_1(S)$ and defining $\tilde{T} = h_2(T)$, our assumption is that $\mu^{\perp}(\tilde{S}, \tilde{T})$ is very large. Approximate duality statements like Lemma 2.10 and the ADC imply the existence of large sets $\hat{S} \subseteq \tilde{S}$ and $\hat{T} \subseteq \tilde{T}$ that have a duality measure of 1 and this implies that E is constant on the pair of large sets $S' = h_1^{(-1)}(\hat{S}), T' = h_2^{(-1)}(\hat{T})$ which contradicts our assumption that E is a two-source disperser.

Approximate duality and the polynomial Freiman Ruzsa conjecture The question addressed by the Freiman-Ruzsa Theorem [Freiman, 1973, Ruzsa, 1999] is the following[4]. Start by recalling that A is an affine subspace of $\mathbb{F}_2^n$ if and only if A does not expand under addition, by which we mean that $|A+A| = |A|$ where $A + A = \{a + a' \mid a, a' \in A\}$. Now suppose $A \subset \mathbb{F}_2^n$ behaves "approximately" like a subspace, i.e., $|A + A| \leq K|A|$ (think of $K \ll |A|$). Can we conclude that A is "close" to a subspace, meaning it contains a large subset A' that is itself a large fraction of a subspace H of $\mathbb{F}_2^n$? The Freiman-Ruzsa theorem gives a positive answer to this question, showing that there exists a subset $A' \subseteq A$ and a subspace $H \subseteq \mathbb{F}_2^n$ such that both fractions $|A'|/|A|$ and $|A'|/|H|$ can be bounded from below by $2^{-\mathrm{poly}(K)}$, the best lower bound on these ratios to date is of the form $K^{-O(K)}$ [Green and Tao, 2009]. The *polynomial Freiman-Ruzsa conjecture* (PFR) postulates that these ratios can be bounded from below by a polynomial, instead of exponential, function in K of the form $K^{-O(1)}$.

The question of approximate duality has a similar flavor: If two sets behave "approximately" like dual sets, do they contain large subsets that are strictly dual? Stated this way it seems natural to explore the connection between approximate duality and PFR, which is what we do later on in the paper. We show that PFR implies ADC and, in the reverse direction, ADC implies a "weak" form of PFR (which we call wPFR) that, although weaker than the PFR, is stronger than what is currently known.

1.5 Open questions

From two-source to affine extractors The question of possible connections between two-source and affine extractors was first raised by Barak et al. [2005] in Section 1.4, where they say about their affine dispersers:

> "Note that the new results here are quantitatively the same as our 2-source results ... The techniques are related as well, but at this point this fact may be surprising — there seem to be little resemblance between the models, and indeed there seem to be no reductions between them in either direction. The similarity in techniques may simply be a byproduct of the fact that we were working on them in parallel, ... however it would be interesting to find any tighter connections between the two models."

Our results address this question only in one direction, that of constructing two-source extractors out of affine ones. The reverse direction, that of constructing in a black-box manner affine extractors from two-source ones, remains wide open. This is somewhat perplexing because we would have guessed that the two-source-to-affine part should be easier. Counting the set of distinct sources that are uniformly distributed

[4] We describe the Freiman-Ruzsa Theorem for linear spaces over $\mathbb{F}_2$, the case most relevant to our study, whereas the Freiman-Ruzsa Theorem applies to arbitrary subsets of groups. See Green [2005b] and references within for more information.

over sets of size $2^{\rho n}$ we see there are $\binom{2^n}{2^{\rho n}}^2 \approx 2^{n \cdot 2^{\rho n}}$ of them, and this is much larger than the size of the set of affine sources, of which there are at most 2^{n^2}. All things considered it should be easier to go from extractors that work against a large set of sources to ones that work against a smaller set. We leave the problem of constructing affine extractors and dispersers from two-source ones as an interesting question for future research.

Constructing "good" affine extractors So far most work on affine extractors and dispersers has focused on reducing the min-entropy rate and significant progress has been made along this line, as surveyed in the previous section. But the question of minimizing the min-entropy loss rate and the ℓ_∞-error of affine extractors has received much less attention. Our work shows that at least as far as two-source constructions are concerned, it is the min-entropy loss rate and the ℓ_∞-error that should be minimized while the min-entropy rate can be set to be a pretty large constant, like $\frac{1}{2}$. It would be interesting to see if, for instance, affine extractors for small min-entropy rate and the tools used to analyze them could be converted into constructions for large min-entropy rate (like $\frac{1}{2}$) but with small min-entropy loss rate, and small ℓ_∞-error.

The approximate duality conjecture (ADC) We find this conjecture interesting both because its connection to the PFR conjecture and because its possible application to the task of constructing two-source extractors. Interesting avenues for future research are to pin down the exact versions of PFR and ADC that are equivalent (assuming they exist) and to study the ADC as a means to obtain a possibly weaker, though better than currently known, version of PFR. We have shown here that the ADC would imply better two-source extractors, and the wPFR implied by it could be sufficient for some of the other applications of the PFR (see Green [2005a] for a survey of some of them). A good starting point would be to prove a weakening of the ADC conjecture for the case in which the duality measure is small, but not exponentially small, e.g. when the duality measure is a small constant or even polynomially small in n. This will imply in turn that every two-source disperser of linear rank is a two-source extractor with small constant error or polynomially small error, respectively.

Affine extractors as fundamental building blocks of extractors In recent years we have seen a number of interesting constructions of extractors for structured sources of randomness, including "bit-fixing", "samplable" and "low-degree" sources, to name a few [Chor et al., 1985, Gabizon et al., 2006, Trevisan and Vadhan, 2000, Kamp and Zuckerman, 2007, Dvir, 2009, Dvir et al., 2009]. Our work suggests exploring the use of affine extractors in constructing extractors for these structured sources of randomness. A related question is to construct "seeded" extractors out of affine ones.

2 Main results

We start by defining the main objects of study in this paper — affine and two-source extractors (and dispersers) — and to do so introduce a bit of notation. We identify $\{0,1\}$ with the two-element field $\mathbb{F}_2$ and $\{0,1\}^n$ with $\mathbb{F}_2^n$. Given $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ and $x' = (x'_1, \ldots, x'_m) \in \mathbb{F}_2^m$ let $(x \circ x')$ denote their concatenation, i.e., $(x \circ x') = (x_1, \ldots, x_n, x'_1, \ldots, x'_m)$. For two sequences $x, y \in \mathbb{F}_2^k$ let $\langle x, y \rangle$ denote the $\mathbb{F}_2$-bilinear form $\langle x, y \rangle = \sum_{i=1}^k x_i \cdot y_i$, commonly referred to as the *inner-product function*. For $A \subset \mathbb{F}_2^n$ let $A^\perp$ denote the space that is dual to $\mathrm{span}\,(A)$, i.e., $A^\perp = \{b \in \mathbb{F}_2^n \mid \langle a, b \rangle = 0 \text{ for all } a \in A\}$. For $x \in \mathbb{F}_2^n$ we let $x + A := \{x + a | a \in A\}$. For a function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$, we denote by $f^{-1}(z)$ the set of preimages of the string z under the function f. For $A \subseteq \mathbb{F}_2^n$ we denote by $f(A)$ the image of A under f, i.e.,

$f(A) = \{f(a) \mid a \in A\}$.

A *source* over n bits is a distribution X over $\mathbb{F}_2^n$. The *min-entropy* of X is denoted by $H_\infty(X)$ and the *min-entropy rate* of X is $h_\infty(X) = H_\infty(X)/n$. If X is distributed uniformly over an affine subspace of $\mathbb{F}_2^n$ of dimension d we call X a d-dimensional *affine source*. Throughout the paper we reserve the letter E to denote various extractors, and $\mathbb{E}$ denotes expectation.

Definition 2.1. [Extractors and dispersers] Let $\mathcal{S}$ be a set of N-bit sources. A $[N, m, \mathcal{S}, \epsilon]$-extractor is a function $f : \mathbb{F}_2^N \to \mathbb{F}_2^m$ satisfying for every source $S \in \mathcal{S}$ that the distance between the random variable $f(S)$ and the uniform distribution over $\mathbb{F}_2^m$, measured with the ℓ_∞-norm, is at most ϵ. Namely, for every source $S \in \mathcal{S}$ and $y \in \mathbb{F}_2^m$ we require that

$$\left|\Pr[f(S) = y] - 2^{-m}\right| \leq \epsilon.$$

The function f is called an $[N, m, \mathcal{S}]$-*disperser* if f is nonconstant on every source $S \in \mathcal{S}$. An alternative definition is to say that the support of the random variable $f(S)$ is larger than 1 for every $S \in \mathcal{S}$. We call N the *source length*, m is the *output length*, and ϵ is the ℓ_∞-*error*, or simply the *error*, of the extractor. We shall be interested in extractors for two special kinds of sources:

- **Two-source extractors and dispersers:** When $N = 2n$ and $\mathcal{S}$ is the set of product distributions $S = X \times Y$ where both X and Y have min-entropy rate greater than ρ, we refer to $f : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^m$ as a $[n, m, \rho, \epsilon]$-*two source extractor*, or $[n, m, \rho]$-*two source disperser*.
- **Affine extractors and dispersers:** When $\mathcal{S}$ is the set of uniform distributions on affine subspaces of $\mathbb{F}_2^N$ of dimension greater than ρN, we refer to f as an $[N, m, \rho, \epsilon]$-*affine extractor*, or $[N, m, \rho]$-*affine disperser*.

When comparing the min-entropy of a source S to that of $f(S)$, it will be convenient for us to use the *entropy loss rate* parameter defined us

$$\lambda = \max_{S \in \mathcal{S}} 1 - \frac{H_\infty(f(S))}{H_\infty(S)}. \tag{2}$$

Intuitively, the loss rate measures how much (relative) min-entropy is lost when applying f to a source S in $\mathcal{S}$. Smaller loss λ corresponds to better extractors, ones that retain a larger min-entropy rate.

We end this part with two remarks on non-standard definitions:

Remark 2.2. [Dispersers] A more standard definition of a disperser, as appearing in, say, Shaltiel [2002], requires that for every source $S \in \mathcal{S}$, the support of the random variable $f(S)$ equals the full range of $\mathbb{F}_2^m$. Notice that for the case of $m = 1$ the two definitions match. All our constructions give two-source dispersers according to definition 2.1.

Remark 2.3. [ℓ_∞-error] Typically, the error of the extractor is measured with the ℓ_1-norm, and is defined to be the statistical distance between $f(S)$ and the uniform distribution over $\mathbb{F}_2^m$. We chose to measure the error with the ℓ_∞-norm, and henceforth the term "error" will refer to the ℓ_∞-error unless stated otherwise. The reason for this is that it will be relatively easy to analyze our constructions using this measure. For instance, we shall argue (in Section 6) that existing affine extractors, which are typically stated as one-output bit extractors, can be easily converted into m-output bit extractors with a relatively small loss in the ℓ_∞-error. For example, the original statement in Bourgain [2007] gives a family of $[n, 1, \delta, 2^{-\Omega(n)}]$-affine extractors, i.e., the output length is 1. It is nonetheless rather straightforward to obtain a linear number of output bits with essentially the same ℓ_∞-error (cf. Lemma 6.1).

Theorem 2.4. *[Bourgain's affine extractor] For every $\delta > 0$ there exists $\lambda_\delta < 1$ that depends only on δ such that there exists an explicit family of $[n, m = (1-\lambda_\delta)\delta n, \delta, 2^{-m}]$-affine extractors.*

Notice that the min-entropy loss rate of the construction above is $\lambda_\delta + (\delta n)^{-1} = \lambda_\delta + o(1)$. Using similar reasoning we will also show how to get multi-output bit two-source extractors out of our constructions (cf. Lemma 2.19). Notice that if f is an extractor with output length m and ℓ_∞-error ϵ then f has "standard" error (i.e., its ℓ_1 distance from uniform is) at most $\epsilon 2^m$.

2.1 Our two-source constructions

All results stated in this paper refer to the following two candidate constructions of two-source extractors.

Definition 2.5. [Concatenated construction] Given functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2^m$, the (f,g)-*concatenated construction* is the function $E^{\mathrm{c}}_{f,g} : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ defined for $x, y \in \mathbb{F}_2^n$ by

$$E^{\mathrm{c}}_{f,g}(x,y) = \langle (x \circ f(x)), (y \circ g(y)) \rangle. \tag{3}$$

Definition 2.6. [Preimage construction] Given functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2^m$ let $n' = n - m$. Let $z, z' \in \mathbb{F}_2^m$ satisfy $|f^{-1}(z)| \geq 2^{n'}$ and $|g^{-1}(z')| \geq 2^{n'}$. Let $F : \mathbb{F}_2^{n'} \to f^{-1}(z)$ and $G : \mathbb{F}_2^{n'} \to g^{-1}(z')$ be injective. The (F,G)-*preimage construction* is the function $E^{\mathrm{p}}_{F,G} : \mathbb{F}_2^{n'} \times \mathbb{F}_2^{n'} \to \mathbb{F}_2$ defined for $x, y \in \mathbb{F}_2^{n'}$ by

$$E^{\mathrm{p}}_{F,G}(x,y) = \langle F(x), G(y) \rangle. \tag{4}$$

2.2 Part 1 - Two-source dispersers

Our first pair of results is that both the concatenated and preimage constructions are two-source dispersers, or bipartite Ramsey graphs, for min-entropy rate below half. The preimage construction can reach arbitrarily small min-entropy rate provided the entropy loss of the affine extractor, together with its ℓ_∞-error, are sufficiently small. For $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $m' \leq m$, the m'-*bit projection* of f is obtained by taking the first m' bits of $f(x)$. Formally, if $f(x) = (y_1, \ldots, y_m)$ where $y_i \in \mathbb{F}_2$ then $f'(x) = (y_1, \ldots, y_{m'})$. To better understand the selection of parameters in the following theorem we point out that if f is an $[n, m = (1-\lambda)\delta n, \delta, 2^{-m}]$-affine extractor, then for any $m' \leq m$ the m'-bit projection of f is an $[n, m', \delta, 2^{-m'}]$-affine extractor.

Theorem 2.7. *[Concatenated two-source disperser from affine extractor] Suppose f and g are $[n, m = (1-\lambda)\delta n, \delta, 2^{-m}]$-affine extractors for $\delta < \frac{1}{2}$ and $\lambda < 1$. Let $\lambda' = \max\left\{\lambda, \frac{5}{3} - \frac{1}{3\delta}\right\}$ (noticing $\lambda' < 1$) and $\rho = \frac{1-\delta(1-\lambda')}{2}$ (noticing $\rho < \frac{1}{2}$). Set $m' = \lfloor \delta(1-\lambda')n \rfloor - 2$ and let f', g' be m'-bit projections of f, g respectively. Then $E^{\mathrm{c}}_{f',g'}$ is a $[n, 1, \rho]$-two-source disperser.*

The above theorem is proved in Section 3. Plugging Bourgain's affine extractor from Theorem 2.4 for min-entropy rate $\frac{1}{5}$ in the previous theorem, and noticing that for rate $\frac{1}{5}$ we have $\lambda' = \lambda$ we get:

Corollary 2.8. *[A two-source disperser for min entropy rate below half] Take $f = g$ to be Bourgain's $[n, m = (1-\lambda_{\frac{1}{5}})n/5, \frac{1}{5}, 2^{-m}]$-affine extractor, with min-entropy loss rate $\lambda_{\frac{1}{5}} < 1$. Then the concatenated construction $E^{\mathrm{c}}_{f,f}$ is a $[n, 1, \rho]$-two source disperser for min-entropy rate $\rho = \frac{2}{5} + \frac{\lambda_{\frac{1}{5}}}{10} < \frac{1}{2}$.*

Inspecting Theorem 2.7 we see that, even if we assume minimal loss $\lambda = 0$ and a min-entropy rate $\delta = \frac{1}{5}$ for the affine extractor which maximizes the min-entropy rate of the resulting concatenated two-source extractor, we end up with a two-source extractor for min-entropy rate $\frac{2}{5}$ (the proof appears in Section 4). This min-entropy rate barrier can be broken by the preimage construction.

Theorem 2.9. *[Preimage two-source disperser from affine extractor] If f and g are $[n, m = (1-\lambda)n/2, \frac{1}{2}, 2^{-m}]$-affine extractors and F, G are as in Definition 2.6, then $E^{\mathrm{p}}_{F,G}$ is a $[n' = \frac{1+\lambda}{2}n, 1, \frac{\lambda}{1+\lambda}]$-two-source disperser.*

2.3 Part 2 - Extractor-to-disperser reduction

2.3.1 Constant bounds on error by approximate duality for nearly-dual sets

We first show that the error of both dispersers presented above is bounded by a constant for a slightly larger min-entropy. This bound follows from a version of the ADC that we prove in Section 4.1. Recall the definition of the duality measure $\mu^{\perp}(A, B)$ of two sets given in (1). The version we prove requires $\mu^{\perp}(A, B)$ to be close to 1, i.e., A, B have to be "nearly-dual".

Lemma 2.10. *[Approximate-duality for nearly-dual sets] For every $\delta > 0$ there exists a constant $\epsilon > 0$ that depends only on δ, such that if $A, B \subseteq \mathbb{F}_2^n$ satisfy $\mu^{\perp}(A, B) \geq 1 - \epsilon$ then there exist subsets $A' \subseteq A, |A'| \geq \frac{1}{2}|A|$ and $B' \subseteq B, |B'| \geq 2^{-\delta n}|B|$, such that $\mu^{\perp}(A', B') = 1$.*

This lemma allows us to convert two-source dispersers of linear rank into two-source extractors with a large, though nontrivial, bound on error.

Definition 2.11. [Rank of a binary function] The *rank* of a function $E(x, y) : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ is the least integer r such that there exist functions $h_1, h_2 : \mathbb{F}_2^n \to \mathbb{F}_2^r$, satitfying $E(x, y) = \langle h_1(x), h_2(y) \rangle$ for all $x, y \in F_2^n$. Equivalently, this is the rank over $\mathbb{F}_2$ of the $2^n \times 2^n$ matrix whose (x, y)-entry is $E(x, y)$.

Lemma 2.12. *[Bounding the error of two-source dispersers of linear rank] For every $\gamma, \nu > 0$ there exists $\gamma' < 1/2$ such that the following holds for sufficiently large n. Every $[n, 1, \rho]$-two-source disperser of rank $\frac{n}{\nu}$ is a $[n, 1, \rho + \gamma, \gamma']$-two source extractor.*

The proof of the above lemma also appears in Section 4.1. Combining this lemma with Theorems 2.7 and 2.9 gives the following corollary.

Theorem 2.13. *[Nontrivial bounds disperser error] For all $\gamma, \lambda > 0$ there exists $\gamma' < 1/2$, depending only on γ and λ such that the following holds.*

1. *If f, g are $[n, m = (1 - \lambda)\delta n, \delta n, 2^{-m}]$-affine extractors, then $E^{\mathrm{c}}_{f',g'}$ defined in Theorem 2.7 is a $[n, 1, \rho + \gamma, \gamma']$-two source extractor for ρ as defined in the same theorem.*
2. *If f, g are $[n, m = (1-\lambda)n/2, \frac{1}{2}, 2^{-m}]$-affine extractors and F, G are as in Definition 2.6, then $E^{\mathrm{p}}_{F,G}$ is a $[n' = \frac{1+\lambda}{2}n, 1, \frac{\lambda}{1+\lambda} + \gamma, \gamma']$-two source extractor.*

2.3.2 Exponentially small bounds on error using the approximate duality conjecture

Both constructions — concatenated and primage — are two-source extractors with exponentially small error if the following well-known conjecture from additive combinatorics is true.

Conjecture 2.14. *[Polynomial Freiman-Ruzsa (PFR)] There exists an integer r such that if $A \subset \mathbb{F}_2^n$ has $|A + A| \leq K|A|$, then A may be covered by at most K^r cosets of some subspace of size at most $K^r|A|$.*

It will be easier to work with the following conjecture, which we later on show is implied by PFR.

Conjecture 2.15. *[Approximate Duality (ADC)] For every pair of constants $\alpha, \delta > 0$ there exists a constant $\zeta > 0$, depending only on α and δ such that the following holds for sufficiently large n. If $A, B \subseteq \mathbb{F}_2^n$ satisfy $|A|, |B| > 2^{\alpha n}$ and $\mu^{\perp}(A, B) \geq 2^{-\zeta n}$, then there exist subsets $A' \subseteq A$, $|A'| \geq 2^{-\delta n}|A|$ and $B' \subseteq B$, $|B'| \geq 2^{-\delta n}|B|$ such that $\mu^{\perp}(A', B') = 1$.*

Remark 2.16. [Exponential loss is necessary] It may seem that the factor $2^{-\delta n}$ appearing in the bound on the relative size of A', B' can be avoided or perhaps replaced by a polynomial factor in n. While there should be some room for improvement in the parameters of the conjecture, Shachar Lovett pointed out to us [personal communication] that a factor of $2^{-\Omega(\sqrt{n})}$ is unavoidable even for large values of $\mu^{\perp}(A, B)$: Take $n = 3k$ and let S denote the subset of k-bit strings of Hamming weight $\gamma\sqrt{k}$ for constant $\gamma > 0$. Consider $A = \mathbb{F}_2^k \times S \times \left\{0^k\right\}, B = \left\{0^k\right\} \times S \times \mathbb{F}_2^k$. It is not hard to verify that $|A| = |B| \approx 2^{n/3}$ and $\mu^{\perp}(A, B) \geq \epsilon$, where $\epsilon > 0$ depends on γ and can be set to be arbitrarily close to 1, but the largest equal-sized dual subsets of A, B have size at most $2^{-\Omega(\sqrt{n})} \cdot |A|$ and $2^{-\Omega(\sqrt{n})} \cdot |B|$ respectively.

The ADC, which is implied by the PFR, allows us to argue (in Section 4.2) that any two-source disperser of linear rank is actually a two-source extractor, with exponentially small error, for roughly the same min-entropy as the disperser.

Lemma 2.17. *[Two-source dispersers of linear rank are extractors] Assuming ADC (Conjecture 2.15), for every $\rho, \gamma, \nu > 0$ there exists $\zeta > 0$ such that the following holds for sufficiently large n: Every $[n, 1, \rho]$-two source disperser of rank $\frac{n}{\nu}$ is a $[n, 1, \rho + \gamma, 2^{-\zeta n}]$-two-source extractor.*

Theorem 2.18. *[Two-source extractors from affine ones] Assuming ADC, for every $\delta, \lambda, \gamma > 0$ there exists $\zeta > 0$ such that the following holds for sufficiently large n.*

1. *If f, g are $[n, m = (1-\lambda)\delta n, \delta, 2^{-m}]$-affine extractors then $E^{\mathrm{c}}_{f',g'}$ defined in Theorem 2.7 is a $[n, 1, \rho + \gamma, 2^{-\zeta n}]$-two source extractor for ρ defined in that theorem.*
2. *If f, g are $[n, m = (1-\lambda)n/2, \frac{1}{2}, 2^{-m}]$-affine extractors and F, G are as in Definition 2.6, then $E^{\mathrm{p}}_{F,G}$ is a $[n' = \frac{1+\lambda}{2}n, 1, \frac{\lambda}{1+\lambda} + \gamma, 2^{-\zeta n}]$-two source extractor.*

2.3.3 Multi-output two-source extractors

We end by pointing out that since our constructions are *linear invariant* of rank $\Omega(n)$, we can use the following lemma to increase their number of output bits to a number linear in n, while maintaining the same ℓ_∞-error.

Let $h_1, h_2 : \mathbb{F}_2^n \to \mathbb{F}_2^r$. In what follows call a $[n, 1, \rho, \epsilon]$-two source extractor $E(x, y) = \langle h_1(x), h_2(y) \rangle$ *linear invariant with respect to* h_1, h_2 if for every full rank matrix $M \in \mathbb{F}_2^{r \times r}$, $E(x, y) = \langle h_1(x), Mh_2(y) \rangle$ is also a $[n, 1, \rho, \epsilon]$-two source extractor. Inspection of the proofs of Theorems 2.7, 2.9, 2.13, 2.18 reveals that both our constructions are linear invariant with respect to the functions h_1, h_2 with which they were defined (since in Theorems 2.7, 2.9 we only use the fact that the sets $h_1(S), h_2(T)$ have large span, while in Theorems 2.13, 2.18 we only use the fact that $h_1(S), h_2(T)$ are large enough subsets of $\{0, 1\}^r$).

Call a set of matrices $M_1, \ldots, M_r \in \mathbb{F}_2^{r \times r}$ *independent* if they satisfy the following property: For every $v_1, \ldots, v_r \in \mathbb{F}_2$ not all zero, the matrix $\sum_i v_i M_i$ has full rank. In Section 6 we explain how a collection of independent matrices can be obtained. There we also prove the following statement.

Lemma 2.19. *[Multi-output extractors] Let $h_1, h_2 : \mathbb{F}_2^n \to \mathbb{F}_2^r$ be such that the rank r function*

$$E(x, y) = \langle h_1(x), h_2(y) \rangle$$

is a $[n, 1, \rho, \epsilon]$-two source extractor which is linear invariant with respect to h_1, h_2. Then for $t \leq r$, and independent matrices $M_1, \ldots, M_t \in \mathbb{F}_2^{r \times r}$, the function $E : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^t$ defined by

$$E(x, y) = (\langle h_1(x), M_1 h_2(y) \rangle, \ldots, \langle h_1(x), M_t h_2(y) \rangle)$$

is a $[n, t, \rho, \epsilon]$-two source extractor.

2.4 On the polynomial Freiman Ruzsa and the approximate duality conjectures

We end the description of our main results by describing the relationship between the PFR and ADC. We have already said that the PFR conjecture implies the ADC one. To prove this implication it is crucial to us that the exponent r in the PFR conjecture be close to 1, i.e., that the *polynomial* in the "polynomial Freiman Ruzsa" conjecture be nearly-linear. To achieve this, we are willing to assume not only that $2A$ is small but even that $\ell A = \left\{ \sum_{i=1}^{\ell} a_i \mid a_i \in A \right\}$ is small for some constant $\ell > 2$. In other words, to prove the ADC what we really need is the *Nearly-linear Freiman Ruzsa (NLFR) conjecture*:

Conjecture 2.20. *[Nearly-linear Freiman-Ruzsa (NLFR)] For every $\rho > 0$ there exists an integer ℓ which depends only on ρ, such that if $A \subset \mathbb{F}_2^n$ has $|\ell A| \leq K|A|$, then A may be covered by at most K^ρ cosets of some subspace of size at most $K|A|$.*

In Section 5 we show that the NLFR and PFR are equivalent. (The implication NLFR $\Rightarrow$ PFR is relatively easy but the other direction is nontrivial.) We also show that NLFR $\Rightarrow$ ADC. Regarding the reverse direction, we show that the ADC implies the following weaker form of PFR:

Conjecture 2.21. *[Weak PFR for dense sets (wPFR)] For every $1 > \alpha > 0$ and $1 > \delta > 0$, there exists an integer r which depends only on α and δ, such that if $A \subset \mathbb{F}_2^n$ has $2^{\alpha n} \leq |A| \leq 2^{(1-\alpha)n}$ and $|A+A| \leq K|A|$, then A may be covered by at most $2^{\delta n} \cdot K$ cosets of some subspace of size at most $2^{\delta n} K^r |A|$.*

The above conjecture differs from the standard PFR conjecture in two ways. First, in the above conjecture the set A must be of high density, and the exponent r depends on the density of the set. Second, the number of cosets and the size of the subspace are multiplied by an exponential factor. However, this exponential factor can set to be arbitrarily small, at the cost of enlarging r.

The relation between these conjectures can be summarized by:

$$(\text{PFR} \Leftrightarrow \text{NLFR}) \Rightarrow \text{ADC} \Rightarrow \text{wPFR}$$

2.5 Organization of the rest of the paper

In the next section we prove that our constructions are two-source dispersers and then in Section 4 we show that this also implies that they are extractors for roughly the same min-entropy. In Section 5 we discuss the relation between PFR, NLFR and ADC in more detail. Finally, in Section 6 we analyze the ℓ_∞-error and the number of output bits of existing affine extractors and multi-output bit two-source extractors.

3 Part 1 - Two-source dispersers

In this section we prove that plugging an affine extractor with sufficiently good parameters into our two-source constructions results in a two-source disperser for min-entropy rate that is related to the parameters of the affine extractor. Our proofs in this section rely on the following elementary lemma, whose proof we bring here for completeness. For $S \subseteq \mathbb{F}_2^n$ we denote by $\dim(S)$ the dimension of $\text{span}\,(S)$.

Lemma 3.1. *Let A, B be subsets of $\mathbb{F}_2^n$ such that*

$$\dim(A) + \dim(B) > n + 1.$$

Then the binary inner product function $IP(x, y) = \langle x, y \rangle$ is non-constant on $A \times B$.

Proof. Suppose in contradiction that $IP(A, B)$ is constant. Then there are two cases. The first case is when $IP(A, B) = 0$. In this case A is contained in $B^{\perp}$, which implies in turn that $\dim(A) + \dim(B) \leq n$, a contradiction. The second case is when $IP(A, B) = 1$. In this case let $a \in A$. Then for any other element $a' \in A$ we have that $\langle a' - a, b \rangle = \langle a, b \rangle - \langle a', b \rangle = 0$. It follows that the set $A - a$ is contained in $B^{\perp}$, and therefore $\dim(A - a) + \dim(B) \leq n$. This in turn implies that $\dim(A) + \dim(B) \leq n + 1$, which is again a contradiction. □

3.1 Concatenated two-source disperser — Proof of Theorem 2.7

The main step in the proof is the following lemma. Before proving the lemma we show how it implies Theorem 2.7.

Lemma 3.2. *[Affine extractors lead to dimension expansion] Suppose f is an $[n, m, \delta, 2^{-m}]$-affine extractor. Then for every $S \subseteq \mathbb{F}_2^n$ of size greater than $2^{m+\delta n}$, denoting*

$$\overline{S} = \{(x \circ f(x)) | x \in S\},$$

we have

$$\dim(\overline{S}) \geq \lfloor \log |S| \rfloor + m - 1.$$

Proof of Theorem 2.7. Our choice of $\rho = \frac{1-\delta(1-\lambda')}{2}$ and $\lambda' \geq \frac{5}{3} - \frac{1}{3\delta}$ implies that

$$\rho \geq \delta(2 - \lambda'). \tag{5}$$

Given two n-bit sources X, Y of min-entropy rate greater than ρ let $S, T \subseteq \mathbb{F}_2^n$ denote their respective supports. Recalling $m' = \delta(1 - \lambda')n - 2$ we conclude from (5) that

$$|S|, |T| > 2^{\rho n} \geq 2^{\delta n + m'}.$$

Letting $\overline{S} = \{(x \circ f'(x)) | x \in S\}$ and $\overline{T} = \{(y \circ g'(y)) | y \in T\}$, Lemma 3.2 implies

$$\dim(\overline{S}), \dim(\overline{T}) > \rho n + m' - 1 \geq \frac{n + m'}{2}.$$

The last inequality follows because $\rho n = \frac{1-\delta(1-\lambda')}{2} n > \frac{n-m'}{2} + 1$. We conclude that $\dim(\overline{S}) + \dim(\overline{T}) > m' + n + 1$. Lemma 3.1 then implies that $E^{c}_{f',g'}(x, y)$ is non-constant on $S \times T$, thereby completing the proof. □

And now we give the proof of Lemma 3.2.

Proof of Lemma 3.2. Denote $\dim(\overline{S})$ by d, noticing $d \geq m + \delta n$. To prove the lemma we will show

$$d \geq \lfloor \log |S| \rfloor + m - 1.$$

Let $\pi_1 : \mathbb{F}_2^{n+m} \to \mathbb{F}_2^n$ be the linear operator which projects $\mathbb{F}_2^{n+m}$ onto the first n bits[5] and let $\pi_2 : \mathbb{F}_2^{n+m} \to \mathbb{F}_2^m$ be the projection onto the last m bits.

Start with a basis $v_1, \ldots, v_d$ for $\mathrm{span}\left(\overline{S}\right)$. Without loss of generality suppose that its last r vectors are such that $\pi_2(v_{d-r+1}), \ldots, \pi_2(v_d)$ form a basis to $\pi_2(\mathrm{span}\left(\overline{S}\right))$ (noticing that $0 < r \leq m$). Use Gaussian elimination to make the first $d-r$ elements in this basis have their support in the first n bits. Let $V_1 := \mathrm{span}\left(\{\pi_1(v_1), \ldots, \pi_1(v_{d-r})\}\right)$, and $V_2 := \mathrm{span}\left(\{\pi_2(v_{d-r+1}), \ldots, \pi_2(v_d)\}\right)$.

First we note that $f(S) \subseteq V_2$, which implies that $|\ f(S)\ | \leq 2^r$. Our goal will be to show that for every string $z \in f(S)$ the number of preimages of z under f in the set S is at most $2^{d-r-m+1}$, that is, $|f^{-1}(z) \cap S| \leq 2^{d-r-m+1}$. This will conclude the proof of the lemma as this will imply that

$$|S| = \sum_{z \in f(S)} |f^{-1}(z) \cap S| \leq \sum_{z \in f(S)} 2^{d-r-m+1} \leq 2^r \cdot 2^{d-r-m+1} \leq 2^{d-m+1}$$

and the proof is completed by taking logarithm of both sides.

Let $z \in f(S)$. Our main observation that will allow us to bound the size of $f^{-1}(z) \cap S$ is that $f^{-1}(z) \cap S$ is contained in an affine shift of V_1 by some vector $w \in \mathbb{F}_2^n$. Assuming that this is true, and noticing that the dimension of V_1 is $d - r \geq (m + \delta n) - m = \delta n$, the fact that f is an $[n, m, \delta, 2^{-m}]$-affine extractor implies that z cannot have too many preimages under f in the set S. In particular,

$$|f^{(-1)}(z) \cap S| \leq |f^{(-1)}(z) \cap (w + V_1)| \leq 2^{d-r} \cdot 2^{-m+1} = 2^{d-r-m+1}$$

It remains to show that $f^{-1}(z) \cap S$ is contained in an affine shift of V_1. Since $z \in f(S) \subseteq V_2$, and since the vectors $\pi_2(v_{d-r+1}), \ldots, \pi_2(v_d)$ form a basis for V_2, we conclude that there exists a unique vector $y \in \mathrm{span}\left(\{v_{d-r+1}, \ldots, v_d\}\right)$ such that $\pi_2(y) = z$. We argue that $f^{-1}(z) \cap S \subseteq \pi_1(y) + V_1$. Indeed, suppose that $s \in f^{-1}(z) \cap S$. Then from the uniqueness of y, and since the basis vectors $v_1, \ldots, v_{d-r}$ have their support in the first n bits, we have that $(s, f(s)) = (s, z) = x + y$, where $x \in \mathrm{span}\left(\{v_1, \ldots, v_{d-r}\}\right)$. But this implies in turn that $s = \pi_1(x) + \pi_1(y)$, where $\pi_1(x) \in V_1$, and hence $s \in \pi_1(y) + V_1$, which concludes the proof of the lemma. □

3.2 Preimage two-source dispersers — Proof of Theorem 2.9

Proof of Theorem 2.9. Let $m = \frac{1-\lambda}{2}n$ and recall $n' = n - m = \frac{1+\lambda}{2}n$. Let $z, z' \in \mathbb{F}_2^m$ be the strings from Definition 2.6, such that F is an injective mapping of $\mathbb{F}_2^{n'}$ into $f^{-1}(z)$ and G is an injective mapping of $\mathbb{F}_2^{n'}$ into $g^{(-1)}(z')$. Given two n'-bit sources X, Y of min-entropy rate greater than $\frac{\lambda}{1+\lambda}$ let $S, T \subseteq \mathbb{F}_2^n$ denote the respective supports of $F(X), G(Y)$, noticing $|S|, |T| > 2^{\frac{\lambda}{1+\lambda}n'} = 2^{\frac{\lambda}{2}n}$. We shall show that $\dim(S), \dim(T) > \frac{n}{2}$, thereby completing our proof due to Lemma 3.1.

By symmetry it suffices to prove the claim only for S. We will prove that S is not contained in any affine space of dimension $n/2$. Let A be such a space. By Definition 2.1 we get

$$|A \cap S| \leq |A \cap f^{-1}(z)| \leq 2 \cdot 2^{-\frac{1-\lambda}{2}n} \cdot 2^{\frac{n}{2}} = 2^{\frac{\lambda}{2}n+1} < |A|.$$

where the last inequality is true for sufficiently large n due to the fact that $\lambda < 1$. We conclude that $S \nsubseteq A$ and since this holds for all affine spaces of dimension $n/2$ our proof is complete. □

[5]Formally, letting $\left\{e_1^{(t)}, \ldots, e_t^{(t)}\right\}$ denote the standard basis for $\mathbb{F}_2^t$ and representing elements of $\mathbb{F}_2^{n+m}$ in this basis for $t = n + m$, we define $\pi_1(\sum_{i=1}^{n+m} a_i e_i^{(n+m)}) = \sum_{j=1}^{n} a_j e_j^{(n)}$ and $\pi_2(\sum_{i=1}^{n+m} a_i e_i^{(n+m)}) = \sum_{j=n+1}^{n+m} a_j e_{j-n}^{(m)}$.

4 Part 2 - Extractor-to-disperser reduction

4.1 Constant bounds on error by approximate duality for nearly-dual sets

In this Section we prove Lemma 2.12 which shows that any two-source disperser of linear rank is actually a two-source extractor with a nontrivial, alas large, bound on the error. This result is implied by the Approximate-Duality Lemma for nearly-dual sets (Lemma 2.10). We start with the proof of this lemma and then prove Lemma 2.12 which is implied by it. We shall need the notion of the *spectrum* of a set which we take from [Tau and Vu, 2006, Chapter 4]. This concept will be used also in proofs that appear later on.

Definition 4.1. [Spectrum] For a set $B \subseteq \mathbb{F}_2^n$ and $\alpha \in [0,1]$ let the α-spectrum of B be the set

$$\mathrm{spec}_\alpha(B) := \left\{ x \in \mathbb{F}_2^n \mid \mathbb{E}_{b \in B}\left[(-1)^{\langle x,b\rangle}\right] \geq \alpha \right\}.$$

Proof of Lemma 2.10. We assume without loss of generality that $\mathbb{E}_{a\in A, b\in B}\left[(-1)^{\langle a,b\rangle}\right] > 0$, the proof for the case in which $\mathbb{E}_{a\in A, b\in B}\left[(-1)^{\langle a,b\rangle}\right] < 0$ is similar. Let $A' = A \cap \mathrm{spec}_{1-2\epsilon}(B)$. The assumption $\mu^\perp(A,B) \geq 1-\epsilon$ together with Markov's inequality shows $|A'| \geq \frac{1}{2}|A|$.

Let $a_1, \ldots, a_d \in A'$ form a basis for $\mathrm{span}\,(A')$. The elements $a_1, a_2, \ldots, a_d$ partition $\mathbb{F}_2^n$ into 2^d sets, where each set is an affine coset of $\{a_1, a_2, \ldots, a_d\}^\perp = A'^\perp$. For a vector $x \in \mathbb{F}_2^n$ let $w(x)$ denote the fraction of zeros in the set $\{\langle x, a_i\rangle \mid i \in \{1,2,\ldots,d\}\}$, i.e.,

$$w(x) = \frac{|\{\langle x, a_i\rangle = 0 \mid i \in \{1,\ldots,d\}\}|}{d}$$

Note that in every affine coset of $A'^\perp$ all elements have the same value $w(x)$. Moreover, for every $0 \leq t \leq d$, there are precisely $\binom{d}{t}$ cosets H of $A'^\perp$ such that $w(x) = 1 - \frac{t}{d}$ for every $x \in H$.

Our main observation is that since $A' \subseteq \mathrm{spec}_{1-2\epsilon}(B)$, a large fraction of elements $b \in B$ have large value $w(b)$. Consequently they cannot participate in too many different affine cosets of $A'^\perp$, and in particular there exists one such affine coset which contains a large fraction of b's in B. If we let $\hat{B}$ denote the set which contains all elements that lie in this coset, then we obtain a large subset $\hat{B} \subseteq B$, such that $\hat{B}$ lies in an affine shift of $A'^\perp$, and this almost gives the desired set B'. Details follow.

In what follows let $H : (0,1) \to (0,1)$ denote the binary entropy function given by:

$$H(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$$

Choose $0 < \epsilon \leq \frac{1}{8}$ such that $H(\sqrt{2\epsilon}) < \delta$. Let $\alpha = \frac{d}{n}$ denote the fractional dimension of $\mathrm{span}\,(A')$.

Since $a_1, \ldots, a_d$ are all contained in $\mathrm{spec}_{1-2\epsilon}(B)$ we have $\mathbb{E}_{b\in B}\left[w(b)\right] \geq 1 - 2\epsilon$. From Markov's inequality, this implies that at least $\left(1-\sqrt{2\epsilon}\right)$-fraction of b's in B satisfy $w(b) \geq 1 - \sqrt{2\epsilon}$. We let $\tilde{B}$ denote the subset of B which contains all elements in B which satisfy $w(b) \geq 1 - \sqrt{2\epsilon}$. We are forced to pick $\tilde{B}$ from affine cosets H of $A'^\perp$ such that $w(x) \geq 1 - \sqrt{2\epsilon}$ for every $x \in H$. The number of such cosets is at most:

$$\sum_{0 \leq t \leq \sqrt{2\epsilon}\alpha n} \binom{\alpha n}{t} \leq 1 + \sqrt{2\epsilon}\alpha n \cdot \binom{\alpha n}{\sqrt{2\epsilon}\alpha n} = 2^{\left(H\left(\sqrt{2\epsilon}\right) + o(1)\right)\alpha n} \leq 2^{\left(H\left(\sqrt{2\epsilon}\right) + o(1)\right)n},$$

where the first inequality is due to our choice of $\epsilon \leq \frac{1}{8}$, which implies $\sqrt{2\epsilon}\alpha n \leq \frac{1}{2}\alpha n$.

This in turn implies the existence of an affine coset of $A'^{\perp}$ which contains at least a $2^{-\left(H\left(\sqrt{2\epsilon}\right)+o(1)\right)n}$-fraction of b's in $\tilde{B}$. Let $\hat{B}$ denote the subset of $\tilde{B}$ which is contained in this affine coset. Recalling we set $\delta > H(\sqrt{2\epsilon})$ we get for sufficiently large n

$$|\hat{B}| \geq 2^{-\left(H\left(\sqrt{2\epsilon}\right)+o(1)\right)n} \cdot |\tilde{B}| \geq 2^{-\left(H\left(\sqrt{2\epsilon}\right)+o(1)\right)n} \cdot |B| \geq 2 \cdot 2^{-\delta n}|B|.$$

We have almost concluded the proof. We have at hand a pretty large set $\hat{B}$ that is contained in $a + A'$ for some $a \in \mathbb{F}_2^n$. Partition $\hat{B}$ into $\hat{B}_0 = \left\{b \in \hat{B} \mid \langle b, a\rangle = 0\right\}$ and $\hat{B}_1 = \left\{b \in \hat{B} \mid \langle b, a\rangle = 1\right\}$. To complete the proof of the lemma take B' to be the larger of $\hat{B}_0, \hat{B}_1$ and notice $|B'| \geq 2^{-\delta n}|B|, |A'| \geq \frac{1}{2}|A|$ and $\mu^{\perp}(B', A') = 1$. □

Proof of Lemma 2.12. Let $\delta = (1/\nu + 2)^{-1}\gamma$ and set $\gamma' = \frac{1-\epsilon}{2}$ where $\epsilon = \epsilon(\delta) > 0$ is the constant guaranteed by Lemma 2.10. We argue by way of contradiction. Let X and Y be two sources of min-entropy rate $> \rho + \gamma$ which we assume without loss of generality to be uniformly distributed over sets A,B respectively, each of size greater than $2^{(\rho+\gamma)n}$, and for which the error of $E(X, Y)$ is greater than γ'.

Assuming E has rank n/ν there exist functions $h_1, h_2 : \mathbb{F}_2^n \to \mathbb{F}_2^{n/\nu}$ such that $E(x, y) = \langle h_1(x), h_2(y)\rangle$. For our purposes we need h_1 and h_2 to be bijective. For this end we let $\tilde{h}_1, \tilde{h}_2 : \mathbb{F}_2^n \to \mathbb{F}_2^{n/\nu+2n}$ be the functions defined by $\tilde{h}_1(x) = (h_1(x) \circ x \circ 0_n)$ and $\tilde{h}_2(y) = (h_2(y) \circ 0_n \circ y)$, where 0_n denotes the all-zeros vector of length n. One can check that $\tilde{h}_1, \tilde{h}_2$ are bijective functions, while it still holds that $E(x, y) = \langle \tilde{h}_1(x), \tilde{h}_2(y)\rangle$.

Let $\overline{A} = \{\tilde{h}_1(a) | a \in A\}$ and $\overline{B} = \{\tilde{h}_2(b) | b \in B\}$. Assuming the ℓ_∞-error of $E(X, Y)$ is greater than $\frac{1-\epsilon}{2}$, is equivalent to saying $\mu^{\perp}(\overline{A}, \overline{B}) > 1 - \epsilon$. Consequently, Lemma 2.10 implies the existence of subsets $A' \subseteq \overline{A}, |A'| \geq \frac{1}{2}|\overline{A}|$ and $B' \subseteq \overline{B}, |B'| \geq 2^{-\gamma(1/\nu+2)^{-1}(n/\nu+2n)}|\overline{B}| = 2^{-\gamma n}|\overline{B}| \geq 2^{\rho n}$ such that $\mu^{\perp}(A', B') = 1$. Let $\hat{A} := \tilde{h}_1^{(-1)}(A'), \hat{B} := \tilde{h}_2^{(-1)}(B')$. Then $\hat{A}$ and $\hat{B}$ are sets of size at least $2^{\rho n}$ each, such that $|E(\hat{A}, \hat{B})| = 1$, contradiction.

□

4.2 Exponentially small bounds on error using the approximate duality conjecture

We now show that, assuming ADC, our two-source dispersers are extractors with exponentially small error.

Proof of Lemma 2.17. Let $E : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ be the $[n, 1, \rho]$ two-source disperser of rank n/ν which is defined by $E(x, y) = \langle h_1(x), h_2(y)\rangle$. Let $\tilde{h}_1, \tilde{h}_2 : \mathbb{F}_2^n \to \mathbb{F}_2^{n/\nu+2n}$ be the functions defined by $\tilde{h}_1(x) = (h_1(x) \circ x \circ 0_n)$ and $\tilde{h}_2(y) = (h_2(y) \circ 0_n \circ y)$, where 0_n denotes the all-zeros vector of length n. Let ζ' be the constant guaranteed by Conjecture 2.15 for the constants $\alpha = (\rho + \gamma)(1/\nu + 2)^{-1}$ and $\delta = \gamma(1/\nu + 2)^{-1}$, and let $\zeta = \zeta'(1/\nu+2)$. Our proof goes by way of contradiction, along the lines of the proof of Lemma 2.12.

Let X and Y be two n-bit sources of min-entropy rate $> \rho + \gamma$, we assume without loss of generality these sources to be uniform distributions over sets A,B respectively, each of size greater than $2^{(\rho+\gamma)n}$. Let $\overline{A} = \left\{\tilde{h}_1(a) \mid a \in A\right\}$ and $\overline{B} = \left\{\tilde{h}_2(b) \mid b \in B\right\}$. Notice $\overline{A}, \overline{B} \subseteq \mathbb{F}_2^{n/\nu+2n}$ and $|\overline{A}|, |\overline{B}| \geq 2^{(\rho+\gamma)n} = 2^{\alpha(n/\nu+2n)}$. Assume by way of contradiction that the error of $E(X, Y)$, which equals $\frac{1}{2}\mu^{\perp}(\overline{A}, \overline{B})$, is greater than $2^{-\zeta n} = 2^{-\zeta'(n/\nu+2n)}$. Applying ADC to $\overline{A}, \overline{B}$ we conclude the existence of subsets $A' \subseteq \overline{A}, B' \subseteq \overline{B}$ such that $\mu^{\perp}(A', B') = 1$ and A', B' are quite large,

$$|A'| \geq \frac{|\overline{A}|}{2^{\delta(n/\nu+2n)}} > \frac{2^{(\rho+\gamma)n}}{2^{\gamma(1/\nu+2)^{-1}(n/\nu+2n)}} = \frac{2^{(\rho+\gamma)n}}{2^{\gamma n}} > 2^{\rho n}$$

and similarly, $|B'| > 2^{\rho n}$. But since $\tilde{h}_1, \tilde{h}_2$ are injective we deduce that X', Y', which are uniformly distributed over $\tilde{h}_1^{(-1)}(A')$ and $\tilde{h}_2^{(-1)}(B')$, are a pair of n-bit sources of min-entropy rate greater than ρ on which E is constant, contradiction. □

5 On the polynomial and nearly-linear Freiman Ruzsa conjectures and the approximate duality conjecture

In this section we study the relation between the PFR, NLFR, and ADC. We shall prove the following relations between these conjectures:

$$(\text{PFR} \Leftrightarrow \text{NLFR}) \Rightarrow \text{ADC} \Rightarrow \text{wPFR}$$

We start by showing the equivalence of PFR and NLFR in the next two sections. Then, in Section 5.3 we move to the rightmost implication. We end in Section 5.4 with the most complicated proof, that of the middle implication (NLFR ⇒ ADC).

5.1 The nearly-linear polynomial Freiman Ruzsa conjecture implies the polynomial one

The implication NLFR ⇒ PFR is a relatively easy consequence of the following inequality of Plunnecke [1969], a new proof of which was found by Ruzsa [1989]:

Theorem 5.1. *[Plunnecke's inequality] Let A, B, be finite sets in a commutative group, and suppose that $|A+B| \leq K|A|$. Then for arbitrary nonnegative integers m,n we have:*

$$|mB - nB| \leq K^{m+n}|A|$$

To show NLFR ⇒ PFR choose $\rho = 1$ in Conjecture 2.20 (NLFR), and let ℓ be the integer guaranteed by this conjecture for $\rho = 1$. Assuming $|A + A| \leq K|A|$, Theorem 5.1 implies that $|\ell A| \leq K^{\ell}|A|$. So NLFR (Conjecture 2.20) implies that A may be covered by at most K^{ℓ} cosets of some subspace of size at most $K^{\ell}|A|$. This shows NLFR ⇒ PFR.

5.2 The polynomial Freiman Ruzsa conjecture implies the nearly-linear one

For this implication, as well as for proofs that appear later on, we will need Ruzsa's covering lemma, appearing in [Tau and Vu, 2006] as Lemma 2.14.

Lemma 5.2. *[Ruzsa's covering lemma] Let A, B be subsets of an abelian group such that $|A+B| \leq K|A|$. Then there is a set $X \subseteq B$, $|X| \leq K$, such that $B \subseteq A - A + X$.*

This powerful lemma has a short and elegant proof, which we bring here for the sake of completeness.

Proof of Lemma 5.2. Pick a maximal set $X \subseteq B$ such that the sets $A + x,\ x \in X$, are pairwise disjoint. Since $\bigcup_{x \in X}(A + x) \subseteq A + B$, we have that $|A||X| \leq K|A|$, which implies that $|X| \leq K$. Suppose that $b \in B$. By maximality there must be some $x \in X$ such that $(A + b) \cap (A + x) \neq \emptyset$, which means that $b \in A - A + X$. □

To show PFR $\Rightarrow$ NLFR let $\rho > 0$ be the constant stated in NLFR (Conjecture 2.20) and let r be the constant guaranteed by PFR (Conjecture 2.14). Choose the integer ℓ referred to in NLFR to be the smallest power of 2 satisfying

$$\left(\frac{1}{1+\rho/(4r)}\right)^{\log(\ell)} \leq \frac{\rho}{r}. \tag{6}$$

We say that a set B *expands under addition with respect to* A if

$$\frac{|B+B|}{|A|} \geq \left(\frac{|B|}{|A|}\right)^{1+\rho/(4r)}. \tag{7}$$

The idea of the proof is the following: for every integer $1 \leq t \leq \log(\ell)$ we check whether $B_t := 2^t A$ expands under addition with respect to A. The proof splits into two cases. The first is the case in which B_t expands under addition with respect to A for all t, namely the size of $B_{t+1} = B_t + B_t$ is large compared to the size of B_t for all t. In this case we shall see that the size of $B_1 = 2A$ is very small compared to the size of $B_{\log(\ell)} = \ell A$. Applying PFR to the set A we conclude that it can be covered by a few cosets of a small subspace. The second case is the case in which there exists some integer t for which B_t does not expand under addition with respect to A. In this case we have that $B_{t+1} = B_t + B_t$ is not too large compared to the size of B_t. Applying PFR to the set B_t together with Ruzsa's covering lemma we conclude that in this case too A can be covered by a few cosets of a small subspace. Details follow.

Case I — All sets B_t expand under addition with respect to A: Equation (7) applied to $t = 1 \ldots \log \ell$ gives

$$\frac{|B_1|}{|A|} \leq \left(\frac{|B_2|}{|A|}\right)^{\frac{1}{1+\rho/(4r)}} \leq \left(\frac{|B_3|}{|A|}\right)^{\left(\frac{1}{1+\rho/(4r)}\right)^2} \leq \ldots \leq \left(\frac{|B_{\log(\ell)}|}{|A|}\right)^{\left(\frac{1}{1+\rho/(4r)}\right)^{\log(\ell)}}$$

The assumption $|\ell A| \leq K|A|$ gives

$$\frac{|2A|}{|A|} \leq \left(\frac{|\ell A|}{|A|}\right)^{\left(\frac{1}{1+\rho/(4r)}\right)^{\log(\ell)}} \leq K^{\left(\frac{1}{1+\rho/(4r)}\right)^{\log(\ell)}} \leq K^{\rho/r}$$

where the last inequality is due to our choice of ℓ in Equation (6).

We conclude that in this case $|2A| \leq K^{\rho/r}|A|$. Applying PFR (Conjecture 2.14) we conclude that A may be covered by K^{ρ} cosets of some subspace of size at most $K^{\rho}|A|$, and this shows PFR $\Rightarrow$ NLFR with even better parameters than stated in NLFR (Conjecture 2.20).

Case II — There exists B_t which does not expand under addition with respect to A: For this t we have

$$|B_t + B_t| \leq \left(\frac{|B_t|}{|A|}\right)^{\rho/(4r)} |B_t| \tag{8}$$

Applying PFR (Conjecture 2.14) to the set B_t we conclude that it may be covered by $\left(\frac{|B_t|}{|A|}\right)^{\rho/4}$ cosets of a subspace L of size at most $\left(\frac{|B_t|}{|A|}\right)^{\rho/4}|B_t|$. By the pigeonhole principle there exists a subset $\tilde{A} \subseteq B_t$ which is contained in an affine shift of L — denote this shift by $a+L$ — such that

$$|\tilde{A}| \geq \left(\frac{|B_t|}{|A|}\right)^{-\rho/4}|B_t| \tag{9}$$

and

$$|L| \leq \left(\frac{|B_t|}{|A|}\right)^{\rho/4}|B_t| \leq \left(\frac{|B_{\log(\ell)}|}{|A|}\right)^{1+\rho/4}|A| = \left(\frac{|\ell A|}{|A|}\right)^{1+\rho/4}|A| \leq K^{1+\rho/4}|A|. \tag{10}$$

The last inequality follows from our assumption that $|\ell A| \leq K|A|$. We shall apply Ruzsa's Covering Lemma 5.2 with the sets A and $\tilde{A}$, so we compute

$$\begin{aligned}
|A+\tilde{A}| &\leq |A+B_t| \quad \text{(since } \tilde{A} \subseteq B_t\text{)} \\
&\leq |B_{t+1}| \quad \text{(since } A \subseteq a+B_t \text{ for some } a \in \mathbb{F}_2^n\text{)} \\
&\leq \left(\frac{|B_t|}{|A|}\right)^{\rho/(4r)}|B_t| \quad \text{(by Equation (8))} \\
&\leq \left(\frac{|B_t|}{|A|}\right)^{\rho/(4r)}\left(\frac{|B_t|}{|A|}\right)^{\rho/4}|\tilde{A}| \quad \text{(by Equation (9))} \\
&\leq \left(\frac{|B_t|}{|A|}\right)^{\rho/2}|\tilde{A}| \\
&\leq \left(\frac{|\ell A|}{|A|}\right)^{\rho/2}|\tilde{A}| \\
&\leq K^{\rho/2}|\tilde{A}| \quad \text{(by the assumption } |\ell A| \leq K|A|\text{)}
\end{aligned}$$

Ruzsa's covering lemma (Lemma 5.2) now implies the existence of a set $X \subseteq A$ of size at most $K^{\rho/2}$ such that

$$A \subseteq X + \tilde{A} - \tilde{A} \subseteq X + (a+L) - (a+L) = X + L$$

Concluding, in this case we have that A may be covered by at most $K^{\rho/2}$ cosets of the subspace L, where $|L| \leq K^{1+\rho/4}|A|$ (Equation (10)). Finally, if we write L as a direct sum of subspaces L' and L'', where L'' is a subspace of size $K^{\rho/4}$, and let $X' = X + L''$, we get that A may be covered by at most $K^{3\rho/4}$ cosets of the subspace L', where $|L'| \leq K|A|$ (the cosets are of the form $x' + L'$ where $x' \in X'$).

5.3 The approximate duality conjecture implies the weak polynomial Freiman Ruzsa conjecture

To prove this implication we need to recall the definition of the spectrum of a set given in Definition 4.1. Our proof uses the following lemma from Tau and Vu [2006] (appearing there as Lemma 4.38) which shows that a set having a small sum set must have large spectrum:

Lemma 5.3. *[Small sumset forces large spectrum] Let A be a subset of a finite abelian group Z, and let $0 < \epsilon \leq 1$. Then we have the following lower bound on the sum set:*

$$|A - A| \geq \frac{|A||Z|}{|A||\mathrm{spec}_\epsilon(A)| + |Z|\epsilon^2}$$

We shall also need the following easy consequence of Ruzsa's Covering lemma (Lemma 5.2):

Lemma 5.4. *[Covering] Suppose that $A \subset \mathbb{F}_2^n$ is a subset with the property that $|A + A| \leq K|A|$. Suppose furthermore that there exists a subset A' of A of size at least $\frac{1}{K_1}|A|$, such that $|\mathrm{span}\,(A')| \leq K_2|A|$. Then A may be covered by at most KK_1 cosets of a subspace of size at most $K_2|A|$.*

Proof. We apply Ruzsa's covering lemma to the sets A' and A:

$$|A + A'| \leq |A + A| \leq K|A| \leq KK_1|A'|$$

Hence Ruzsa's covering lemma implies the existence of a subset X of size at most KK_1 such that $A \subseteq X + A' - A'$. The proof is completed by noticing that $A' - A'$ is contained in a subspace of size at most $K_2|A|$. □

The idea of the proof of ADC ⇒ wPFR is as follows. Lemma 5.4 implies that it is enough to prove that if A has a small sumset then there exists a large subset A' of A which has small span. Suppose that A has a small sum set. Then Lemma 5.3 implies that A has large spectrum, denote the spectrum set by B. Assuming the approximate-duality conjecture, we have that A and B contain large subsets A', B' respectively which lie in affine shifts of dual subspaces. But this implies in turn that $\dim(A') \leq n - \dim(B')$, i.e. A' has a small span, and setting the parameters correctly we arrive at the desired result.

Let ζ be the constant guaranteed by Conjecture 2.15 for the constants $\alpha/2$ and δ. Our goal will be to show that A may be covered by at most $2^{\delta n} \cdot K$ cosets of a subspace of size at most $2^{\delta n}K^r$, where $r := \max\left\{\frac{1}{\zeta}, \frac{4}{\alpha}\right\}$.

First we observe that without loss of generality we may assume that

$$K \leq \min\left\{2^{\alpha n/4}, 2^{\zeta n}\right\} \tag{11}$$

since otherwise from our choice of r we have that $K^r \geq 2^n$, and hence the desired conclusion holds trivially.

Next, in Lemma 5.3 set $\epsilon = 1/K$. Then from the lemma and the assumption that $|A + A| \leq K|A|$ we have:

$$K|A| \geq |A - A| \geq \frac{|A|2^n}{|A||\mathrm{spec}_{1/K}(A)| + 2^n K^{-2}}$$

And rearranging we obtain:

$$|\mathrm{spec}_{1/K}(A)| \geq \frac{2^n}{|A|K}\frac{K-1}{K} \geq \frac{2^n}{|A|K^2} \tag{12}$$

We would like to apply ADC (Conjecture 2.15) to the sets A and $\mathrm{spec}_{1/K}(A)$. Obviously, $\mu^\perp(A, \mathrm{spec}_{1/K}(A)) \geq 1/K$, where $1/K \geq 2^{-\zeta n}$ (Equation (11)). Also, our choice of $K \leq 2^{\alpha n/4}$ in (11) and our assumption that $|A| \leq 2^{(1-\alpha)n}$, together with Equation (12) imply that

$$|\mathrm{spec}_{1/K}(A)| \geq \frac{2^n}{|A|K^2} \geq \frac{2^n}{2^{(1-\alpha)n}2^{2\alpha n/4}} = 2^{\alpha n/2}$$

Obviously, we also have that A is of size at least $2^{\alpha n/2}$. Hence ADC (Conjecture 2.15) implies the existence of subsets $A' \subseteq A$, $B' \subseteq \mathrm{spec}_{1/K}(A)$ which lie in affine shifts of dual spaces such that

$$|A'| \geq \frac{|A|}{2^{\delta n}}, \quad |B'| \geq \frac{|\mathrm{spec}_{1/K}(A)|}{2^{\delta n}}.$$

But this implies in turn that $\dim(A') + \dim(B') \leq n$, and consequently

$$|\mathrm{span}\left(A'\right)| \leq \frac{2^n}{|B'|} \leq \frac{2^{\delta n} \cdot 2^n}{|\mathrm{spec}_{1/K}(A)|} \leq 2^{\delta n} K^2 |A|$$

where the last inequality is due to Equation (12).

Concluding, we have that $|\mathrm{span}\,(A')| \leq 2^{\delta n} K^2 |A|$ where A' is a subset of A of size at least $|A|/2^{\delta n}$. Using Lemma 5.4 we conclude A can be covered by $\leq 2^{\delta n} \cdot K$ cosets of $\mathrm{span}\,(A')$ which is of size $\leq 2^{\delta n} K^2 |A|$ and this completes the proof of ADC $\Rightarrow$ wPFR.

5.4 The polynomial Freiman Ruzsa conjecture implies the approximate duality conjecture

Our proof of the implication PFR $\Rightarrow$ ADC uses the following lemma, which shows that whenever $A \subseteq \mathrm{spec}_\epsilon(B)$ for sufficiently large ϵ, and assuming the PFR conjecture, we can find a large set $A' \subset A$ and a set $B' \subset B$ such that $\mu^\perp(A, B') = 1$ and, most importantly, the size of B' is proportional to $\frac{|\mathrm{span}(A')|}{|\mathrm{span}(A)|}$. This last property is important because it allows us to make the following iterative argument: Either $\mathrm{span}\,(A')$ is large relatively to $\mathrm{span}\,(A)$ in which case B' is also large and we have proved the ADC, or $\mathrm{span}\,(A')$ is small relatively to $\mathrm{span}\,(A)$ and then we apply the lemma again with A' instead of A. We prove that this process terminates eventually, and that when it terminates the sets A and B' are the desired sets we need.

Lemma 5.5. *[ADC as function of span] For every constants $1 > \delta' > 0$, $1 > \alpha' > 0$ there exist a constant $\zeta' > 0$ and an integer k, both depend only on δ', α', such that the following holds for sufficiently large n assuming the PFR conjecture. If $A, B \subseteq \mathbb{F}_2^n$ satisfy $|A| \geq 2^{\alpha' n}$ and $A \subseteq \mathrm{spec}_\epsilon(B)$ for $\epsilon \geq 2^{-\zeta' n}$, then there exist subsets $A' \subseteq A$ and $B' \subseteq B$ satisfying*

1. $|A'| \geq |A|^{1-\delta'}$.
2. $|B'| \geq \epsilon^{2k} \frac{1}{8} \frac{|\mathrm{span}(A')|}{|\mathrm{span}(A)|} |A|^{-\delta'} |B|$.
3. $\mu^\perp(A, B') = 1$.

The proof of the above lemma is deferred to the next subsection. To prove the implication NLFR $\Rightarrow$ ADC we need to set a few parameters that will be used later on. Choose $\delta' < \delta^2/4$ and $\alpha' < \alpha - \delta$ in Lemma 5.5, and let $\zeta' > 0$ and k be the constant and the integer guaranteed by this lemma for the constants δ', α'. Suppose that $\epsilon \geq 4 \cdot 2^{-\zeta n}$, where $\zeta = \min\{\zeta', \delta/4k\}$. We assume $\mathbb{E}_{a \in A, b \in B}\left[(-1)^{\langle a,b \rangle}\right]$ is positive, the proof for the negative case is similar.

Now we describe the iterative process. Start with the set $A_0 := A \cap \mathrm{spec}_{\epsilon/2}(B)$, which, by Markov's inequality is of size at least $|A|/2$. For $i = 0, 1, \ldots$ let $A_i' \subseteq A_i, B_i' \subseteq B$ be the subsets guaranteed by Lemma 5.5 with respect to A_i, B. Let

$$\sigma_i = \frac{|\mathrm{span}\,(A_i')|}{|\mathrm{span}\,(A_i)|} |A_i|^{-\delta'}. \tag{13}$$

While $\sigma_i \leq 2^{-(\delta/2)n}$ we set $A_{i+1} := A'_i$. Let t be the first time $\sigma_t > 2^{-(\delta/2)n}$. We argue that such a t exists and that in fact $t < 4/\delta$. Indeed, if $\sigma_i \leq 2^{-(\delta/2)n}$ we have that

$$|\mathrm{span}\left(A'_i\right)| \leq 2^{-(\delta/2)n}|A_i|^{\delta'}|\mathrm{span}\left(A_i\right)| \leq 2^{-(\delta/2-\delta')n}|\mathrm{span}\left(A_i\right)| \leq 2^{-(\delta/4)n}|\mathrm{span}\left(A_i\right)| \tag{14}$$

where the last inequality is due to our choice of $\delta' < \delta^2/4$. Since $1 \leq |\mathrm{span}\left(A'_i\right)|$ and $|\mathrm{span}\left(A_i\right)| \leq 2^n$, the above equation implies that $t \leq 4/\delta$.

We next show that the set A_t is of pretty large size. The first bullet of Corollary 5.5 implies that for all $i \leq t$

$$|A'_i| \geq |A_i|^{1-\delta'} \geq 2^{-\delta' n}|A_i| \geq 2^{-(\delta^2/4)}|A_i|$$

where the last inequality is again due to our choice of $\delta' < \delta^2/4$. Consequently, our assumption that $|A| \geq 2^{\alpha n}$ together with the fact that $t \leq 4/\delta$ imply for sufficiently large n that

$$|A_t| \geq 2^{-\delta n}|A|/2 \geq 2^{(\alpha-\delta)n-1} \geq 2^{\alpha' n}$$

where the last inequality is due to our choice of $\alpha' < \alpha - \delta$.

Applying Lemma 5.5 one final time with A_t and B and using the assumption $\sigma_t > 2^{-(\delta/2)n}$ we conclude the existence of $B' \subseteq B, |B'| > \left(\frac{\epsilon}{2}\right)^{2k} \frac{1}{8} 2^{-(\delta/2)n}|B| \geq 2^{-\delta n}|B|$ (the last inequality is due to our choice of $\epsilon > 4 \cdot 2^{-(\delta/4k)n}$) such that $\mu^{\perp}(A_t, B') = 1$. So A_t and B' are the two sets promised by ADC and this shows PFR $\Rightarrow$ ADC.

5.5 Proof of main technical lemma

In this section we prove the main technical lemma used in the proof of PFR $\Rightarrow$ ADC, Lemma 5.5. The proof breaks down to two lemmas (Lemmas 5.6 and 5.7) stated next. The first lemma can be seen as a version of the ADC which applies when $|\mathrm{span}\left(A\right)|$ is not much larger than $|A|$.

Lemma 5.6. *[Approximate-duality for sets with small span] Given $B \subseteq \mathbb{F}_2^n$ and $A \subseteq \mathrm{spec}_{\epsilon}(B)$, there exists a subset B' of B of size at least $\epsilon^2 \frac{1}{2} \frac{|A|}{|\mathrm{span}(A)|}|B|$ such that $\mu^{\perp}(A, B') = 1$.*

The second main step in the proof is to show that if $A \subseteq \mathrm{spec}_{\epsilon}(B)$, then assuming the PFR conjecture, $\tilde{A} := \mathrm{spec}_{\epsilon^k/2}(B) \cap \mathrm{span}\left(A\right)$ ia large for some constant k. By showing this we will be able to apply the above lemma also to sets that have large span relative to their size, by applying it to the sets $\tilde{A}$ and B. A lower bound on the on the size of $\tilde{A}$ is given by the following lemma:

Lemma 5.7. *For every $1 > \delta' > 0$, $1 > \alpha' > 0$, there exist a constant $\zeta' > 0$ and an integer k, both depend only on δ', α', such that the following holds for sufficiently large n, assuming the PFR conjecture holds. Given $B \subseteq \mathbb{F}_2^n$ and $A \subseteq \mathrm{spec}_{\epsilon}(B), |A| \geq 2^{\alpha' n}$ where $\epsilon \geq 2^{-\zeta' n}$, there exists a subset A' of A of size at least $|A|^{1-\delta'}$ such that*

$$|\mathrm{span}\left(A'\right)||A|^{-\delta'} \leq |\mathrm{span}\left(A\right) \cap \mathrm{spec}_{\epsilon^k/2}(B)|. \tag{15}$$

Given these two lemmas we can complete the proof of Lemma 5.5. Then we prove the two lemmas.

Proof of Lemma 5.5. Noticing the assumptions of Lemma 5.5 and Lemma 5.7 are the same, let A' be the subset of A which is of size at least $|A|^{1-\delta'}$ and satisfies Equation (15). Notice A' satisfies bullet 1 of Lemma 5.5.

Let $\tilde{A} := \mathrm{span}\,(A) \cap \mathrm{spec}_{\epsilon^k/2}(B)$. Apply Lemma 5.6 to $\tilde{A}, B$ and conclude the existence of $B' \subseteq B$ such that $\mu^{\perp}(\tilde{A}, B')$, and which satisfies

$$|B'| \geq \epsilon^{2k}\frac{1}{8}\frac{|\tilde{A}|}{\left|\mathrm{span}\left(\tilde{A}\right)\right|} \cdot |B| \geq \epsilon^{2k}\frac{1}{8}\frac{|\mathrm{span}\,(A')\,||A|^{-\delta'}}{\left|\mathrm{span}\left(\tilde{A}\right)\right|} \cdot |B|.$$

The last inequality above uses Equation (15). To show that B' satisfies bullets 2 and 3 of Lemma 5.5 notice $\tilde{A} \supseteq A$ because $A \subseteq \mathrm{spec}_{\epsilon}(B) \subseteq \mathrm{spec}_{\epsilon^k}(B)$ which implies $\mathrm{span}\left(\tilde{A}\right) = \mathrm{span}\,(A)$ and, also, $\mu^{\perp}(A, B') = \mu^{\perp}(\tilde{A}, B)$. This completes the proof. □

For the proof of Lemma 5.6 we shall use Fourier analysis and recall the standard notations for it. For a function $f : \mathbb{F}_2^n \to \mathbb{C}$ and $\alpha \in \mathbb{F}_2^n$ we denote by $\hat{f}(\alpha)$ the α-coefficient of the Fourier expansion of f over $\mathbb{F}_2^n$, defined by

$$\hat{f}(\alpha) = \mathbb{E}_{\beta \in \mathbb{F}_2^n}\left[f(\beta)(-1)^{\langle \beta, \alpha \rangle}\right].$$

We shall need Parseval's equality which says that for a function $f : \mathbb{F}_2^n \to \mathbb{C}$,

$$\sum_{\alpha \in \mathbb{F}_2^n} (\hat{f}(\alpha))^2 = 2^{-n} \sum_{\beta \in \mathbb{F}_2^n} (f(\beta))^2. \tag{16}$$

Proof of Lemma 5.6. Let $d := \dim(A)$ and choose an arbitrary basis $a_1, a_2, \ldots, a_d$ of A. For every vector $\beta = (\beta_1, \beta_2, \ldots, \beta_d) \in \mathbb{F}_2^d$, we denote by S_β the following coset of $A^{\perp}$:

$$S_\beta = \{\gamma \in \mathbb{F}_2^n | \langle a_i, \gamma \rangle = \beta_i \text{ for all } i = 1, 2, \ldots, d\}$$

For every $\beta \in \mathbb{F}_2^d$ we denote the relative weight of B inside S_β by:

$$w(\beta) = Pr_{b \in B}[b \in S_\beta] = \frac{|B \cap S_\beta|}{|B|}$$

Our goal will be to show the existence $\beta \in \mathbb{F}_2^d$ such that $w(\beta) \geq \epsilon^2 \frac{|A|}{|\mathrm{span}(A)|}$, since in this case $\hat{B} := B \cap S_\beta$ is a subset of B of size at least $\epsilon^2 \frac{|A|}{|\mathrm{span}(A)|}|B|$ which is contained in an affine coset of $A^{\perp}$ of the form $a + A$. Partition $\hat{B}$ into $\hat{B}_0 = \left\{b \in \hat{B} \mid \langle b, a \rangle = 0\right\}$ and $\hat{B}_1 = \left\{b \in \hat{B} \mid \langle b, a \rangle = 1\right\}$. To complete the proof of the lemma take B' to be the larger of $\hat{B}_0, \hat{B}_1$ and notice $|B'| \geq \epsilon^2 \frac{1}{2} \frac{|A|}{|\mathrm{span}(A)|}|B|$.

Our main observation is that for every $a \in A$, if we write $a = \sum_{i=1}^d \alpha_i a_i$, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$, then we have

$$\hat{w}(\alpha) = \mathbb{E}_{\beta \in \mathbb{F}_2^d}\left[w(\beta)(-1)^{\langle \beta, \alpha \rangle}\right] = 2^{-d} \sum_{\beta \in \mathbb{F}_2^d} w(\beta)(-1)^{\langle \beta, \alpha \rangle} = 2^{-d}\mathbb{E}_{b \in B}\left[(-1)^{\langle a, b \rangle}\right] \tag{17}$$

The equality above allows us to prove the lemma by bounding the sum $\sum_{\alpha \in \tilde{A}} (\hat{w}(\alpha))^2$ from above and from below, where:

$$\tilde{A} = \left\{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbb{F}_2^d | \sum_{i=1}^d \alpha_i a_i \in A\right\}$$

For obtaining the lower bound we use Equation (17) together with our assumption that $\mu^{\perp}(A,B) \geq \epsilon$, while for the upper bound we use Parseval's equality together with the fact that w is a distribution, i.e., $\sum_{\beta \in \mathbb{F}_2^d} w(\beta) = 1$. We start with bounding the sum $\sum_{\alpha \in \tilde{A}} (\hat{w}(\alpha))^2$ from below

$$\begin{aligned}
\sum_{\alpha \in \tilde{A}} (\hat{w}(\alpha))^2 &\geq |\tilde{A}| \Big(\mathbb{E}_{\alpha \in \tilde{A}}[\hat{w}(\alpha)] \Big)^2 \quad \text{(by convexity)} \\
&= |\tilde{A}| \Big(2^{-d} \mathbb{E}_{a \in A} \mathbb{E}_{b \in B} \big[(-1)^{\langle a,b \rangle} \big] \Big)^2 \quad \text{(by (17))} \\
&= |A| 2^{-2d} \big(\mu^{\perp}(A,B) \big)^2 \geq |A| 2^{-2d} \epsilon^2 \qquad (18)
\end{aligned}$$

Next we bound the sum $\sum_{\alpha \in \tilde{A}} (\hat{w}(\alpha))^2$ from above:

$$\begin{aligned}
\sum_{\alpha \in \tilde{A}} (\hat{w}(\alpha))^2 &\leq \sum_{\alpha \in \mathbb{F}_2^d} (\hat{w}(\alpha))^2 \qquad (19) \\
&= 2^{-d} \sum_{\beta \in \mathbb{F}_2^d} (w(\beta))^2 \quad \text{(by Parseval's Equality)} \\
&\leq 2^{-d} \max_{\beta \in \mathbb{F}_2^d} w(\beta) \sum_{\beta \in \mathbb{F}_2^d} w(\beta) \\
&= 2^{-d} \max_{\beta \in \mathbb{F}_2^d} w(\beta) \quad (\textstyle\sum_{\beta \in \mathbb{F}_2^d} w(\beta) = 1) \qquad (20)
\end{aligned}$$

Finally, the combination of equations (18) and (20) implies the existence of $\beta \in \mathbb{F}_2^d$ such that

$$w(\beta) \geq |A| 2^{-d} \epsilon^2 = \epsilon^2 \frac{|A|}{|\mathrm{span}\,(A)\,|}$$

which finishes the proof of the lemma. □

For the proof of Lemma 5.7, which follows next, we shall apply the NLFR conjecture, together with an analogous nearly-linear version of the Balog-Szemerédi-Gowers (BSG) theorem. The BSG theorem [Balog and Szemerédi, 1994, Gowers, 1998] says that if A is a subset of an abelian additive group which satisfies $\Pr_{a,a' \in A}[a + a' \in S] > 1/K$ for $|S| \leq C|A|$, then one can find a subset $A' \subseteq A$, such that $|A'| \geq |A|/f(K,C)$, and $|A' + A'| \leq g(K,C)|A|$, where $f(K,C)$ and $g(K,C)$ are polynomials with fixed degrees in K and C.

For our purposes we need that the exponents in the polynomials $f(K,C)$ and $g(K,C)$ would be as small as possible, in particular we would be interested in a bound of the form $C^{\delta}K^d$ for some small constant $\delta > 0$, and an integer d. As was the case with the NLFR conjecture, to achieve this, we are willing to assume not only that $\Pr_{a,a' \in A}[a + a' \in S]$ is large but even that $\Pr_{a_1,a_2,\ldots,a_k \in A}[a_1 + a_2 + \ldots + a_k \in S]$ is large for some constant $k > 2$. For this we need a nearly-linear version of the BSG theorem. Fortunately, such a theorem was proved by Croot and Borenstein [2008].

Theorem 5.8. *[Nearly-linear BSG Theorem, Croot and Borenstein [2008]] For every $1 > \epsilon > 0$ and $c > 1$, there exist a constant η, and an integer k, both depend only on ϵ, c, where ζ is monotonically decreasing in*

c and k is monotonically increasing in c, such that the following holds. If A is a sufficiently large subset of an additive abelian group which satisfies:

$$Pr_{a_1\in A,a_2\in A,\ldots,a_k\in A}\left[\sum_{i=1}^{k} a_i \in S\right] \geq |A|^{-\zeta},\ |S| \leq |A|^c$$

then there exists a subset A' of A of size at least $|A|^{1-\epsilon}$ such that:

$$|\ell A'| \leq |A'|^{c(1+\epsilon\ell)}$$

.

The combination of the above theorem with the NLFR conjecture gives the following corollary:

Corollary 5.9. *[Nearly-linear BSG + NLFR] For every $1 > \delta > 0$ and $c > 1$, there exist a constant $\zeta > 0$ and an integer k, both depend only on δ, c, where ζ is monotonically decreasing in c and k is monotonically increasing in c, and such that the following holds assuming the PFR conjecture. If A is a sufficiently large subset of an additive abelian group which satisfies:*

$$Pr_{a_1\in A,a_2\in A,\ldots,a_k\in A}\left[\sum_{i=1}^{k} a_i \in S\right] \geq |A|^{-\zeta},\ |S| \leq |A|^c$$

then there exists a subset A' of A of size at least $|A|^{1-\delta}$ such that:

$$|\mathrm{span}\left(A'\right)| \leq |A|^{c+\delta}$$

Proof. Let $\rho \leq \delta/(4c)$, and let ℓ be the integer guaranteed by Conjecture 2.20 for the constant ρ. Let $\epsilon \leq \delta/(2c\ell)$, and let k and ζ be the integer and the constant guaranteed by Theorem 5.8 for the constants c and ϵ.

Noticing that the assumptions in Theorem 5.8 and Corollary 5.9 are the same, we have that there exists a subset $A' \subseteq A$ of size at least $|A|^{1-\epsilon}$ such that $|\ell A'| \leq |A'|^{c(1+\epsilon\ell)}$. Applying Conjecture 2.20 with the set A' and $K \leq |A'|^{c(1+\epsilon\ell)-1}$ we get that there exists a subset $A'' \subseteq A'$ of size at least $K^{-\rho}|A'|$ such that $|\mathrm{span}\,(A'')| \leq K^{1+\rho}|A'|$. To conclude the proof we compute the sizes of A'' and $\mathrm{span}\,(A'')$:

$$\begin{aligned}
|A''| &\geq K^{-\rho}|A'| \\
&\geq |A'|^{-\rho c(1+\epsilon\ell)}|A'| \quad (K \leq |A'|^{c(1+\epsilon\ell)}) \\
&\geq |A'|^{-\delta/4(1+1)}|A'| \quad (\,\rho \leq \delta/(4c), \epsilon \leq 1/\ell) \\
&= |A'|^{1-\delta/2} \\
&\geq |A|^{(1-\delta/2)(1-\epsilon)} \quad (\,|A'| \geq |A|^{1-\epsilon}) \\
&\geq |A|^{(1-\delta/2)^2} \quad (\,\epsilon \leq \delta/2) \\
&\geq |A|^{1-\delta}
\end{aligned}$$

$$\begin{aligned}
|\mathrm{span}\left(A''\right)| &\leq K^{1+\rho}|A'| \\
&\leq |A'|^{c(1+\epsilon\ell)+\rho c(1+\epsilon\ell)} \quad (\, K \leq |A'|^{c(1+\epsilon\ell)-1}) \\
&\leq |A'|^{c(1+\epsilon\ell)+\delta/2} \quad (\rho \leq \delta/(4c),\, \epsilon \leq 1/\ell) \\
&\leq |A'|^{c+\delta} \quad (\, \epsilon \leq \delta/(2c\ell)) \\
&\leq |A|^{c+\delta} \quad (\, A' \subseteq A)
\end{aligned}$$

So we have that $|A''| \geq |A|^{1-\delta}$ and $|\mathrm{span}\,(A'')| \leq |A|^{c+\delta}$ which concludes the proof of the corollary. □

We now proceed to the proof of Lemma 5.7:

Proof of Lemma 5.7. Let $\delta := \delta'$, and $c = 1/\alpha'$. Let $\zeta > 0$ and k be the constant and the integer guaranteed by Corollary 5.9 for the constants δ and c.

We may assume that k is even (if k is odd replace it by $k+1$ and the proof goes through). From our assumption that $A \subseteq \mathrm{spec}_\epsilon(B)$ and using convexity we get

$$\begin{aligned}
\epsilon^k &\leq \left(\mathbb{E}_{b\in B}\mathbb{E}_{a\in A}\left[(-1)^{\langle a,b\rangle}\right]\right)^k \\
&\leq \mathbb{E}_{b\in B}\left(\mathbb{E}_{a\in A}\left[(-1)^{\langle a,b\rangle}\right]\right)^k \\
&= \mathbb{E}_{b\in B}\mathbb{E}_{a_1\in A,a_2\in A,\ldots,a_k\in A}\left[(-1)^{\langle \sum_{i=1}^k a_i,b\rangle}\right]
\end{aligned}$$

Markov's inequality implies

$$Pr_{a_1\in A,a_2\in A,\ldots,a_k\in A}\left[\sum_{i=1}^k a_i \in \mathrm{spec}_{\epsilon^k/2}(B)\cap \mathrm{span}\,(A)\right] \geq \frac{\epsilon^k}{2}$$

Let $S := \mathrm{spec}_{\epsilon^k/2}(B)\cap \mathrm{span}\,(A)$, and let $c' = \frac{\log|S|}{\log|A|}$. From our assumptions that $A \subseteq \mathrm{spec}_\epsilon(B)$, and $|A| \geq 2^{\alpha' n}$ we have that $1 \leq c' \leq 1/\alpha'$. Since in Corollary 5.9 k is monotonically increasing in c, and ζ is monotonically decreasing in c, we may apply this corollary with the constant c' instead of c without changing the values of k and ζ.

Let $\zeta' = \frac{\alpha'}{k}\zeta$, and suppose that $\epsilon \geq 2^{-\zeta' n}$. From our choice of ϵ and the assumption that $|A| \geq 2^{\alpha' n}$ we have

$$\epsilon^k \geq 2^{-\zeta' k n} \geq 2^{-\frac{\alpha'}{k}\zeta k n} \geq |A|^{-\zeta}$$

We conclude that

$$Pr_{a_1\in A,a_2\in A,\ldots,a_k\in A}\left[\sum_{i=1}^k a_i \in S\right] \geq \frac{1}{2}|A|^{-\zeta}$$

and in addition

$$|S| = |A|^{c'}$$

Corollary 5.9 applies, and we conclude that there exists a subset A' of A of size at least $|A|^{1-\delta'}$ such that

$$|\mathrm{span}\left(A'\right)| \leq |A|^{c'+\delta'} = |S||A|^{\delta'}$$

This completes the proof of the lemma.

□

6 On the ℓ_∞-error of multi-output bit affine and two-source extractors

In this section we explain our use of ℓ_∞ as the measure of error in the definition of extractors (Definition 2.1). In a nutshell, it yields cleaner and tighter analysis than we would have obtained using statistical distance as our measure of error. And it allows us to construct affine and (in Section 6.2) two-source extractors with multiple output bits with essentially no loss in ℓ_∞-error. Details follow.

6.1 On the ℓ_∞-error of existing affine extractors

All known affine extractors, i.e., those of Bourgain [2007], Yehudayoff [2009], Ben-Sasson and Kopparty [2009], Li [2010] have the following property. Each of them is defined as evaluating a certain r-variate polynomial P over a finite field $\mathbb{F}_{2^m}$, where $n = r \cdot m$. The n-input bits are viewed as describing an input $\beta = (\beta_1, \ldots, \beta_r) \in \mathbb{F}_{2^m}^r$. And each of these constructions shows a bound on the error of any nontrivial character applied to $P(\beta)$. Recall that a nontrivial additive character $\chi_\alpha : \mathbb{F}_{2^m} \to \{-1, 1\}$ of $\mathbb{F}_{2^m}$ is a function of the form

$$\chi_\alpha(x) = (-1)^{\sum_{j=1}^m \alpha_j \cdot x_j},$$

where $(x_1, \ldots, x_m)$ is the representation of x according to an arbitrary fixed $\mathbb{F}_2$-basis for $\mathbb{F}_{2^m}$. In other words, for each of the known constructions of affine extractors we have a result of the following form. For every nontrivial character χ_α as above, and every $\mathbb{F}_2$-affine subspace A of $(\mathbb{F}_{2^m})^r$ of dimension at least d, we have

$$|\mathbb{E}_{x \in A}\left[\chi_\alpha(P(x))\right]| \leq \epsilon.$$

We point out that Vazirani's "XOR-lemma" shows that extractors with error bounds as above are also $\epsilon \cdot 2^{m/2}$-close to uniform in statistical distance. This can be converted back to a bound on the ℓ_∞-error of the form $\epsilon \cdot 2^{m/2}$. But using the lemma below we can deduce that the ℓ_∞-error is merely ϵ, i.e., we lose literally nothing from outputting m bits instead of a single bit.

Lemma 6.1. *[Multi-output extractors] Let ζ be a distribution on $\mathbb{F}_2^m$ satisfying for every nontrivial additive character*

$$|\mathbb{E}[\chi_\alpha(x)]| \leq \epsilon, \tag{21}$$

where x is distributed according to ζ. Then for any linearly independent $\alpha_1, \ldots, \alpha_t \in \mathbb{F}_2^m$ and any $b_1, \ldots, b_t \in \mathbb{F}_2$, denoting by S_b the affine space

$$S_b = \{x \in \mathbb{F}_2^m \mid \langle \alpha_1, x\rangle = b_1, \ldots, \langle \alpha_t, x\rangle = b_t\},$$

we have

$$2^{-t} - \epsilon < \zeta(S_b) < 2^{-t} + \epsilon.$$

Consequently, taking $t = m$ and $\alpha_1, \alpha_2, \ldots, \alpha_m$ to be the standard basis and noticing that in this case we have $\zeta(S_b) = \Pr_{X \sim \zeta}[X = b]$ we conclude the ℓ_∞-distance between ζ and the uniform distribution is at most ϵ.

Proof of Lemma 6.1. Consider a vector α of the form $\alpha = \sum_{i=1}^{t} a_i \alpha_i$ where $a_i \in \mathbb{F}_2$. Let a denote the vector $(a_1, \ldots, a_t)$. We have that

$$\langle \alpha, x \rangle = \sum_{i=1}^{t} a_i \cdot \langle \alpha_i, x \rangle.$$

Thus, for $c = (c_1, \ldots, c_t) \in \mathbb{F}_2^t$ and $x \in S_c$ we have $\langle \alpha, x \rangle = \langle a, c \rangle$ which implies

$$\mathbb{E}[\chi_\alpha(x)] = \sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = 0} \zeta(S_c) - \sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = 1} \zeta(S_c) \tag{22}$$

which, by (21), implies that for any $\alpha \in \operatorname{span}(\alpha_1, \ldots, \alpha_t) \setminus \{0\}$ and $\beta \in \mathbb{F}_2$ we have

$$-\epsilon \leq \sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = \beta} \zeta(S_c) - \sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = 1-\beta} \zeta(S_c) \leq \epsilon. \tag{23}$$

For $\alpha = 0$ we get from (22)

$$\sum_{c \in \mathbb{F}_2^t} \zeta(S_c) = 1, \tag{24}$$

because every c satisfies $\langle 0, c \rangle = 0$. Set $\beta_a = \langle b, a \rangle$. Consider the following sum:

$$\sum_{a \in \mathbb{F}_2^t} \left(\sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = \beta_a} \zeta(S_c) - \sum_{c \in \mathbb{F}_2^t, \langle c,a \rangle = 1-\beta_a} \zeta(S_c) \right). \tag{25}$$

Using (23) and (24) we bound (25) from above by $1 + (2^t - 1) \cdot \epsilon$ (the first summand comes from $\alpha = 0$ via (24) and the remaining ones come from $\alpha \neq 0$ via (23)). Similarly, (25) is bounded from below by $1 - (2^t - 1) \cdot \epsilon$. Finally, we observe that (25) is equal to $2^t \cdot \zeta(S_b)$. The reason for this is that we have by definition $\langle b, a \rangle = \beta_a$ for all $a \in \mathbb{F}_2^t$, whereas for any fixed $c \neq b$ we have $\langle c, a \rangle = \beta_a$ if and only if $\langle c, a \rangle = \langle b, a \rangle$ which happens iff $\langle c - b, a \rangle = 0$. Since $c \neq b$ (and both b and c are fixed) this latter event happens for precisely half of the a's and thus the summand $\zeta(S_c)$ appears in (25) equally often positively as negatively and gets canceled.

We have shown that

$$1 - (2^t - 1) \cdot \epsilon \leq 2^t \cdot \zeta(S_b) \leq 1 + (2^t - 1) \cdot \epsilon,$$

and dividing this inequality by 2^t completes the proof. □

6.2 Increasing the output length of our two-source extractors

In this section we prove Lemma 2.19 and show how to obtain two source extractors with multiple output bits. Before doing so we briefly explain how a so-called collection of independent matrices can be obtained.

Let $\mathbb{F}_{2^r}$ denote the finite field with 2^r elements. It is well-known that elements of this field form a $\mathbb{F}_2$-linear space of dimension r. Let $\beta_1, \ldots, \beta_r \in \mathbb{F}_{2^r}$ be a basis for this space. Note that multiplication by any $\beta \in \mathbb{F}_2^r \setminus 0$ is an invertible $\mathbb{F}_2$-linear transformation, and let M_i be the matrix representing multiplication by β_i in our basis. It is now rather straightforward to verify that $M_1, \ldots, M_r$ are independent according to our definition.

We now proceed to prove Lemma 2.19.

Proof of Lemma 2.19. Let $\chi_\alpha : \mathbb{F}_2^t \to [-1, 1]$ be a nontrivial additive character. We have

$$\chi_\alpha(E(x, y)) = (-1)^{\sum_{i=1}^t \alpha_i \langle h_1(x), M_i h_2(y)\rangle} = (-1)^{\langle h_1(x), M h_2(y)\rangle}$$

where $M = \sum_{i=1}^t \alpha_i M_i$. Since M has full rank and $E(x, y)$ is a $[n, 1, \rho, \epsilon]$-two source extractor which is linear invariant with respect to h_1, h_2, we see that

$$|\mathbb{E}_{X,Y} [\chi_\alpha(E(X, Y))]| \leq \epsilon.$$

Applying Lemma 6.1 completes the proof. □

References

Antal Balog and Endre Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14(3):263–268, 1994.

Boaz Barak, Guy Kindler, Ronen Shaltiel, Benjamin Sudakov, and Avi Wigderson. Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In *Proc. 37th STOC*. ACM, 2005.

Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. *SIAM Journal on Computing*, 36(4):1095–1118, 2006a. Preliminary version in FOCS' 04.

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 *Proc. 38th Symposium on Theory of Computing (STOC)*, pages 671–680. ACM, 2006b.

Eli Ben-Sasson and Swastik Kopparty. Affine dispersers from subspace polynomials. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 65–74. ACM, 2009. ISBN 978-1-60558-506-2. URL `http://doi.acm.org/10.1145/1536414.1536426`.

Eli Ben-Sasson, Shlomo Hoory, Eyal Rozenman, Salil Vadhan, and Avi Wigderson. Extractors for affine sources, unpublished manuscript. 2001.

Jean Bourgain. More on the sum-product phenomenon in prime fields and its application. *International Journal of Number Theory*, 1:1–32, 2005.

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

Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, 1988. ISSN 0097-5397. doi: http://dx.doi.org/10.1137/0217015.

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.

Ernie Croot and Evan Borenstein. On a certain generalization of the Balog-Szemerèdi-Gowers theorem, June 25 2008. URL `http://arxiv.org/abs/0805.3305`.

Zeev Dvir. Extractors for varieties. In *IEEE Conference on Computational Complexity*, pages 102–113. IEEE Computer Society, 2009. ISBN 978-0-7695-3717-7.

Zeev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. *Computational Complexity*, 18(1):1–58, 2009.

Paul Erdös. Some remarks on the theory of graphs. *B.A.M.S.*, 53:292–294, 1947.

Gregory A. Freiman. *Foundations of a structural theory of set addition*, volume 37. American Mathematical Society, 1973.

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

William Timothy Gowers. A new proof of Szemerèdi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.

Ben Green. Finite field models in additive combinatorics. In *London Mathematical Society Lecture Note Series*, volume 324. Cambridge University Press, 2005a.

Ben Green. Finite field models in additive combinatorics. In Bridget S. Webb, editor, *Surveys in Combinatorics*, number 327 in London Mathematical Society Lecture Note Series, pages 1–27. Cambridge University press, 2005b.

Ben Green and Terence Tao. A note on the Freiman and Balog–Szemerédi–Gowers theorems in finite fields. *Journal of the Australian Mathematical Society*, 86(01):61–74, 2009.

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

Xin Li. A new approach to affine extractors and dispersers. *Electronic Colloquium on Computational Complexity (ECCC)*, (064), 2010. ISSN 1433-8092. URL `http://eccc.hpi-web.de/eccc-reports/2010/TR10-064/index.html`.

H. Plunnecke. Eigenschaften und abschatzungen von wirkingsfunktionen. *BMwF-GMD22 Gesellschaft fur Mathematik und Datenverarbeitung*, 1969.

Pavel Pudlák and Vojtěch Rödl. Pseudorandom sets and explicit constructions of Ramsey graphs. *Quaderni di Matematica, Dipartimanto di Matematica, Seconda Universita di Napoli, Caserta*, 13:327–346, 2004.

Anup Rao. An exposition of bourgain's 2-source extractor. In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, 2007.

Ran Raz. Extractors with weak random seeds. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2005.

Imre Z. Ruzsa. An application of graph theory to additive number theory. *Scientia, Ser. A*, 3:97–109, 1989.

Imre Z. Ruzsa. An analog of Freiman's theorem in groups. *Astèrique*, 258:323–326, 1999.

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

Terence Tau and Van Vu. *Additive Combinatorics*. Cambridge University Press, Cambridge, 2006.

Yael Tauman Kalai, Xin Li, and Anup Rao. 2-source extractors under computational assumptions and cryptography with defective randomness. In *Proc. 50th Foundations of Computer Science (FOCS)*. IEEE, 2009.

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

Amir Yehudayoff. Affine extractors over prime fields. 2009. URL `http://www.math.ias.edu/~7Eamiry/Affine.pdf`.

David Isaac Zuckerman. *Computing efficiently using general weak random sources*. PhD thesis, University of California, Berkeley, 1991.