Arithmetic progressions with a pseudorandom step

Elad Aigner-Horev Hiệp Hàn

Abstract

Let $\alpha, \sigma > 0$ and let A and S be subsets of a finite abelian group G of densities α and σ independent of $|G|$, respectively. Without additional restrictions A need not contain a 3-term arithmetic progression whose common gap is in S. What is then the least integer $k \geq 2$ for which there exists an $\eta = \eta(\alpha, \sigma)$ such that $\|S\|_{U^k(G)} \leq \eta$ implies that A contains a non-trivial 3-term arithmetic progression with a common gap in S?

For $G = \mathbb{Z}_n$ (n sufficiently large and odd) we show that $k = 3$, while for $G = \mathbb{F}_p^n$ (p an odd prime and n sufficiently large) we show that $k = 2$.

§1. Introduction

I. Given a set $A \subseteq [N]$ with positive density (dense, hereafter), an additional set $S \subseteq [N]$, and an integer $k \geq 2$ we may enquire whether A contains a k-term arithmetic progression (kAP hereafter) whose common difference lies in S (kSAP, hereafter). In the celebrated Szemerédi's theorem [18] we have $S = [N]$. In the far reaching generalisation of this theorem, namely the *polynomial Szemerédi's theorem* [1] we allow a set S as follows.

Theorem 1.1. (Polynomial Szemerédi's theorem [1])
For every $\alpha > 0$ there exists an N_0 such that for all $N \geq N_0$ the following holds.

Let $A \subseteq [N]$ have density α and let $P_1, \ldots, P_k$ be polynomials with integer coefficients all vanishing at zero. Then there exists a $d \neq 0$ such that A contains the (polynomial) configuration $x + P_1(d), \ldots, x + P_k(d)$.

II. Pursuing a quantitative version of Theorem 1.1, Green [12] established the following.

Theorem 1.2. (Green [12])
There exists a constant c such that any subset of $[N]$ of density at least $(\log \log N)^{-c}$ contains the configuration $\{x, x + d_1^2 + d_2^2, x + 2d_1^2 + 2d_2^2\}$ for some integers d_1 and d_2 not both zero.

Currently, the sole non-ergodic proof in the direction of the polynomial Szemerédi theorem is that of Green [12]. Notice that the set of allowed gaps S in Green's result is dense in $[N]$ (see, e.g., [19, Corollary 4.15] and comments thereafter).

III. In view of Theorem 1.2, we focus on dense gap sets S then; yet we decouple S from any number theoretical definition. We consider the emergence of 3SAPs in a dense set A based solely on the density and pseudorandomness level of S.

IV. Requiring only that S is dense is clearly insufficient[1]. Also insisting that S is a relatively dense subset of a random set is insufficient[2]. If to consider the infinite setting as a guide to which additional constraints should the set S satisfy, we see in this setting, already for the case $k = 2$, that $S \cap q\mathbb{Z} \neq \emptyset$, for every $q \in \mathbb{Z}$, is necessary.

V. Pseudorandom steps. Imposing a certain level of pseudorandomness on the set S is then natural. Conceptually, it is fairly obvious that if in addition to being dense S would also be "sufficiently" pseudorandom then we expect[3] an abundance of 3SAPs to emerge in A. The actual question here then is to quantify "sufficiently". We formulate this problem as follows.

Problem 1.3. *Let $\alpha, \sigma > 0$ and let A and S be subsets of an abelian group G of densities α and σ, respectively. What is the least integer $k \geq 2$ for which there exists an $\eta = \eta(\alpha, \sigma)$ such that $\|S\|_{U^k} \leq \eta$ implies that A contains a non-trivial 3SAP?*

Here, $\|\cdot\|_{U^k}$ denotes the kth Gowers norm.

VI. Random steps. The emergence of 3SAPs for a random set S was considered in [5, 3] and these results were improved recently in [6]. By these results we have the following. Let n be sufficiently large and choose a dense set $A \subseteq [1, n]$. Next, draw uniformly at random a set $S \subseteq [1, n]$ of density $\geq \omega(n)n^{-1/2}$, where $\omega(n) \to \infty$ with n. Then, with high probability, A contains 3SAP.

The results [5, 3, 6] are all proved using ergodic methods. It would be interesting to see infinitary proofs of these results.

VII. Throughout our notation is that of [19], with the exception that the characteristic function of a set X is denoted $X(\cdot)$.

§1.1 Our results.

I. We consider Problem 1.3 for the groups $\mathbb{Z}_n$ and $\mathbb{F}_p^n$ (p odd prime). For the former, we show that $k = 3$ (Theorem 1.6) and for the latter we show that $k = 2$ (Theorem 1.7).

II. Two point configurations (2SAPs). Prior to our main results, we address a simpler problem. That is, given two sets A and S in $\mathbb{Z}_n$ let us now consider the emergence of 2SAPs in A: two points $x, y \in A$ such that $x - y \in S$. The following asserts that a weak pseudorandomness assumption is sufficient to ensure the emergence of 2SAPs in $\mathbb{Z}_n$ (unlike the situation for 3SAPs in this group).

Proposition 1.4. *Let $S \subseteq \mathbb{Z}_n$ be symmetric[4]. Then, any set of density at least $\|S\|_u / \|S\|_{L^1}$, contains a 2SAP.*

Here, $\|S\|_u$ denotes the *linear bias* of S given by

$$\|S\|_u = \sup_{\xi \in \widehat{\mathbb{Z}_n} \setminus \widehat{0}} |\widehat{S(\xi)}| = \sup_{\xi \in \widehat{\mathbb{Z}_n} \setminus \widehat{0}} \left| \mathbb{E}_{x \in \mathbb{Z}_n} S(x) \overline{\xi(x)} \right|, \tag{1.5}$$

where $\widehat{S} : \widehat{\mathbb{Z}_n} \to \mathbb{C}$ is the Fourier transform of S.

[1] Consider: $A = (\frac{2}{3}N, N]$ and $S = (\frac{1}{3}N, \frac{2}{3}N]$.

[2] Take A to be the even numbers and S to be the intersection of the odd numbers with a dense random set.

[3] Not in the probabilistic sense.

[4] A set $X \subseteq \mathbb{Z}_n$ is called symmetric if $x \in X \iff x^{-1} \in X$.

Put another way, for a symmetric set S that is pseudorandom in the weak sense of $\|S\|_u = \varepsilon(n)\|S\|_{L^1}$, $\varepsilon(n) > 0$, we have, by Proposition 1.4, that every subsets of $\mathbb{Z}_n$ of density $\geq \varepsilon(n)$ contains a 2SAP.

III. 3SAPs in $\mathbb{Z}_n$. Let us now consider Problem 1.3 for sets A and S taken in $\mathbb{Z}_n$. Unlike the situation for 2SAPs, here an assumption on $\|S\|_{U_2}$ is insufficient; and $k \geq 3$ (k per Problem 1.3) is required.

To see this, fix $\varepsilon < 1/10$ and an irrational number ϑ. Consider the sets

$$A = \{x \mod n : \|x^2\vartheta\| < \varepsilon\}$$

and

$$S = \{d \mod n : \|2d^2\vartheta\| > 1/2 - \varepsilon\},$$

where $\|t\| = \min\{\{t\}, 1 - \{t\}\}$ is the distance from $t \in \mathbb{R}$ to the closest integer and where $\{t\}$ denotes the fractional part of t.

The sets A and S, just defined, are both dense and highly pseudorandom in the sense that $\|A\|_u = o(|A|)$ and $\|S\|_u = o(|S|)$ (so that there U^2 norms are arbitrarily small) [11, pp. $9 - 10$].

Nevertheless, A contains no 3SAPs. Assume towards contradiction that $(x, x+d, x+2d)$ is a 3SAP in A (i.e., $d \in S$). Then

$$\| - 2(x+d)^2\vartheta\| = \|2(x+d)^2\vartheta\| \leq 2\|(x+d)^2\vartheta\| < 2\varepsilon.$$

Observe that

$$\|2d^2\vartheta\| = \|(x^2 - 2(x+d)^2 + (x+2d)^2)\vartheta\| < 4\varepsilon$$

contradicting the assumption that $d \in S$ and satisfying $\|2d^2\vartheta\| > 1/2 - \varepsilon$.

On the other hand we show that $k \leq 3$; in particular we prove the following.

Theorem 1.6. *Let $\alpha > 0$ and $\sigma > 0$ there exists an $\eta > 0$ and n_0 such that for every $n \geq n_0$ the following holds.*

Let A and S be subsets of $\mathbb{Z}_n$ of densities α and σ, respectively, and such that $\|S\|_{U^3(\mathbb{Z}_n)} \leq \eta$. Then there exists a constant C such that A contains at least[5] $C|S|n$ 3SAPs.

In our proof of Theorem 1.6 we may have

$$n_0 = \exp\exp(\alpha^{-K})/\sigma D(\alpha).$$

where K and $D(\alpha)$ are defined in § 3.

IV. 3SAPs in $\mathbb{F}_p^n$. Throughout, p is an odd prime. Unlike $\mathbb{Z}_n$, we show that for $\mathbb{F}_p^n$ an assumption on $\|S\|_{U^2}$ is sufficient. In particular, the following is our main result.

Theorem 1.7. *For every $\alpha > 0$ and $\sigma > 0$ there exist an $\eta > 0$ and an n_0 such that for every integer $n \geq n_0$ the following holds.*

Let A and S be subsets of $\mathbb{F}_p^n$ of densities α and σ, respectively, such that $\|S\|_u \leq \eta\sigma$. Then, there exists a constant C such that A contains at least $C|S|p^n$ 3SAPs.

Here, though possible, we do not quantify n_0.

[5] As for n sufficiently large $|S|n \gg |A|$ this means that non-trivial 3SAPs are captured.

V. About the proofs. For the proof of our main result, Theorem 1.7, we employ the so called arithmetic regularity lemma established by Green and Tao [15]; we apply it for the U^3 norm in $\mathbb{F}_p^n$.

Theorem 1.6 is a consequence of the inverse U^3 theorem [14], and the fact that for a dense subset $A \subseteq \mathbb{Z}_n$, the set

$$S_3(A) = \{d \in \mathbb{Z}_n : \{x, x+d, x+2d\} \subseteq A,\ x \in A\}, \tag{1.8}$$

(consisting of elements $d \in \mathbb{Z}_n$ which form a common difference of some 3-term arithmetic progression in A) is essentially "almost periodic" (in a sense that will be made clear in § 3) as shown recently by Candela [2].

VI. Organisation. In § 2 we prove Proposition 1.4; then in § 3 we prove Theorem 1.6; finally, in § 4 we prove our main result, namely Theorem 1.7.

§2. 2SAPs in subsets of $\mathbb{Z}_n$

In this section we prove Proposition 1.4; this is a special case of the following. In this section only it is more convenient to use the counting measure on $\mathbb{Z}_n$ and the uniform measure of $\widehat{\mathbb{Z}_n}$ when defining Fourier coefficients. That is, in this section only we write

$$\|S\|_u = \sup_{\xi \in \widehat{\mathbb{Z}_n} \setminus \widehat{0}} |\widehat{S(\xi)}| = \sup_{\xi \in \widehat{\mathbb{Z}_n}} \left| \sum_{x \in \mathbb{Z}_n} S(x)\overline{\xi(x)} \right|, \tag{2.1}$$

instead of (1.5).

Proposition 2.2. *Let A and S be subsets of $\mathbb{Z}_n$. If for every $s \in S$ the set A contains no two points x and y such that $x - y = s$ and contains no two points x' and y' such that $x' - y' = s^{-1}$, then $|A|/n < \|S\|_u/|S|$.*

Proof. Given $S \subseteq \mathbb{Z}_n$, let G_S denote the undirected Cayley graph generated by S; that is $G_S = (V, E)$ where $V = \mathbb{Z}_n$ and $E = \{xy : x, y \in \mathbb{Z}_n,\ x - y \in S\}$. A stable[6] set $A \subseteq V$ corresponds to a subset of $\mathbb{Z}_n$ satisfying the property that for every $s \in S$ there exist no two points $x, y \in A$ with $x - y = s$ and there exist no two points $x', y' \in A$ with $x - y = s^{-1}$.

Suffices now to provide an upper bound on $\alpha(G_S)$, the size of the largest stable set in G_S. As G_S is $|S|$-regular, i.e., all vertices have degree $|S|$, we have that

$$\alpha(G_S)/n \le \frac{-\lambda_{\min}}{\lambda_1 - \lambda_{\min}} \tag{2.3}$$

by a result of Hoffman [16]; where here

$$|S| = \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n = \lambda_{\min}$$

are the eigenvalues of the adjacency matrix of G_S.

The adjacency matrix of G_S is circulant and so its eigenvalues are the Fourier coefficients of S; see, e.g., [4] or [17, Chapter 11, exercise 8]. In particular, $-\lambda_{\min} = |\lambda_{\min}| \le \|S\|_u$; as a result (2.3) becomes

$$\alpha(G_S)/n \le \frac{-\lambda_{\min}}{\lambda_1 - \lambda_{\min}} < \frac{\|S\|_u}{|S|}$$

and the claim follows. ■

[6] A set of vertices in a graph spanning no edges between its members is called *stable*.

§3. 3SAPs in subsets of $\mathbb{Z}_n$

In this section, we prove Theorem 1.6. Throughout this section A and S are subsets of $\mathbb{Z}_n$ (n sufficiently large as we shall see) of densities α and σ respectively, where the latter are both independent of n. Also, throughout this section, $S_3 = S_3(A)$ is the set defined in (1.8).

§3.1 Preliminaries.

We require the inverse U^3 theorem in the form [10, Theorem 3.2] and also a result of Candela [2] detailing the structure of the set $S_3(A)$ (see (1.8) for a definition).

I. Quadratic averages. Let G and H be finite abelian group, let $X \subseteq G$, and let $\gamma : X \to H$. The function γ is called *linear on* X if it vanishes along any parallelogram contained in X, that is to say, for any $(x, x+h_1, x+h_2, x+h_1+h_2) \in X^4$ we have that

$$\gamma(x) - \gamma(x+h_1) - \gamma(x+h_2) + \gamma(x+h_1+h_2) = 0.$$

The function γ is called *quadratic on* X if it vanishes along all parallelepipeds contained in X; that is for any

$$(x, x+h_1, x+h_2, x+h_3, x+h_1+h_2, x+h_1+h_3, x+h_2+h_3, x+h_1+h_2+h_3) \in X^8$$

we have that

$$\gamma(x) - \gamma(x+h_1) - \gamma(x+h_2) - \gamma(x+h_3) + \gamma(x+h_1+h_2) + \gamma(x+h_1+h_3) + \gamma(x+h_2+h_3) - \gamma(x+h_1+h_2+h_3) = 0.$$

Let B be a regular Bohr set[7] in $\mathbb{Z}_n$ and let $q : B \to \mathbb{T}$ be quadratic on B. A *quadratic average of base* (B, q) *in* $\mathbb{Z}_n$ is a function $Q : \mathbb{Z}_n \to \mathbb{C}$ given by

$$Q(x) = \mathbb{E}_{y \in x-B}\, e(q_y(x)), \tag{3.1}$$

where for each y, the mapping q_y is a quadratic map on B given by $q_y(x) = q(x-y) + \varphi_y(x-y)$ such that for each y, the mapping $\varphi_y : B \to \mathbb{T}$ is linear on B.

The *complexity* of a Bohr set $B = B(K, \varrho)$ with *frequency set* K and *radius* ϱ is the pair $(|K|, \varrho)$. A Bohr set B is said to have *complexity at most* (d, ϱ) if there exists a set K of cardinality at most d and a $\varrho' \geq \varrho$ such that $B = B(K, \varrho')$. The *complexity of a quadratic average* is that of the Bohr set appearing in its base.

II. Throughout this section, $C_0 = 2^{24}$. The following is the so called *inverse* U^3 *theorem* as this appears in [10, Theorem 3.2].

Theorem 3.2. (The inverse U^3 theorem) *Let* $f : \mathbb{Z}_n \to \mathbb{C}$ *with* $\|f\|_\infty \leq 1$ *and* $\|f\|_{U^3(\mathbb{Z}_n)} \geq \delta$. *Then there exists a regular Bohr set* B *of complexity at most* $((2/\delta)^{C_0}, (\delta/2)^{C_0})$ *and a quadratic map* $q : B \to \mathbb{Z}_N$ *such that* $|\langle f, Q\rangle| \geq (\delta/2)^{C_0}/2$ *where* Q *is a quadratic average with base* (B, q).

[7] We refer the reader to [19, Chapter 4] for the terminology and basic properties involving Bohr sets.

III. The structure of $S_3(A)$. By [2, Lemma 3]

$$\|S_3\|^*_{U^3(G)} \leq \alpha^{3/2}, \tag{3.3}$$

where $\|\cdot\|^*_{U^3(G)}$ denotes the *dual* U^3 *norm* given by

$$\|f\|^*_{U^3(\mathbb{Z}_n)} = \sup_{g:\|g\|_{U^3(\mathbb{Z}_n)}\leq 1} |\langle f, g\rangle| \tag{3.4}$$

with $f, g : \mathbb{Z}_n \to \mathbb{C}$ and $\langle f, g\rangle = \mathbb{E}_{x\in\mathbb{Z}_n} f(x)\overline{g(x)}$.

Functions with bounded dual U^3 norm admit a "simple" structure in the sense that these can be expressed as a small sum of quadratic averages (see (3.6) for a rigours explanation). This qualitative statement was made quantitative by Candela [2, Theorem 4] who specialised the Gowers-Wolf decomposition theorem [10, Theorem 7.5] for functions with bounded dual U^3 norm.

By (3.3) and [2, Theorem 4] we have the following.

Theorem 3.5. (The structure of S_3)
Let $\alpha > 0$ and $\delta > 0$; set

$$d = (2\alpha^{3/2}/\delta)^{C_0},\ \varrho = (\delta/2\alpha^{3/2})^{C_0},$$

and let $\xi \in (0, 2^{-17}d^{-4}\delta^8)$.

If $A \subseteq \mathbb{Z}_n$, n odd, is a subset of density α then:

$$S_3(x) = \sum_{i=1}^{k} U_i(x)Q_i(x) + g(x) \tag{3.6}$$

such that the following holds.

1. *For each $i \in [k]$, Q_i is a quadratic average on G of complexity at most $(d, \xi\varrho/400d4^d)$.*
2. *The functions $U_i : G \to \mathbb{C}$ satisfy the the following.*
   (a) *$\sum_{i=1}^{k} \|U_i\|_\infty \leq 4d$, and*
   (b) *for each $i \in [k]$, $U_i(x) = W_i(x)+V_i(x)$ such that $\|V_i\|_{L^1(G)} \leq 296\xi d$ and $\|W_i\|^*_{U^2(G)} \leq 4\gamma^{-3/4}d$, where γ is the density of a Bohr set of complexity at most*

   $$(K, \varphi) = (2d + 2^{49}\delta^{-24}(4d)^{12}, 2^{-168}\delta^{48}(4d)^{-24}\xi^6 d^{-6}4^{-d}).$$
3. *$k \leq 32(d/\delta)^2$.*
4. *$\|g\|_{L^1(G)} \leq 3\delta$.*

IV. For future reference, let us note here that by a result of Gowers [7, 8], if $A \subseteq \mathbb{Z}_n$ has density α (independent of n), then there exists a constant L such that whenever $n \geq n_\alpha$, then

$$|S_3|/n \geq C(\alpha), \tag{3.7}$$

where

$$n_\alpha = \exp\exp(\alpha^{-L}). \tag{3.8}$$

§3.2 Proof of Theorem 1.6.

We now prove Theorem 1.6. Throughout the proof we shall only use the fact that γ (per Theorem 3.5) satisfies $\gamma \geq \varphi^{|K|}$ (K per Theorem 3.5), by [19, Lemma 4.20]. For future reference, we note here that this lower bound is a function of δ and ξ (per Theorem 3.5) and we denote it by

$$\beta = \beta(\delta, \xi) = \varphi^{|K|}. \tag{3.9}$$

Proof of Theorem 1.6. Given A, S, α, and σ as in the premise of Theorem 1.6, we put $S_3 = S_3(A)$. For every $n \geq n_\alpha$ (per (3.8)), we show that for every $\nu > 0$ there exists a choice for η (per Theorem 1.6) such that

$$\left| \frac{|S \cap S_3|}{n} - \frac{|S|}{n}\frac{|S_3|}{n} \right| \leq \nu, \tag{3.10}$$

and thus bound the deviation of $|S \cap S_3|$ from what one would expect it to be if S was truly random. Theorem 1.6 is then clearly implied by (3.10).

In what follows then we prove (3.10). Given ν, set:

$$\delta = \frac{\nu}{10^6}, \quad \xi = \min\{\frac{\nu\delta^2\varrho^3}{10^6}, \frac{\delta^8}{2^{17}d^4}\} \quad \varepsilon = \frac{\nu\delta^2\varrho^2}{10^6}, \tag{3.11}$$

where d and ϱ are as in Theorem 3.5. In addition, set η so that

$$\eta^{C_0} \leq \min\{\frac{\nu\varepsilon^4\beta^3\delta^2\varrho^6}{10^6}, \frac{\xi\varrho^2}{4004^d}\}, \tag{3.12}$$

where β is as in (3.9). In particular we have: $0 < \eta < \varepsilon < \delta < 1$ and $0 < \eta < \xi < \delta < 1$.

Set $f_S(x) = S(x) - |S|/n$ and note that

$$\left| \frac{|S \cap S_3|}{n} - \frac{|S|}{n}\frac{|S_3|}{n} \right| = |(f_S * S_3)(0)|.$$

Then, Theorem 3.5, applied with A, α, δ, and ξ, yields

$$\begin{aligned} |(f_S * S_3)(0)| &\leq |\sum_{i=1}^{k} \mathbb{E}_{y\in\mathbb{Z}_n} U_i(y)Q_i(y)f_S(-y)| + |\mathbb{E}_{y\in\mathbb{Z}_n} g(y)f_S(-y)| \\ &\leq |\sum_{i=1}^{k} \mathbb{E}_{y\in\mathbb{Z}_n} W_i(y)Q_i(y)f_S(-y)| + |\sum_{i=1}^{k} \mathbb{E}_{y\in\mathbb{Z}_n} V_i(y)Q_i(y)f_S(-y)| + |\mathbb{E}_{y\in\mathbb{Z}_n} g(y)f_S(-y)|, \end{aligned} \tag{3.13}$$

where Q_i, U_i, W_i, V_i, g, and k are as in Theorem 3.5.

In what follows, we provide upper bound estimations for each of the terms appearing on the right hand side of (3.13). The term involving g can be upper bounded by $\|g\|_{L^1(\mathbb{Z}_n)}$; indeed, it is not hard to verify that since $\|S\|_\infty \leq 1$ we have that

$$|\mathbb{E}_{y\in\mathbb{Z}_n} g(y)f_S(-y)| \leq 2\|g\|_{L^1(\mathbb{Z}_n)}. \tag{3.14}$$

Consider the term on the right hand side of (3.13) involving V_i. As $\|S\|_\infty \leq 1$ and since $\|Q_i\|_\infty \leq 1$ for every $i \in [k]$, we have that for every $i \in [k]$

$$|\mathbb{E}_{y\in\mathbb{Z}_n} V_i(y)Q_i(y)f_S(-y)| \leq 2\|V_i\|_{L^1(\mathbb{Z}_n)}, \tag{3.15}$$

so that

$$|\sum_{i=1}^{k} \mathbb{E}_{y\in\mathbb{Z}_n} V_i(y)Q_i(y)f_S(-y)| \leq 2k \max_{i\in[k]} \|V_i\|_{L^1(\mathbb{Z}_n)}. \tag{3.16}$$

Consider the term on the right hand side of (3.13) involving W_i. For each i, we have that $\|W_i\|^*_{U^2(\mathbb{Z}_n)} \leq 4\beta^{-3/4}d$ (by Theorem 3.5), where $\|\cdot\|^*_{U^2(\mathbb{Z}_n)}$ denotes the dual U^2 norm given by

$$\|f\|^*_{U^2(\mathbb{Z}_n)} = \sup_{g:\|g\|_{U^2(\mathbb{Z}_n)}\leq 1} |\langle f, g\rangle|.$$

The following version of [2, Lemma 2] then provides a useful decomposition of W_i.

Lemma 3.17. *Let $f : \mathbb{Z}_n \to \mathbb{C}$ and let $\zeta > 0$. Then there exists a $g : \mathbb{Z}_n \to \mathbb{C}$ satisfying*

$$\|f - g\|_{L^2(\mathbb{Z}_n)} \leq \zeta,\ |\operatorname{supp}\widehat{g}| \leq (\|f\|^*_{U^2(\mathbb{Z}_n)}/\zeta)^4.$$

By Lemma 3.17, applied to W_i with $\zeta = \varepsilon$ (and ε as in (3.11)), we have that $W_i = g_i + h_i$ where $\|h_i\|_{L^2(\mathbb{Z}_n)} \leq \varepsilon$ and $|\operatorname{supp}\widehat{g_i}| \leq L_i$ where $L_i \leq (\|W_i\|^*_{U^2(\mathbb{Z}_n)}/\varepsilon)^4$. Now, the term on the right hand side of (3.13) involving W_i satisfies

$$|\sum_{i=1}^{k} \mathbb{E}_{y\in\mathbb{Z}_n} W_i(y)Q_i(y)f_S(-y)| \leq \sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} g_i(y)Q_i(y)f_S(-y)| + \sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} h_i(y)Q_i(y)f_S(-y)|. \tag{3.18}$$

By Cauchy-Schwarz and the fact that $\|S\|_\infty \leq 1$ and $\|Q_i\|_\infty \leq 1$ for each $i \in [k]$, we have that

$$\sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} h_i(y)Q_i(y)f_S(-y)| \leq 2k \max_{i\in[k]} \|h_i\|_{L^2(\mathbb{Z}_n)}. \tag{3.19}$$

In addition, we see that

$$\sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} g_i(y)Q_i(y)f_S(-y)| \leq \sum_{i=1}^{k} L_i |\mathbb{E}_{y\in\mathbb{Z}_n} Q_i^*(y)f_S(-y)|, \tag{3.20}$$

where Q_i^* is as follows. For a given $y \in \mathbb{Z}_n$ we have that $g_i(y) = r_y e^{i\vartheta_y}$ as this is a complex number. The assumption $|\operatorname{supp}\widehat{g_i}| \leq L_i$ implies that $|r_y| \leq L_i$ for every $y \in \mathbb{Z}_n$. Consequently, we may write

$$\begin{aligned}\sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} g_i(y)Q_i(y)f_S(-y)| &\leq \sum_{i=1}^{k} |\mathbb{E}_{y\in\mathbb{Z}_n} L_i e^{i\vartheta_y} Q_i(y)f_S(-y)| \\ &\leq \sum_{i=1}^{k} L_i |\mathbb{E}_{y\in\mathbb{Z}_n} e^{i\vartheta_y} Q_i(y)f_S(-y)|.\end{aligned}$$

For a quadratic average Q_i, let $q_{y,i}(x) = q_i(x-y) + \varphi_{y,i}(x-y)$ denote the quadratic form defining Q_i (see (3.1)). Put $\varphi^*_{y,i}(z) = \varphi^*_{y,i}(z) + \vartheta_y$; which remains linear. Let, now $q^*_{y,i}(x) = q_i(x-y) + \varphi^*_{y,i}(x-y)$, and define Q_i^* to be the quadratic average defined over $q^*_{y,i}$ (per (3.1)).

Coming back to (3.20), we seek to estimate $|\mathbb{E}_{y\in\mathbb{Z}_n} Q_i^*(y)f_S(-y)|$. We invoke Theorem 3.2 as follows. Put $m(x) = f_{-S}(x)/2$; so that $\|m\|_\infty \leq 1$ (since $\|f_{-S}\|_\infty \leq 2$), and observe that $\|m\|_{U^3(\mathbb{Z}_n)} \leq \|S\|_{U^3(\mathbb{Z}_n)}/2 \leq \eta/2$. Then, by Theorem 3.2, we have that

$$|\langle m_S, Q_i^*\rangle| = |\mathbb{E}_{y\in\mathbb{Z}_n} Q_i^*(y)m_S(y)| \leq (\eta/4)^{C_0}/2, \tag{3.21}$$

provided that

$$d \leq \left(\frac{4}{\eta}\right)^{C_0}, \tag{3.22}$$

and that there exists a $\xi \in (0, 2^{-17}d^{-4}\delta^8)$ for which

$$\left(\frac{\eta}{4}\right)^{C_0} \leq \frac{\xi\varrho}{400d4^d}, \tag{3.23}$$

where d, ξ and ϱ are as in Theorem 3.5, and we recall that each Q_i is defined on a Bohr set of complexity at most $(d, \xi\varrho/400d4^d)$.

By our choice of η (see (3.12)) both of these conditions ((3.22) and (3.23)) are satisfied. Then, (3.21) and the observation that $\|f_S\|_{U^3(\mathbb{Z}_n)} = \|f_{-S}\|_{U^3(\mathbb{Z}_n)}$ yield that

$$|\,\mathbb{E}_{y\in\mathbb{Z}_n}\, Q_i^*(y) f_S(-y)| \leq \left(\frac{\eta}{4}\right)^{C_0}. \tag{3.24}$$

As a result, we now have that if (3.22) and (3.23) both hold then

$$\sum_{i=1}^{k} |\,\mathbb{E}_{y\in\mathbb{Z}_n}\, g_i(y)Q_i(y)f_S(-y)| \leq k\left(\frac{\eta}{4}\right)^{C_0} \max_{i\in[k]} L_i. \tag{3.25}$$

Collecting all of the above estimations (i.e., (3.14), (3.16), (3.19), and (3.25)) we arrive at

$$|(f_S * S_3)(0)| \leq k\,(\eta/4)^{C_0} \max_{i\in[k]} L_i + 2k \max_{i\in[k]} \|h_i\|_{L^2(\mathbb{Z}_n)} + 2k \max_{i\in[k]} \|V_i\|_{L^1(\mathbb{Z}_n)} + 2\|g\|_{L^1(\mathbb{Z}_n)} \tag{3.26}$$

It is not hard not to verify that due to our choice of constants (see (3.11) and (3.12)) each of the terms appearing on the right hand side of (3.26) is at most $\nu/4$.

This concludes our proof of Theorem 1.6. ■

§4. 3SAPs in subsets of $\mathbb{F}_p^n$

In this section, we prove our main result, that is Theorem 1.7. In § 4.1 and § 4.2 we prepare for the proof of Theorem 1.7 presented in § 4.4

§4.1 A generalised von Neumann type lemma.

The aim of this section is to prove Corollary 4.19 stated below.

I. Let G be a finite abelian group. For $h \in G$ and a function $f : G \to \mathbb{C}$, we write

$$\Delta(f,h)(x) = f(x)\overline{f(x+h)}. \tag{4.1}$$

With this notation we have that

$$\|f\|_{U^d(G)}^{2^d} = \mathbb{E}_{h\in G}\, \|\Delta(f,h)\|_{U^{d-1}(G)}^{2^{d-1}}, \tag{4.2}$$

whenever $d \geq 2$ [19, Chapter 11]. Also, for $S \subseteq G$ and any $h \in G$, it is not hard to see that

$$\|\Delta(S,h)\|_{U^2(G)} \leq \|S\|_{U^2(G)} \tag{4.3}$$

In addition, let us record here for future reference the elegant quality

$$\|f\|_{U^2(G)} = \|\widehat{f}\|_{\ell^4(\widehat{G})} \tag{4.4}$$

see e.g., [19, pg. 419].

II. Given $S \subseteq G$, we consider

$$|\mathbb{E}_{x\in G,d\in S}\, f_1(x)f_2(x+d)f_3(x+2d)| = |\mathbb{E}_{x\in G,d\in G}\, f_1(x)f_2(x+d)f_3(x+2d)\mu_S(d)|\,, \tag{4.5}$$

where $\mu_S(x) = S(x)\left(\mathbb{E}_{x\in G}\, S(x)\right)^{-1} = S(x)/\|S\|_{L^1(G)}$, and seek to provide an upper bound estimation in the case that S is dense and pseudorandom in the sense that $\|S\|_{U^2(G)}$ is small. This estimation is presented in Corollary 4.19. In the next lemma we consider (4.5) without the normalisation by $\|S\|_{L^1(G)}$.

Lemma 4.6. *Let $k \geq 3$ be an integer, let $\mathcal{F} = \{f_1, \dots, f_k\} \subset \mathbb{C}^G$ be a collection of complex valued functions over G, and let $S \subseteq G$. If $\|f\|_\infty \leq 1$ for each $f \in \mathcal{F} \setminus \{g\}$ for some $g \in \mathcal{F}$, then*

$$|\mathbb{E}_{x,d\in G}\, f_1(x)\cdots f_k(x+(k-1)d)S(d)| \leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/2^{k-1}} \|g\|_{U^k(G)}. \tag{4.7}$$

Remarks.

1. For our needs, the case $k=3$ is sufficient; nevertheless, treating an arbitrary k is not much of a burden.
2. As in Theorem 1.7 the set S is dense, we can make do with the weaker bound
$$|\mathbb{E}_{x,d\in G}\, f_1(x)\cdots f_k(x+(k-1)d)S(d)| \leq \|g\|_{U^k(G)} \tag{4.8}$$
instead of (4.7). Nevertheless, we find (4.7) more insightful and since its proof does not significantly prolong the argument we provide its proof. In Paragraph § 4.1.III below we indicate what is to be altered in order to attain (4.8) with a slightly shorter argument.
3. Comparing (4.7) or (4.8) with the traditional generalised von Neumann theorem [19, Lemma 11.4] we see that when $S = G$ then one can replace U^k in (4.8) with U^{k-1}. In the sequel this is precisely the reason why our approach of using the arithmetic regularity lemma to prove Theorem 1.7 is applied for the U^3 norm and not with the U^2 norm.

Proof of Lemma 4.6. The proof is by induction on k. We assume, without loss of generality, that $f_2 = g$ and show that

$$|\mathbb{E}_{x,d\in G}\, f_1(x)\cdots f_k(x+(k-1)d)S(d)| \leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/2^{k-1}} \|f_2\|_{U^k(G)}. \tag{4.9}$$

We begin by following the traditional argument for the generalised von Neumann theorem for arithmetic progressions [19, Lemma 11.4]. By Cauchy-Schwarz, we may write

$$\begin{aligned} |\ \mathbb{E}_{x,d\in G} & f_1(x)\cdots f_k(x+(k-1)d)S(d)\ |^2 \\ &\leq \|f_1\|^2_{L^2(G)}\, \mathbb{E}_{x\in G}\, |\mathbb{E}_{d\in G}\, f_2(x+d)\cdots f_k(x+(k-1)d)S(d)|^2 \\ &\leq \mathbb{E}_{x\in G}\, |\mathbb{E}_{d\in G}\, f_2(x+d)\cdots f_k(x+(k-1)d)S(d)|^2\,; \end{aligned} \tag{4.10}$$

For the right hand side of (4.10) we have

$$\begin{aligned} \mathbb{E}_{x\in G}\, |\ & \mathbb{E}_{d\in G}\, f_2(x+d)\cdots f_k(x+(k-1)d)S(d)\ |^2 \\ &= \mathbb{E}_{x\in G}\left(\mathbb{E}_{d\in G}\, f_2(x+d)\cdots f_k(x+(k-1)d)S(d)\right)\overline{\left(\mathbb{E}_{d'\in G}\, f_2(x+d')\cdots f_k(x+(k-1)d)S(d')\right)} \\ &= \mathbb{E}_{x,d,d'\in G}\, f_2(x+d)\overline{f_2(x+d')}\cdots f_k(x+(k-1)d)\overline{f_k(x+(k-1)d')}S(d)\overline{S(d')}. \end{aligned}$$

Setting $y = x + d$ and $h = d' - d$ in the last expression yields

$$\begin{aligned}
&= \mathbb{E}_{y,h,d\in G}\, f_2(y)\overline{f_2(y+h)}\cdots f_k(y+(k-2)d)\overline{f_k(y+(k-1)h+(k-2)d)}S(d)\overline{S(d+h)}\\
&= \mathbb{E}_{y,h,d\in G}\,\Delta(f_2,h)(y)\cdots\Delta(f_k,(k-1)h)(y+(k-2)d)\Delta(S,h)(d)\\
&\le \mathbb{E}_{h\in G}\left|\mathbb{E}_{y,d\in G}\,\Delta(f_2,h)(y)\cdots\Delta(f_k,(k-1)h)(y+(k-2)d)\Delta(S,h)(d)\right|. \qquad (4.11)
\end{aligned}$$

Combining (4.10) with (4.11) yields

$$\begin{aligned}
&|\ \mathbb{E}_{x,d\in G} f_1(x)\cdots f_k(x+(k-1)d)S(d)\ |^2\\
&\qquad \le \mathbb{E}_{y,h,d\in G}\,\Delta(f_2,h)(y)\cdots\Delta(f_k,(k-1)h)(y+(k-2)d)\Delta(S,h)(d). \qquad (4.12)
\end{aligned}$$

The induction base. At this point we depart from the traditional argument for the generalised von Neumann theorem for arithmetic progressions [19, Lemma 11.4]. To establish the induction base, we now consider (4.12) for $k = 3$. By Fourier inversion, the right hand side of (4.12) can be written as follows

$$\begin{aligned}
&\mathbb{E}_{y,d,h}\left(\sum_{\xi_1\in\widehat{G}}\widehat{\Delta(f_2,h)}(\xi_1)\xi_1(y)\right)\left(\sum_{\xi_2\in\widehat{G}}\widehat{\Delta(f_3,2h)}(\xi_2)\xi_2(y+d)\right)\left(\sum_{\xi_3\in\widehat{G}}\widehat{\Delta(S,h)}(\xi_3)\xi_3(d)\right)\\
&\quad = \mathbb{E}_{h\in G}\sum_{\xi_1,\xi_2,\xi_3\in\widehat{G}}\widehat{\Delta(f_2,h)}(\xi_1)\widehat{\Delta(f_3,2h)}(\xi_2)\widehat{\Delta(S,h)}(\xi_3)\,\mathbb{E}_{y\in G}\,\xi_1(y)\xi_2(y)\,\mathbb{E}_{d\in G}\,\xi_2(d)\xi_3(d)\\
&\quad = \mathbb{E}_{h\in G}\sum_{\xi\in\widehat{G}}\widehat{\Delta(f_2,h)}(\xi)\widehat{\Delta(f_3,2h)}(-\xi)\widehat{\Delta(S,h)}(\xi), \qquad (4.13)
\end{aligned}$$

where the last equality is due to orthogonality relations.

The absolute value of the summand in the sum appearing on the right hand side of (4.13) is at most $\|\widehat{\Delta(f_2,h)}\widehat{\Delta(f_3,2h)}\widehat{\Delta(S,h)}\|_{\ell^1(\widehat{G})}$; so that

$$|\ \mathbb{E}_{x,d\in G}\, f_2(x+d)f_3(x+2d)S(d)\ |^2 \le \mathbb{E}_{h\in G}\,\|\widehat{\Delta(f_2,h)}\widehat{\Delta(f_3,2h)}\widehat{\Delta(S,h)}\|_{\ell^1(\widehat{G})} \qquad (4.14)$$

so that by Hölder's inequality we arrive at

$$\le \mathbb{E}_{h\in G}\,\|\widehat{\Delta(f_2,h)}\|_{\ell^4(\widehat{G})}\|\widehat{\Delta(f_3,2h)}\|_{\ell^4(\widehat{G})}\|\widehat{\Delta(S,h)}\|_{\ell^2(\widehat{G})}.$$

Then, by (4.4) and the Parseval equality:

$$= \mathbb{E}_{h\in G}\,\|\Delta(f_2,h)\|_{U^2(G)}\|\Delta(f_3,2h)\|_{U^2(G)}\|\Delta(S,h)\|_{L^2(G)}.$$

By (4.3), $\|\Delta(f_3,2h)\|_{U^2(G)} \le \|f_3\|_{U^2(G)}$; and as $\|f_3\|_\infty \le 1$ then $\|f_3\|_{U^2(G)} \le 1$ so that:

$$\le \mathbb{E}_{h\in G}\,\|\Delta(f_2,h)\|_{U^2(G)}\|\Delta(S,h)\|_{L^2(G)}. \qquad (4.15)$$

At this point we note that as S is a boolean function so is $\Delta(S,h)$; then for any fixed $h \in G$

$$\|\Delta(S,h)\|^2_{L^2(G)} = \mathbb{E}_{x\in G}\,S(x)S(x+h) = \frac{\#\{(x,x+h)\in S\}}{N},$$

i.e., $\|\Delta(S,h)\|^2_{L^2(G)}$ is simply the density of configurations of the form $(x, x+h)$ in the set[8] S. Now, as

$$\mathbb{E}_{x\in G}\,S(x)S(x+h) = |\widehat{S}(0)|^2 \pm \|S\|_u\|S\|^2_{L^2(G)} = \|S\|^2_{L^1(G)} \pm \|S\|_u\|S\|_{L^1(G)},$$

[8] At this point it is insightful to think of S as a truly random set in which case we expect $\|\Delta(S,h)\|^2_{L^2(G)} \approx \|S\|^2_{L^1}$.

(where we again use the fact that S is boolean), we arrive at

$$\|\Delta(S,h)\|_{L^2(G)} = \left(\|S\|^2_{L^1(G)} \pm \|S\|_u \|S\|_{L^1(G)}\right)^{1/2};$$

in particular, (4.15) now reads

$$|\,\mathbb{E}_{x,d\in G}\, f_2(x+d)f_3(x+2d)S(d)\,|^2 \leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/2} \mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^2(G)}.$$

Observe now that

$$\begin{aligned}(\,\mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^2(G)}\,)^4 &= (\,(\,\mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^2(G)}\,)^2\,)^2 \\ &\leq (\,\mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|^2_{U^2(G)}\,)^2 \\ &\leq \mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|^4_{U^2(G)} \overset{(4.2)}{=} \|f_2\|^8_{U^3(G)}, \qquad (4.16)\end{aligned}$$

where both inequalities are due to Cauchy-Schwarz. Consequently,

$$|\,\mathbb{E}_{x,d\in G}\, f_2(x+d)f_3(x+2d)S(d)\,| \leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/4} \|f_2\|_{U^3(G)}$$

yielding (4.9) for $k=3$ as required.

The induction step. Suppose then that $k>3$ and that (4.7) holds for $k-1$. From (4.12) we have that

$$\begin{aligned}&|\,\mathbb{E}_{x,d\in G} f_1(x)\cdots f_k(x+(k-1)d)S(d)\,|^2 \\ &\qquad \leq \mathbb{E}_{h\in G}\,|\,\mathbb{E}_{y,d\in G}\,\Delta(f_2,h)(y)\cdots\Delta(f_k,(k-1)h)(y+(k-2)d)\Delta(S,h)(d)\,|\,. \qquad (4.17)\end{aligned}$$

The function $\Delta(S,h)$ is boolean. Moreover, $\|\Delta(f_i,(i-1)h)\|_\infty \leq 1$ since $\|f_i\|_\infty \leq 1$ for each $i\in[3,k]$. Consequently, the inner expectation in (4.17) is

$$\leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/2^{k-2}} \|\Delta(f_2,h)\|_{U^{k-1}(G)},$$

by the induction hypothesis. Observe now that

$$\left(\mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^{k-1}(G)}\right)^{2^{k-1}} \leq \mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|^{2^{k-1}}_{U^{k-1}(G)} \overset{(4.2)}{=} \|f_2\|^{2^k}_{U^k(G)}, \qquad (4.18)$$

where the first inequality follows by $k-1$ applications of Cauchy-Schwarz. Assertion (4.9) now follows. ■

III. Regarding (4.8). Now that we have seen a proof for (4.7), let us revisit the weaker (4.8). To prove the latter with a slightly shorter argument, consider the induction basis of the argument above; in particular, consider (4.14) which we recall here:

$$|\,\mathbb{E}_{x,d\in G}\, f_2(x+d)f_3(x+2d)S(d)\,|^2 \leq \mathbb{E}_{h\in G}\, \|\widehat{\Delta(f_2,h)}\widehat{\Delta(f_3,2h)}\widehat{\Delta(S,h)}\|_{\ell^1(\widehat{G})}.$$

Now, apply Hölder's inequality as to attain

$$\leq \mathbb{E}_{h\in G}\, \|\widehat{\Delta(f_2,h)}\|_{\ell^4(\widehat{G})} \|\widehat{\Delta(f_3,2h)}\|_{\ell^2(\widehat{G})} \|\widehat{\Delta(S,h)}\|_{\ell^4(\widehat{G})}$$

(compare the application of Hölder's inequality here and that used in the proof above). Then by (4.4) and Parseval's equality

$$= \mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^2(G)} \|\Delta(f_3,2h)\|_{L^2(G)} \|\Delta(S,h)\|_{U^2(G)}.$$

As S is boolean, then $\|\Delta(S,h)\|_{U^2(G)} \leq \|S\|_{U^2(G)} \leq 1$, and as $\|f_3\|_\infty \leq 1$, by assumption, then $\|\Delta(f_3,2h)\|_{L^2(G)} \leq 1$; consequently we get

$$\leq \mathbb{E}_{h\in G}\, \|\Delta(f_2,h)\|_{U^2(G)}.$$

From here continue as in the argument above (with the necessary modifications to the induction step).

To summarise the difference between the proofs for (4.7) and (4.8), we see that for the former we allow the nature of S to take part in the estimate by considering the count of some two point configurations in it. For the latter, we ignore S by taking the trivial bound $\|\Delta(S,h)\|_{U^2(G)} \leq 1$.

IV. As mentioned above, we are in fact interested in an upper bound for (4.5). The following corollary of Lemma 4.6 provides such an upper bound.

Corollary 4.19. *Let $k \geq 3$ be an integer, $S \subseteq G$ a subset, and let $\mathcal{F} = \{f_1,\dots,f_k\} \subset \mathbb{C}^G$ be a collection of complex valued functions over G. Suppose that $\|f\|_\infty \leq 1$ for each $f \in \mathcal{F}\backslash\{g\}$ for some $g \in \mathcal{F}$. Then,*

$$|\mathbb{E}_{x\in G,d\in G}\, f_1(x)\cdots f_k(x+(k-1)d)\mu_S(d)| \leq \left(\|S\|^2_{L^1(G)} + \|S\|_u\|S\|_{L^1(G)}\right)^{1/2^{k-1}} \|g\|_{U^k(G)}\|S\|^{-1}_{L^1(G)} \tag{4.20}$$

V. As mentioned in the remarks made after the statement of Lemma 4.6, for our needs the weaker bound

$$|\mathbb{E}_{x\in G,d\in G}\, f_1(x)\cdots f_k(x+(k-1)d)\mu_S(d)| \leq \|g\|_{U^k(G)}\|S\|^{-1}_{L^1(G)} \tag{4.21}$$

shall suffice.

VI. For pseudorandom sets S in the sense that $\|S\|_u$ is dominated by $\|S\|_{L^1(G)}$ we have that $\|S\|^2_{L^1(G)}$ dominates $\|S\|_u\|S\|_{L^1(G)}$ so that (4.20) for $k=3$ becomes

$$|\mathbb{E}_{x\in G,d\in G}\, f_1(x)f_2(x+d)f_3(x+2d)\mu_S(d)| \ll \|g\|_{U^k(G)}\|S\|^{-1/2}_{L^1(G)}.$$

This type of bound is too weak to handle sets S of density $o(1)$.

§4.2 The arithmetic regularity lemma for U^3 in $\mathbb{F}_p^n$.

The aim of this section is to state Theorem 4.25; the so called arithmetic regularity lemma for the U^3 norm for functions over $\mathbb{F}_p^n$. The general version of this decomposition theorem was established by Green and Tao in [15]. For us, the version of [13] is sufficient and we state it here. Some preparation is required.

I. Factors. Let $f_1, \dots, f_k \in \mathbb{C}^{\mathbb{F}_p^n}$. A σ-algebra of $\mathbb{F}_p^n$ with each of its atoms of the form

$$\{x \in \mathbb{F}_p^n : f_1(x) = z_1, \dots, f_k(x) = z_k\},$$

where $(z_1, \dots, z_k) \in \mathbb{C}$, is called a *factor* of $\mathbb{F}_p^n$. A factor of $\mathbb{F}_p^n$ each of whose atoms has the form $\{x \in \mathbb{F}_p^n : (r_1^{\mathrm{T}}x, \dots, r_k^{\mathrm{T}}x) = a\}$ where $r_1, \dots, r_k, a \in \mathbb{F}_p^k$ is called a *linear factor of complexity* k, and we say that this linear factor is *generated* by the values $r_1, \dots, r_k \in \mathbb{F}_p^n$. We use the members of $\mathbb{F}_p^k$ in order to represent the various atoms of such a linear factor by mapping a member $x \in \mathbb{F}_p^n$ to the vector $(r_1^{\mathrm{T}}x, \dots, r_k^{\mathrm{T}}x) \in \mathbb{F}_p^k$.

Definition 4.22. [Quadratic factor]
Let $r_1, \dots, r_{d_1} \in \mathbb{F}_p^n$ and let $M_1, \dots, M_{d_2}$ be symmetric $n \times n$ matrices over $\mathbb{F}_p$. Let $\mathcal{B}_1$ be the linear factor generated by $r_1, \dots, r_{d_1}$, and let $\mathcal{B}_2$ be the factor generated[9] *by the quadratic forms $x^{\mathrm{T}}M_1x, \dots, x^{\mathrm{T}}M_{d_2}x$ and the linear forms $r_1^{\mathrm{T}}x, \dots, r_{d_1}^{\mathrm{T}}x$. The pair $(\mathcal{B}_1, \mathcal{B}_2)$ is called a* quadratic factor of complexity (d_1, d_2).

Let $(\mathcal{B}_1, \mathcal{B}_2)$ be as in Definition 4.22. The atoms of $\mathcal{B}_1$ are indexed using the elements of $\mathbb{F}_p^{d_1}$ as described above. In a similar way, the atoms of $\mathcal{B}_2$ are indexed using the elements of $\mathbb{F}_p^{d_1} \times \mathbb{F}_p^{d_2}$; that is, we map an $x \in \mathbb{F}_p^n$ to the pair

$$(\Gamma(x), \Phi(x)) = ((r_1^{\mathrm{T}}x, \dots, r_{d_1}^{\mathrm{T}}x), (x^{\mathrm{T}}M_1x, \dots, x^{\mathrm{T}}M_{d_2}x)) \in \mathbb{F}_p^{d_1} \times \mathbb{F}_p^{d_2}.$$

We identify an atom of a quadratic factor with its index. Given a pair $(a, b) \in \mathbb{F}_p^{d_1} \times \mathbb{F}_p^{d_2}$ that indexes an atom of $\mathcal{B}_2$ and an element $x \in \mathbb{F}_p^n$, we write $x \in (a, b)$ to denote that x belongs to the atom whose index is (a, b), i.e., that $(\Gamma(x), \Phi(x)) = (a, b)$ holds.

II. For a σ-algebra $\mathcal{B}$ of $\mathbb{F}_p^n$ and $x \in \mathbb{F}_p^n$, we write $\mathcal{B}(x)$ to denote the atom of $\mathcal{B}$ containing x. Then, given $f \in \mathbb{C}^{\mathbb{F}_p^n}$ we write

$$\mathbb{E}\left(f|\mathcal{B}\right)(x) = \mathbb{E}_{\mathcal{B}(x)}\, f = \frac{1}{|\mathcal{B}(x)|} \sum_{y \in \mathcal{B}(x)} f(y) \tag{4.23}$$

to denote the average of f over the atom of $\mathcal{B}$ containing x. The function $\mathbb{E}\left(f|\mathcal{B}\right) : G \to \mathbb{C}$ is called the *conditional expectation of f with respect to $\mathcal{B}$.*

We say that $g \in \mathbb{C}^{\mathbb{F}_p^n}$ is *measurable with respect to $\mathcal{B}$* if it is constant on each atom of $\mathcal{B}$. By definition, the conditional expectation $\mathbb{E}(g|\mathcal{B})$ of any function $g \in \mathbb{C}^{\mathbb{F}_p^n}$ is measurable with respect to $\mathcal{B}$.

III. Let $A \subseteq \mathbb{F}_p^n$. Let us record here for future reference the fact that

$$\mathbb{E}(A|\mathcal{B}) \in [0, 1] \tag{4.24}$$

since $\mathbb{E}(\mathbf{1}_A|\mathcal{B})(x) = |A \cap \mathcal{B}(x)|/|\mathcal{B}(x)|$.

IV. We write $\operatorname{rk} M$ to denote the rank of a matrix M. A quadratic factor of complexity (d_1, d_2) satisfying

$$\operatorname{rk}\left(\lambda_1 M_1 + \cdots + \lambda_{d_2} M_{d_2}\right) \geq r$$

for any $\lambda_1, \dots, \lambda_{d_2} \in \mathbb{F}_p$ not all zero, where $M_1, \dots, M_{d_2}$ are the symmetric matrices involved in its generation, is said to have *rank* at least r.

[9] The atoms of $\mathcal{B}_2$ have the form

$$\{x \in \mathbb{F}_p^n : r_1^{\mathrm{T}}x = c_1, \dots, r_{d_1}^{\mathrm{T}}x = c_{d_1} \text{and } x^{\mathrm{T}}M_1x = z_1, \dots, x^{\mathrm{T}}M_{d_2}x = z_{d_2}\},$$

where $(r_1, \dots, r_{d_1}) \in \mathbb{F}_p^{d_1}$ and $(z_1, \dots, z_{d_2}) \in \mathbb{F}_p^{d_2}$.

V. We are ready to state the regularity lemma that we shall be using. This can be found in [13, Proposition 3.12] and also in [9, Theorem 3.5].

Theorem 4.25. (The arithmetic regularity lemma for U^3 in $\mathbb{F}_p^n$)
For every real $\delta > 0$ and every two growth functions $\omega_{\rm rk}, \omega_{\rm uni} : \mathbb{R}_+ \to \mathbb{R}_+$ (which may be independent of δ) there exists an n_0 such that for every integer $n \geq n_0$ the following holds.

For every function $f : \mathbb{F}_p^n \to [-1,1]$ there exists a constant d_0, a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$, and a decomposition $f = f_{\rm str} + f_{\rm uni} + f_{\rm neg}$ satisfying the following terms.

1. *The complexity of $(\mathcal{B}_1, \mathcal{B}_2)$ is at most (d_1, d_2) where $d_1, d_2 \leq d_0$;*
2. *the rank of $(\mathcal{B}_1, \mathcal{B}_2)$ is at least $\omega_{\rm rk}(d_1 + d_2)$;*
3. *and*
$$f_{\rm str} = \mathbb{E}(f|\mathcal{B}_2), \quad \|f_{\rm neg}\|_{L^2(\mathbb{F}_p^n)} \leq \delta, \quad \textit{and} \quad \|f_{\rm uni}\|_{U^3(\mathbb{F}_p^n)} \leq 1/\omega_{\rm uni}(d_1 + d_2).$$

§4.3 Counting 3SAPs along atoms of quadratic factors.

I. The aim of this section is to establish Lemma 4.31. The purpose of this lemma is to count 3SAPs along a function of the form of $f_{\rm str}$ (see Theorem 4.25) that we shall obtain after regularising the set A in the proof of Theorem 1.7 (see § 4.4).

II. As $f_{\rm str}$ is constant on the atoms of the quadratic factor (over which it is defined) this task of counting reduces to considering three atoms $(a^{(0)}, b^{(0)})$, $(a^{(1)}, b^{(1)})$, and $(a^{(2)}, b^{(2)})$ of a quadratic factor and counting triplets of the form $\{x, x+d, x+2d\}$ satisfying $d \in S$, $x \in (a^{(0)}, b^{(0)})$, $x + d \in (a^{(1)}, b^{(1)})$, and $x + 2d \in (a^{(2)}, b^{(2)})$.

III. For reasons arising in the proof of Theorem 1.7 we shall need to consider quadruplets of atoms instead of triplets of atoms. For further discussion on this issue see Paragraph § 4.4. IX.

We require some preparation.

IV. Throughout, $\omega = e^{2\pi i/p}$. The following is the well-known estimate for Gauss sums over $\mathbb{F}_p^n$ (cf., [9, Lemma 3.2] or [13, Lemma 3.1]).

Lemma 4.26. (Gauss sums over $\mathbb{F}_p^n$)
Let M be a symmetric $n \times n$ matrix over $\mathbb{F}_p$ of rank r and let $b \in \mathbb{F}_p^n$. Then,
$$\left| \mathbb{E}_{x \in \mathbb{F}_p^n} \omega^{x^{\rm T} M x + b^{\rm T} x} \right| \leq p^{-r/2}.$$

V. The size of an atom of a quadratic factor of $\mathbb{F}_p^n$ can be estimated as follows; see [9, Corollary 3.9] or [13, Lemma 4.2].

Lemma 4.27. (Size of an atom)
We have
$$\frac{\left|\{x \in \mathbb{F}_p^n : x \in (a,b)\}\right|}{p^n} = p^{-(d_1+d_2)} \pm p^{-r/2} \tag{4.28}$$
whenever $(a,b) \in \mathbb{F}_p^{d_1} \times \mathbb{F}_p^{d_2}$ is an atom of a quadratic factor of rank at least r and complexity at most (d_1, d_2).

VI. Viable quadruples of atoms. Let $S \subseteq \mathbb{F}_p^n$, and let $(\mathcal{B}_1, \mathcal{B}_2)$ be a quadratic factor of $\mathbb{F}_p^n$ of complexity at most (d_1, d_2). a 4AP $\{x, x+d, x+2d, x+3d\}$ with $d \in S$ satisfies

$$x \in (a^{(0)}, b^{(0)}), x+d \in (a^{(1)}, b^{(1)}), x+2d \in (a^{(2)}, b^{(2)}), x+3d \in (a^{(3)}, b^{(3)})$$

(and $(a^{(i)}, b^{(i)})$ are atoms of $(\mathcal{B}_1, \mathcal{B}_2)$) if and only if

$$(a^{(0)}, a^{(1)}, a^{(2)}, a^{(3)}) \text{ is a } 3\Gamma(S)\text{AP in } \mathbb{F}_p^{d_1}, \tag{4.29}$$

where Γ is as in § 4.2, and $\Gamma(S)$ is the image of S under Γ. In addition,

$$b^{(0)} - 3b^{(1)} + 3b^{(2)} - b^{(3)} = 0. \tag{4.30}$$

The necessity of (4.29) and (4.30) is trivial and can be seen by simply considering the expressions $\{\Gamma(x), \Gamma(x+d), \Gamma(x+2d), \Gamma(x+3d)\}$ and $\{\Phi(x), \Phi(x+d), \Phi(x+2d), \Phi(x+3d)\}$. The sufficiency can be seen through the proof of Lemma 4.31 below (shall indicate this); alternatively, one may consult [13, Lemma 4.3].

A quadruple of atoms $((a^{(0)}, b^{(0)}), (a^{(1)}, b^{(1)}), (a^{(2)}, b^{(2)}), (a^{(3)}, b^{(3)}))$ satisfying (4.29) and (4.30) is called *viable*.

VII. Counting. We arrive at the main lemma of this section. Here, given a viable quadruple of atoms $((a^{(0)}, b^{(0)}), (a^{(1)}, b^{(1)}), (a^{(2)}, b^{(2)}), (a^{(3)}, b^{(3)}))$ we count 3APs with gap in S along the first three atoms within the quadruple.

Lemma 4.31. (Counting 3SAPs along quadratic factors)
Let $S \subseteq \mathbb{F}_p^n$, let $(\mathcal{B}_1, \mathcal{B}_2)$ be a quadratic factor of $\mathbb{F}_p^n$ of rank at least r and complexity at most (d_1, d_2), and let $((a^{(0)}, b^{(0)}), (a^{(1)}, b^{(1)}), (a^{(2)}, b^{(2)}), (a^{(3)}, b^{(3)}))$ be viable. Then,

$$\begin{aligned} &|\,\{(x,d) \in \mathbb{F}_p^n \times S : x + jd \in (a^{(j)}, b^{(j)}), 0 \le j \le 2\}\,| = \\ &\qquad \left[p^{-2d_1 - 3d_2} \pm \left(\|S\|_u \|S\|^{-1}_{L^1(\mathbb{F}_p^n)} + 4p^{-r/2}\right)\right] p^n |S|. \end{aligned} \tag{4.32}$$

Proof. Put

$$X = \{(x,d) \in \mathbb{F}_p^n \times S : x + jd \in (a^{(j)}, b^{(j)}), 0 \le j \le 2\};$$

so that $|X|$ denotes the number of 3SAPs spanned by the triplet $((a^{(0)}, b^{(0)}), (a^{(1)}, b^{(1)}), (a^{(2)}, b^{(2)}))$.

We express $|X|$ as follows. By orthogonality relations of the characters of $\mathbb{F}_p^{d_1}$, a pair $(x,d) \in \mathbb{F}_p^n \times S$ belongs to the set

$$\{(x,d) \in \mathbb{F}_p^n \times S : \Gamma(x+jd) = a^{(j)}, 0 \le j \le 2\}$$

provided that

$$p^{-3d_1} \prod_{j=0}^{2} \sum_{\mu \in \mathbb{F}_p^{d_1}} \omega^{\mu^{\mathrm{T}}(\Gamma(x+jd) - a^{(j)})} = 1;$$

(sufficiency of (4.29) for 3SAPs is now given through the exponent of the summands). Similarly, by orthogonality relations of the characters of $\mathbb{F}_p^{d_2}$, a pair $(x,d) \in \mathbb{F}_p^n \times S$ belongs to the set

$$\{(x,d) \in \mathbb{F}_p^n \times S : \Phi(x+jd) = b^{(j)}, 0 \le j \le 2\}$$

provided that

$$p^{-3d_2} \prod_{j=0}^{2} \sum_{\lambda \in \mathbb{F}_p^{d_2}} \omega^{\lambda^{\mathrm{T}}(\Phi(x+jd) - b^{(j)})} = 1;$$

(if this term would have been written for 4APs then the sufficiency of (4.30) would be seen through the exponent of the summands).

Then

$$|X| = p^{-3d_1-3d_2} \sum_{x\in\mathbb{F}_p^n, d\in S} \prod_{j=0}^{2} \left[\prod_{i=1}^{d_1} \sum_{\mu_i^{(j)}\in\mathbb{F}_p} \omega^{\mu_i^{(j)}\left(r_i^{\mathrm{T}}(x+jd)-a_i^{(j)}\right)} \right] \left[\prod_{\ell=1}^{d_2} \sum_{\lambda_\ell^{(j)}\in\mathbb{F}_p} \omega^{\lambda_\ell^{(j)}\left((x+jd)^{\mathrm{T}}M_\ell(x+jd)-b_\ell^{(j)}\right)} \right].$$

Rearranging the above sum we arrive at

$$|X| = p^{-3d_1-3d_2} \sum_{x\in\mathbb{F}_p^n, d\in S} \sum_{\mu_i^{(j)},\lambda_\ell^{(j)}} \omega^{x^{\mathrm{T}}Px+d^{\mathrm{T}}Rx+d^{\mathrm{T}}Ld+u^{\mathrm{T}}x+v^{\mathrm{T}}d-h}, \tag{4.33}$$

where the inner sum ranges over $\mu_i^{(j)}, \lambda_\ell^{(j)} \in \mathbb{F}_p$ for $0 \le j \le 2$, $1 \le i \le d_1$, $1 \le \ell \le d_2$, and where

$$\begin{aligned}
P &= \sum_{i=1}^{d_2} \left(\lambda_i^{(0)} + \lambda_i^{(1)} + \lambda_i^{(2)}\right) M_i;\\
R &= \sum_{i=1}^{d_2} \left(2\lambda_i^{(1)} + 4\lambda_i^{(2)}\right) M_i;\\
L &= \sum_{i=1}^{d_2} \left(\lambda_i^{(1)} + 4\lambda_i^{(2)}\right) M_i;\\
u &= \sum_{i=1}^{d_1} \left(\mu_i^{(0)} + \mu_i^{(1)} + \mu_i^{(2)}\right) r_i;\\
v &= \sum_{i=1}^{d_1} \left(\mu_i^{(1)} + 2\mu_i^{(2)}\right) r_i;\\
h &= \sum_{j=0}^{2}\sum_{i=1}^{d_1} \mu_i^{(j)} a_i^{(j)} + \sum_{j=0}^{2}\sum_{i=1}^{d_2} \lambda_i^{(j)} b_i^{(j)}.
\end{aligned}$$

Note that each of the matrices P, R, and L is the sum of symmetric matrices. Also, as $(\mathcal{B}_1, \mathcal{B}_2)$ has rank at least r, then each of these matrices that is not identically zero has rank at least r.

Note that

$$\begin{aligned}
&\|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-2n} |X| = \\
&\qquad \|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-3d_1-3d_2} \sum_{\mu_i^{(j)},\lambda_\ell^{(j)}} \mathbb{E}_{x,d\in\mathbb{F}_p^n}\, S(d)\omega^{x^{\mathrm{T}}Px+d^{\mathrm{T}}Rx+d^{\mathrm{T}}Ld+u^{\mathrm{T}}x+v^{\mathrm{T}}d-h},
\end{aligned} \tag{4.34}$$

and let us focus now on the expectations appearing as summands on the right hand side of (4.34). In particular, let $E(P = 0, L = 0)$ denote the term

$$\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, S(d)\omega^{x^{\mathrm{T}}Px+d^{\mathrm{T}}Rx+d^{\mathrm{T}}Ld+u^{\mathrm{T}}x+v^{\mathrm{T}}d-h}$$

with P and L set to zero. In a similar manner, let us introduce the terms $E(P \neq 0, L = 0)$,

$E(P=0, L\neq 0)$, $E(P\neq 0, L\neq 0)$. We may rewrite (4.34) as

$$\begin{aligned}
&\|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-2n}|X| = \\
&\qquad \|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-3d_1-3d_2} \sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=L=0}} E(P=0, L=0) \\
&\qquad + \|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-3d_1-3d_2} \sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P\neq 0, L=0}} E(P\neq 0, L=0) \\
&\qquad + \|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-3d_1-3d_2} \sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=0, L\neq 0}} E(P=0, L\neq 0) \\
&\qquad + \|S\|_{L^1(\mathbb{F}_p^n)}^{-1} p^{-3d_1-3d_2} \sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P\neq 0, L\neq 0}} E(P\neq 0, L\neq 0)
\end{aligned} \tag{4.35}$$

1. The term $E(P=0, L=0)$. Let us write $E(P=0, L=0, R=0)$ and $E(P=0, L=0, R\neq 0)$ to denote the term $E(P=0, L=0)$ with R set to zero or otherwise, respectively. Recall that

   $$E(P=0, L=0) = \mathbb{E}_{d\in\mathbb{F}_p^n}\, S(d)\omega^{v^{\mathrm{T}}d}\, \mathbb{E}_{x\in\mathbb{F}_p^n}\, \omega^{d^{\mathrm{T}}Rx+u^{\mathrm{T}}x-h}. \tag{4.36}$$

   By orthogonality,

   $$E(P=0, L=0, R=0) = \mathbb{E}_{d\in\mathbb{F}_p^n}\, S(d)\omega^{v^{\mathrm{T}}d}\, \mathbb{E}_{x\in\mathbb{F}_p^n}\, \omega^{u^{\mathrm{T}}x-h}$$

   vanishes unless $u=0$. For $u=0$, we have

   $$E(P=0, L=0, R=0) = \widehat{S}(v)\omega^{-h}. \tag{4.37}$$

   Next, if $R\neq 0$, then the inner expectation of (4.36) vanishes unless $Rd+u=0$. This occurs to at most a p^{-r}-fraction of the values of d since R has rank at least r. As a result,

   $$\begin{aligned}
   |E(P=0, L\neq 0, R\neq 0)| &\leq p^{-r}\left|\mathbb{E}_{d\in\mathbb{F}_p^n}\, S(d)\omega^{v^{\mathrm{T}}d}\right| \\
   &= p^{-r}|\widehat{S}(v)| \\
   &\leq p^{-r}|\widehat{S}(0)| \\
   &= p^{-r}\|S\|_{L^1(\mathbb{F}_p^n)}.
   \end{aligned} \tag{4.38}$$

2. The term $E(P\neq 0, L=0)$. In this case, we have

   $$\begin{aligned}
   |E(P\neq 0, L=0)| &= \left|\mathbb{E}_{x\in\mathbb{F}_p^n}\, \omega^{x^{\mathrm{T}}Px+u^{\mathrm{T}}x+h}\, \mathbb{E}_{d\in\mathbb{F}_p^n}\, S(d)\omega^{x^{\mathrm{T}}Rd+v^{\mathrm{T}}d}\right| \\
   &= \left|\widehat{S}(Rx+v)\, \mathbb{E}_{x\in\mathbb{F}_p^n}\, \omega^{x^{\mathrm{T}}Px+u^{\mathrm{T}}x-h}\right| \\
   &\leq |\widehat{S}(0)|\left|\mathbb{E}_{x\in\mathbb{F}_p^n}\, \omega^{x^{\mathrm{T}}Px+u^{\mathrm{T}}x-h}\right| \\
   &\leq p^{-r/2}\|S\|_{L^1(\mathbb{F}_p^n)},
   \end{aligned} \tag{4.39}$$

   where the last inequality is due to Lemma 4.26.

3. The term $E(P=0, L\neq 0)$.

$$|E(P=0, L\neq 0)| = \left|\mathbb{E}_{d\in\mathbb{F}_p^n} S(d)\omega^{d^{\mathrm{T}}Ld+v^{\mathrm{T}}d-h}\,\mathbb{E}_{x\in\mathbb{F}_p^n}\,\omega^{d^{\mathrm{T}}Rx+u^{\mathrm{T}}x}\right|$$

The inner expectation in the above expression vanishes unless $Rd+u=0$ (in which case it is equal to one). Since R has rank at least r, we have that $Rd+u=0$ for at most a p^{-r}-fraction of the values of d. This then yields

$$|E(P=0, L\neq 0)| \;\leq p^{-r}\left|\mathbb{E}_{d\in\mathbb{F}_p^n} S(d)\omega^{d^{\mathrm{T}}Ld+v^{\mathrm{T}}d}\right| \leq p^{-r}\|S\|_{L^1(\mathbb{F}_p^n)}. \tag{4.40}$$

4. The $E(P\neq 0, L\neq 0)$.

$$\begin{aligned}|E(P\neq 0, L\neq 0)| &= \left|\mathbb{E}_{d\in\mathbb{F}_p^n} S(d)\omega^{d^{\mathrm{T}}Ld+v^{\mathrm{T}}d}\,\mathbb{E}_{x\in\mathbb{F}_p^n}\,\omega^{x^{\mathrm{T}}Px+d^{\mathrm{T}}Rx+u^{\mathrm{T}}x-h}\right|\\ &\leq \mathbb{E}_{d\in\mathbb{F}_p^n}\left|S(d)\omega^{d^{\mathrm{T}}Ld+v^{\mathrm{T}}d}\right|\left|\mathbb{E}_{x\in\mathbb{F}_p^n}\,\omega^{x^{\mathrm{T}}Px+d^{\mathrm{T}}Rx+u^{\mathrm{T}}x-h}\right|\\ &\leq p^{-r/2}\,\mathbb{E}_{d\in\mathbb{F}_p^n}\left|S(d)\omega^{d^{\mathrm{T}}Ld+v^{\mathrm{T}}d}\right|\\ &\leq p^{-r/2}\|S\|_{L^1(\mathbb{F}_p^n)},\end{aligned} \tag{4.41}$$

where the second inequality is due to Lemma 4.26.

Now, assertions (4.35) through (4.41) yield

$$\begin{aligned}&\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}p^{-2n}|X| = \\ &\quad \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=L=R=0\\ u=v=0}}|\widehat{S}(0)|\\ &\quad \pm\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=L=R=0\\ u=0,v\neq 0}}\|S\|_u\\ &\quad \pm p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=L=0,R\neq 0}}p^{-r}\\ &\quad \pm p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P\neq 0,L=0}}p^{-r/2}\\ &\quad \pm p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P=0,L\neq 0}}p^{-r}\\ &\quad \pm p^{-3d_1-3d_2}\sum_{\substack{\mu_i^{(j)},\lambda_\ell^{(j)}\\ P\neq 0,L\neq 0}}p^{-r/2}.\end{aligned} \tag{4.42}$$

Observe now that the total number of choices for $\mu_i^{(j)}$ and $\lambda_\ell^{(j)}$ (for i,j,ℓ ranging in their respective ranges) is $p^{3d_1+3d_2}$. Consequently, (4.42) assumes the form

$$\begin{aligned}&\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}p^{-2n}|X| = \\ &\quad p^{-3d_1-3d_2}\left|\{\mu_i^{(j)},\lambda_\ell^{(j)} : P=L=R=0, u=v=0\}\right|\\ &\quad \pm\left(\|S\|_u\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}+4p^{-r/2}\right).\end{aligned} \tag{4.43}$$

To prove (4.32) and thus conclude the proof of this lemma, it remains to estimate the size of the set $\{\mu_i^{(j)}, \lambda_\ell^{(j)} : P = L = R = 0, u = v = 0\}$. We begin by estimating the number of viable vectors $\mu^{(0)}$, $\mu^{(1)}$, and $\mu^{(2)}$. Here, the viable vectors are those that yield $u = 0$ and $v = 0$ and thus must satisfy

$$\begin{aligned}\mu_i^{(1)} + 2\mu_i^{(2)} &= 0,\\ \mu_i^{(0)} + \mu_i^{(1)} + \mu_i^{(2)} &= 0,\end{aligned}$$

for every $1 \leq i \leq d_1$ (where the equations are over $\mathbb{F}_p$). For the first equation, there are p options to choose $\mu_i^{(1)}$ and each determines a unique $\mu_i^{(2)}$ that together yield a solution over $\mathbb{F}_p$ for the first equation. With $\mu_i^{(1)}$ and $\mu_i^{(2)}$ determined by the first equation, the value of $\mu_i^{(0)}$ that would yield a solution to the second equation is unique. Thus, in total there are p^{d_1} options to choose the vectors $\mu^{(0)}$, $\mu^{(1)}$, and $\mu^{(2)}$ as to have $u = v = 0$.

We proceed to the estimation of the number of viable vectors $\lambda^{(0)}$, $\lambda^{(1)}$, and $\lambda^{(2)}$ that would yield $P = L = R = 0$. Here, the vectors must satisfy

$$\begin{aligned}\lambda_i^{(0)} + \lambda_i^{(1)} + \lambda_i^{(2)} &= 0,\\ \lambda_i^{(1)} + 2\lambda_i^{(2)} &= 0,\\ \lambda_i^{(1)} + 4\lambda_i^{(2)} &= 0,\end{aligned}$$

for every $0 \leq i \leq d_2$ (where the equations are over $\mathbb{F}_p$). The second and third equations show that here only the choice $\lambda^{(0)} = \lambda^{(1)} = \lambda^{(2)} = 0$ is valid.

We have then that $|\{\mu_i^{(j)}, \lambda_\ell^{(j)} : P = L = R = 0, u = v = 0\}| = p^{d_1}$, so that (4.32) now follows and this completes the proof of the lemma. ■

§4.4 Proof of Theorem 1.7.

I. Given A, S, α, and σ as in Theorem 1.7 we give a lower bound on

$$|\mathbb{E}_{x\in\mathbb{F}_p^n, d\in S}\, A(x)A(x+d)A(x+2d)|;$$

in particular we show that

$$\left|\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, A(x)A(x+d)A(x+2d)\mu_S(d)\right| \geq \alpha^4/2^6, \tag{4.44}$$

where $\mu_S(x) = S(x)\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}$.

II. Constants. Given α and σ set

$$\delta = \frac{\sigma\alpha^4}{6\cdot 2^6}, \quad \omega_{\text{uni}} = \delta^{-1}. \tag{4.45}$$

Next, set ω_{rk} so that for any two positive integers s and t,

$$p^{-\omega_{\text{rk}}(s+t)/2} \leq p^{-2s-3t}/8 - p^{-s}/4 \tag{4.46}$$

is satisfied; which in particular means that

$$p^{-\omega_{\text{rk}}(s+t)/2} \leq p^{-2(s+t)}/2 \tag{4.47}$$

Apply Theorem 4.25 with $\delta, \omega_{\text{rk}}$, and ω_{uni} in order to obtain a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ of $\mathbb{F}_p^n$ with complexity (d_1, d_2) and rank $r \geq \omega_{\text{rk}}(d_1 + d_2)$ together with a decomposition $A = f_{\text{str}} + f_{\text{uni}} + f_{\text{neg}}$ satisfying the conditions specified in Theorem 4.25. In addition, set

$$\eta = p^{-d_1}/2. \tag{4.48}$$

III. Consider the left hand side of (4.44). Replace the occurrences of A in (4.44) with $f_{\text{str}} + f_{\text{uni}} + f_{\text{neg}}$ one after the other as to attain (4.49) below. This is done as follows. First write

$$\begin{aligned} A(x)A(x+d)A(x+2d) = &f_{\text{str}}(x)A(x+d)A(x+2d) \\ &+ f_{\text{uni}}(x)A(x+d)A(x+2d) \\ &+ f_{\text{neg}}(x)A(x+d)A(x+2d). \end{aligned}$$

The terms not containing f_{str} are left as is; continue replacing the occurrences of A with its decomposition only in the term involving f_{str}; eventually one arrives at (4.49).

$$\begin{aligned} |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} A(x)A(x+d)A(x+2d)\mu_S(d)| = \\ &|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{str}}(x)f_{\text{str}}(x+d)f_{\text{str}}(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{neg}}(x)A(x+d)A(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{uni}}(x)A(x+d)A(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{str}}(x)f_{\text{neg}}(x+d)A(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{str}}(x)f_{\text{uni}}(x+d)A(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{str}}(x)f_{\text{str}}(x+d)f_{\text{neg}}(x+2d)\mu_S(d)| \\ &\pm |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{str}}(x)f_{\text{str}}(x+d)f_{\text{uni}}(x+2d)\mu_S(d)| \end{aligned} \tag{4.49}$$

The reason we replace the occurrences of A in this manner instead of replacing each occurrence of A in (4.44) and then cross multiply is due to the fact that our generalised von Neumann theorem (i.e., Corollary 4.19) requires that at most one function may have its infinity norm not bounded by one.

The fact that $\|A\|_\infty \le 1$ and $\|f_{\text{str}}\|_\infty \le 1$ (see (4.24)) has the following two consequences. The first of which is that each of the three terms above involving f_{uni} is at most

$$\|f_{\text{uni}}\|_{U^3(\mathbb{F}_p^n)}\|S\|^{-1}_{L^1(\mathbb{F}_p^n)} \le \omega_{\text{uni}}^{-1}\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}, \tag{4.50}$$

by (4.21).

The second consequence is that each of the three terms above involving f_{neg} is upper bounded by

$$\|f_{\text{neg}}\|_{L^1(\mathbb{F}_p^n)}\|S\|^{-1}_{L^1(\mathbb{F}_p^n)} \le \|f_{\text{neg}}\|_{L^2(\mathbb{F}_p^n)}\|S\|^{-1}_{L^1(\mathbb{F}_p^n)} \le \delta\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}. \tag{4.51}$$

To see this, consider, for example, the term $|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{neg}}(x)f_{\text{str}}(x+d)A(x+2d)\mu_S(d)|$ which is equivalent to one of the terms above by change of variable. We may write

$$\begin{aligned} &|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\text{neg}}(x)f_{\text{str}}(x+d)A(x+2d)\mu_S(d)| \\ &\qquad = \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\text{neg}}(x)f_{\text{str}}(x+d)A(x+2d)S(d)| \\ &\qquad \le \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}\,\mathbb{E}_{x\in\mathbb{F}_p^n}\,|f_{\text{neg}}(x)||\,\mathbb{E}_{d\in\mathbb{F}_p^n}\, f_{\text{str}}(x+d)A(x+2d)S(d)|. \end{aligned}$$

Observe that $|f_{\text{str}}(x+d)A(x+2d)S(d)| \le 1$ for any $x, d \in \mathbb{F}_p^n$; consequently, the function $E(x) = \mathbb{E}_{d\in\mathbb{F}_p^n}\, f_{\text{str}}(x+d)A(x+2d)S(d)$ satisfies $\|E\|_\infty \le 1$. This then yields that

$$\begin{aligned} &|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\text{neg}}(x)f_{\text{str}}(x+d)A(x+2d)\mu_S(d)| \\ &\qquad \le \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}\,\mathbb{E}_{x\in\mathbb{F}_p^n}\,|f_{\text{neg}}(x)| = \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}\|f_{\text{neg}}\|_{L^1(\mathbb{F}_p^n)}; \end{aligned}$$

so that (4.51) follows.

From (4.49), (4.50), and (4.51) it follows that

$$\begin{aligned} |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, A(x)A(x+d)A(x+2d)\mu_S(d)| & \\ \geq |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n}\, f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)|\|S\|^{-1}_{L^1(\mathbb{F}_p^n)} & \\ - 3\left(\delta+\omega_{\mathrm{uni}}^{-1}\right)\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}. & \end{aligned} \tag{4.52}$$

IV. We focus on the expectation seen on the right hand side of (4.52). To that end, let us write $F_{\mathrm{str}}(a,b)$ to denote the (single) value that f_{str} assumes on the atom $(a,b)\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}$. Then,

$$\begin{aligned} |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)| & \\ \geq p^{-2n}\sum_{\substack{(a^{(i)},b^{(i)})\\ 0\leq i\leq 3\\ \text{viable}}} \Lambda(0,1,2)F_{\mathrm{str}}(a^{(0)},b^{(0)})F_{\mathrm{str}}(a^{(1)},b^{(1)})F_{\mathrm{str}}(a^{(2)},b^{(2)}); & \end{aligned} \tag{4.53}$$

here the sum ranges over viable quadruples of atoms $(a^{(i)},b^{(i)})\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}$ and where

$$\Lambda(0,1,2)=|\{(x,d)\in\mathbb{F}_p^n\times S : x+id\in(a^{(i)},b^{(i)}), 0\leq i\leq 2\}|$$

is the number of 3SAPs spanned by the first three atoms in the viable quadruple $(a^{(i)},b^{(i)})\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}$, $0\leq i\leq 3$, so that

$$\Lambda(0,1,2)=\left[p^{-2d_1-3d_2}\pm\left(\|S\|_u\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}+4p^{-r/2}\right)\right]p^n|S|,$$

by Lemma 4.31.

As $F_{\mathrm{str}}(a,b)\in[0,1]$ for every atom $(a,b)\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}$, by (4.24), the right hand side of (4.53) is

$$\begin{aligned} \geq \sigma\,[p^{-2d_1-3d_2}-(\ \|S\|_u\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}+4p^{-r/2}\ )\,]\times & \\ \sum_{\substack{(a^{(i)},b^{(i)})\\ 0\leq i\leq 3\\ \text{viable}}} F_{\mathrm{str}}(a^{(0)},b^{(0)})F_{\mathrm{str}}(a^{(1)},b^{(1)})F_{\mathrm{str}}(a^{(2)},b^{(2)})F_{\mathrm{str}}(a^{(3)},b^{(3)}), & \end{aligned}$$

which is

$$\begin{aligned} \geq \sigma\,[p^{-2d_1-3d_2}-(\ \eta+4p^{-r/2}\ )\,]\times & \\ \sum_{\substack{(a^{(i)},b^{(i)})\\ 0\leq i\leq 3\\ \text{viable}}} F_{\mathrm{str}}(a^{(0)},b^{(0)})F_{\mathrm{str}}(a^{(1)},b^{(1)})F_{\mathrm{str}}(a^{(2)},b^{(2)})F_{\mathrm{str}}(a^{(3)},b^{(3)}). & \end{aligned}$$

Note now that $\eta+4p^{-r/2}\leq p^{-2d_1-3d_2}/2$ by (4.46) and (4.48). Consequently, we have

$$\begin{aligned} |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)|\geq & \\ (\sigma p^{-2d_1-3d_2}/2)\sum_{\substack{(a^{(i)},b^{(i)})\\ 0\leq i\leq 3\\ \text{viable}}} F_{\mathrm{str}}(a^{(0)},b^{(0)})F_{\mathrm{str}}(a^{(1)},b^{(1)})F_{\mathrm{str}}(a^{(2)},b^{(2)})F_{\mathrm{str}}(a^{(3)},b^{(3)}). & \end{aligned} \tag{4.54}$$

V. Let now $H = \langle r_1, \ldots, r_{d_1} \rangle \le \mathbb{F}_p^n$ be the subgroup generated by all of the linear forms defining the atoms of $\mathcal{B}_1$; and let $H^\perp = \{x \in \mathbb{F}_p^n : \ell^{\mathrm{T}} x = 0 \ \forall \ell \in H\}$ be the orthogonal complement of H. Then $H^\perp \le \mathbb{F}_p^n$. Next, as $|H||H^\perp| = p^n$ and $|H| \le p^{d_1}$, we have that

$$|H^\perp| \ge p^{n-d_1}. \tag{4.55}$$

We claim that

$$S \cap H^\perp \text{ is nonempty.} \tag{4.56}$$

To see this suffices to show that

$$\left| \mathbb{E}_{x\in\mathbb{F}_p^n} S(x) H^\perp(x) \right| > 0.$$

Indeed, by Fourier inversion and orthogonality relations of the characters of $\mathbb{F}_p^n$ we have that

$$\begin{aligned}\left| \mathbb{E}_{x\in\mathbb{F}_p^n} S(x) H^\perp(x) \right| &= \left| \sum_{\xi\in\widehat{\mathbb{F}_p^n}} \widehat{S}(\xi)\widehat{H^\perp}(-\xi) \right| \\ &= \widehat{S}(0)\widehat{H^\perp}(0) \pm \|S\|_u \sum_{\xi\in\widehat{Fpn}} \left| \widehat{H^\perp}(\xi) \right|.\end{aligned}$$

Recall now that

$$|\widehat{H^\perp}(\xi)| = \begin{cases} \frac{|H^\perp|}{p^n}, & \xi \in (H^\perp)^\perp = H, \\ 0, & \text{otherwise;} \end{cases}$$

so that $\sum_{\xi\in\widehat{Fpn}} \left| \widehat{H^\perp}(\xi) \right| = 1$. Then

$$\left| \mathbb{E}_{x\in\mathbb{F}_p^n} S(x) H^\perp(x) \right| = \sigma \tfrac{|H^\perp|}{p^n} \pm \eta\sigma;$$

now as $\frac{|H^\perp|}{p^n} \ge p^{-d_1}$ and $\eta < p^{-d_1}$, by (4.48), the assertion (4.56) follows.

VI. We return to (4.54). We have that $0 \in \Gamma(S)$, by (4.56), so that for every $a \in \mathbb{F}_p^{d_1}$ the quadruple (a,a,a,a) forms a $4\Gamma(S)$AP in $\mathbb{F}_p^{d_1}$. Consequently,

$$\begin{aligned}&\sum_{\substack{(a^{(i)},b^{(i)})\\ 0\le i\le 3\\ \text{viable}}} F_{\mathrm{str}}(a^{(0)},b^{(0)})F_{\mathrm{str}}(a^{(1)},b^{(1)})F_{\mathrm{str}}(a^{(2)},b^{(2)})F_{\mathrm{str}}(a^{(3)},b^{(3)}) \\ &\quad\ge \sum_{a\in\mathbb{F}_p^{d_1}} \sum_{\substack{b^{(i)})\in\mathbb{F}_p^{d_2}\\ 0\le i\le 3\\ b^{(0)}-3b^{(1)}+3b^{(2)}-b^{(3)}=0}} F_{\mathrm{str}}(a,b^{(0)})F_{\mathrm{str}}(a,b^{(1)})F_{\mathrm{str}}(a,b^{(2)})F_{\mathrm{str}}(a,b^{(3)}) \\ &\quad= \sum_{a\in\mathbb{F}_p^{d_1}} \sum_{x\in\mathbb{F}_p^{d_2}} \sum_{\substack{b,b'\in\mathbb{F}_p^{d_2}\\ b'-3b=x}} \left(F_{\mathrm{str}}(a,b)F_{\mathrm{str}}(a,b')\right)^2.\end{aligned}$$

Next, we apply Cauchy-Schwarz twice to obtain

$$\geq p^{-d_2} \sum_{a\in\mathbb{F}_p^{d_1}} \left(\sum_{x\in\mathbb{F}_p^{d_2}} \sum_{\substack{b,b'\in\mathbb{F}_p^{d_2}\\ b'-3b=x}} F_{\mathrm{str}}(a,b)F_{\mathrm{str}}(a,b') \right)^2$$

$$\geq p^{-d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{x\in\mathbb{F}_p^{d_2}} \sum_{\substack{b,b'\in\mathbb{F}_p^{d_2}\\ b'-3b=x}} F_{\mathrm{str}}(a,b)F_{\mathrm{str}}(a,b') \right)^2 .$$

Rewriting the sums we attian:

$$= p^{-d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{x\in\mathbb{F}_p^{d_2}} \sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b)F_{\mathrm{str}}(a,x+3b) \right)^2$$

$$= p^{-d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \sum_{x\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,x+3b) \right)^2 .$$

For a fixed $b \in \mathbb{F}_p^{d_2}$ the term $x+3b$ ranges over the entire group $\mathbb{F}_p^{d_2}$ as x ranges over $\mathbb{F}_p^{d_2}$ so we may write

$$= p^{-d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \sum_{y\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,y) \right)^2$$

$$= p^{-d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \left(\sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \right)^2 \right)^2 .$$

Applying Cauchy-Schwarz once more yields

$$\geq p^{-d_1-d_2} \left(p^{-d_1} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \right)^2 \right)^2$$

$$= p^{-2d_1-d_2} \left(\sum_{a\in\mathbb{F}_p^{d_1}} \sum_{b\in\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \right)^4 . \tag{4.57}$$

VII. By (4.54) and (4.57) we now have

$$|\, \mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)| \geq (\sigma p^{-4d_1-4d_2}/2) \left(\sum_{(a,b)\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \right)^4$$

$$= (\sigma/2) \left(\mathbb{E}_{(a,b)\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) \right)^4 . \tag{4.58}$$

Lemma 4.27 implies that

$$\mathbb{E}_{(a,b)\in\mathbb{F}_p^{d_1}\times\mathbb{F}_p^{d_2}} F_{\mathrm{str}}(a,b) = \alpha\left(1 \pm p^{2d_1+2d_2-r/2}\right) \tag{4.59}$$

(see, e.g., [13, Equation (4.13)]). This together with (4.58) implies that

$$|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)| \geq \tfrac{1}{2}\sigma\alpha^4\left(1 - p^{2d_1+2d_2-r/2}\right)^4.$$

By (4.47), $p^{2d_1+2d_2-r/2} \leq 1/2$; so we get

$$|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)| \geq 2^{-5}\sigma\alpha^4. \tag{4.60}$$

VIII. We are in a position to conclude our proof. By (4.52) and (4.60), we have

$$\begin{aligned}|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} A(x)A(x+d)A(x+2d)\mu_S(d)| &\geq \|S\|^{-1}_{L^1(\mathbb{F}_p^n)}2^{-5}\sigma\alpha^4 - 3(\delta+\omega_{\mathrm{uni}}^{-1})\|S\|^{-1}_{L^1(\mathbb{F}_p^n)}\\ &\overset{(4.45)}{=} 2^{-5}\alpha^4 - 6\sigma^{-1}\delta\\ &\overset{(4.45)}{\geq} 2^{-6}\alpha^4.\end{aligned}$$

This concludes our proof of Theorem 1.7. ■

IX. One now sees that viable quadruples of atoms were useful in order to reach (4.57) where we see that the density of f_{str} appears; over this density we have a lower bound of the form (4.59).

Is the use of viable quadruples "necessary" in our approach? Could an alternative approach be that we ignore quadratic nature of the quadratic factors and then count 3SAPs in the counting Lemma 4.31 only on the linear atoms along a coarser approximation of A than f_{str}, namely $\tilde{f}_{\mathrm{str}}(a) = \sum_{b\in\mathbb{F}_p^{d_2}} f_{\mathrm{str}}(a,b) = \mathbb{E}(A|\mathcal{B}_1)(a)$ and $a \in \mathbb{F}_p^{d_1}$?

Our answer is No. The "weakness"[10] of our generalised von Neumann theorem (i.e., Lemma 4.6) is that it forces us to apply the arithmetic regularity lemma of the U^3 norm and to control the pseudorandom part of the decomposition the regularity lemma requires the structured approximation to have a quadratic nature (unless one comes up with a new proof of the regularity lemma).

Thus, in this approach one is forced to confront the term

$$|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)|.$$

For the "alternative" approach to work within the framework here, we would now have to prove that

$$|\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} f_{\mathrm{str}}(x)f_{\mathrm{str}}(x+d)f_{\mathrm{str}}(x+2d)S(d)| \geq |\,\mathbb{E}_{x,d\in\mathbb{F}_p^n} \tilde{f}_{\mathrm{str}}(x)\tilde{f}_{\mathrm{str}}(x+d)\tilde{f}_{\mathrm{str}}(x+2d)S(d)|.$$

Translating this to "atom language" one sees that this is entirely not clear due to the fact that f_{str} and $\tilde{f}_{\mathrm{str}}$ are constant on different sets. In particular, given a quadratic atom (a,b) we may have that $f_{\mathrm{str}}(a,b) = 0$ while $\tilde{f}_{\mathrm{str}}(a) > 0$; put another way the distribution of A would have to be taken into account and over this we have no assumptions.

[10] In fact, we do not see how to improve it.

References

[1] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753.

[2] P. Candela, *On the structure of steps of three-term arithmetic progressions in a dense set of integers*, Bull. Lond. Math. Soc. **42** (2010), no. 1, 1–14.

[3] M. Christ, *On random multilinear operator inequalities*, 2011.

[4] P. J. Davis, *Circulant matrices*, John Wiley & Sons, New York-Chichester-Brisbane, 1979, A Wiley-Interscience Publication, Pure and Applied Mathematics.

[5] N. Frantzikinakis, E. Lesigne, and M. Wierdl, *Random sequences and pointwise convergence of multiple ergodic averages*, Indiana Univ. Math. J. **61** (2012), no. 2, 585–617.

[6] ———, *Random differences in szemer\ 'edi's theorem and related results*, 2013.

[7] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), no. 3, 529–551.

[8] ———, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.

[9] W. T. Gowers and J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. (3) **100** (2010), no. 1, 155–176.

[10] ———, *Linear forms and quadratic uniformity for functions on $\mathbb{Z}_N$*, J. Anal. Math. **115** (2011), 121–186.

[11] A. Granville and Z. Rudnick (eds.), *Equidistribution in number theory, an introduction*, NATO Science Series II: Mathematics, Physics and Chemistry, vol. 237, Dordrecht, Springer, 2007.

[12] B. Green, *On arithmetic structures in dense sets of integers*, Duke Math. J. **114** (2002), no. 2, 215–238.

[13] B. Green, *Montréal notes on quadratic Fourier analysis*, Additive combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 69–102.

[14] B. Green and T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinb. Math. Soc. (2) **51** (2008), no. 1, 73–153.

[15] ———, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.

[16] Alan J. Hoffman, *On eigenvalues and colorings of graphs*, Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., 1969), Academic Press, New York, 1970, pp. 79–91.

[17] L. Lovász, *Combinatorial problems and exercises*, second ed., North-Holland Publishing Co., Amsterdam, 1993.

[18] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[19] T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2010, Paperback edition [of MR2289012].