CONSTELLATIONS IN $\mathbb{P}^d$

BRIAN COOK ÁKOS MAGYAR

Abstract. Let A be a subset of positive relative upper density of $\mathbb{P}^d$, the d-tuples of primes. We prove that A contains an affine copy of any set $e \subseteq \mathbb{Z}^d$, as long as e is in *general position* in the sense that the set $e \cup \{0\}$ has at most one point on every coordinate hyperplane.

1. Introduction.

1.1. Background. The celebrated theorem of Green and Tao [4] states that subsets of positive relative upper density of the primes contain an affine copy of any finite set of the integers, in particular contain arbitrary long arithmetic progressions. It is natural to ask if similar results hold in the multi-dimensional settings, especially in light of the multi-dimensional extensions of the closely related theorem of Szemerédi [8] on arithmetic progressions in dense subsets of the integers. Indeed such a result was obtained by Tao [9], showing that the Gaussian primes contain arbitrary constellations. In the same paper the problem of finding constellations in dense subsets of $\mathbb{P}^d$ was raised and briefly discussed.

The difficulty in this setting comes from two facts. First, the natural majorant of the d-tuples of primes is not pseudo-random with respect to the box norms, which replace the Gowers' uniformity norms in the multi-dimensional case. This may be circumvented by assuming the set e is in *general position* as described below, as is already suggested in [9]. However even under the this non-degeneracy assumption, the so-called *correlation conditions* in [4] do not seem to be sufficient, and a key observation of this note is to use more general correlation conditions to obtain the dual function estimates in the multi-dimensional case. Also, we need a more abstract form of the *transference principle* of Green and Tao [4]. The formulation we use is due to Gowers [3], however essentially equivalent results have been obtained by Tao and Ziegler [10], as well as by Reingold et al. [7].

Finally let us note that we expect the main result of this paper remains true for sets which are not in general position. For example in the simplest case, when $e = \{(x, y), (x + d, y), (x, y + d)\}$, it is easy to see that both subsets of the form $A = B \times C$ and random subsets $A \subseteq \mathbb{P}^2$ of positive relative density, contain many affine copies of e. However to prove such a result, our approach needs to modified in an essential way, as the box norms do not seem to control such constellations in the relative setting.

1.2. Main Results. Let $e = \{e_1, \dots, e_l\} \in (\mathbb{Z}^d)^l$ be a set of vectors; a constellation defined by e is then a set $e' = \{x, x + te_1, \dots, x + te_l\}$ where $t \neq 0$ is a scalar, that is an affine image of the set $e \cup \{0\}$.

Definition 1.1. *We say that a set of l vectors $e \in (\mathbb{Z}^d)^l$ is in general position, if $|\pi_i(e \cup \{0\})| = l+1$ for each i, where π_i is the orthogonal projection to the i^{th} coordinate axis.*

Let us also recall that a subset A of the d-tuples of primes $\mathbb{P}^d$ is of positive upper relative density if

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]^d|}{\pi(N)^d} > 0$$

The second author was supported by NSERC Grant 22R44824.

Our main result is then the following

Theorem 1.1. *Given any set $A \subseteq \mathbb{P}^d$ of positive relative upper density, we have that A contains infinitely many constellations defined by a set of vectors $e \in (\mathbb{Z}^d)^l$ in general position.*

Remarks: We note that for $d = 1$ this translates back the above described theorem of Green and Tao [4], as any finite subset of $\mathbb{Z}$ is in general position.

Also, one may assume that $l = d$ and the set $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}^d$ forms a basis in $\mathbb{R}^d$ besides being in general position, by passing to higher dimensions. Indeed, if $e \in (\mathbb{Z}^d)^l$ then let $\{f_1, \ldots, f_l\} \subseteq \mathbb{Z}^l$ be linearly independent vectors, and define a basis $e' = \{e'_1 = (e_1, f_1), \ldots, e'_l = (e_l, f_l), e'_{l+1}, \ldots, e'_{l+d}\} \subseteq \mathbb{Z}^{d+l}$ by extending the linearly independent set of vectors $e'_i = (e_i, f_i)$, $(1 \leq i \leq l)$. If e was in general position then it is easy to make the construction so that e' is also in general position, and if the set $A' := A \times \mathbb{P}^l$ contains a constellation $x' + te'$, then A contains $x + te$. Thus from now on we will always assume that e is also a basis of $\mathbb{R}^d$.

Theorem 1.1 may be viewed as a relative version of the so-called Multidimensional Szemerédi Theorem [1], stating that any subset of $\mathbb{Z}^d$ of positive upper density contains infinitely many constellations defined by any finite set of vectors $e \subseteq \mathbb{Z}^d$. As is customary, we will work in the "finitary" settings, when the underlying space is the group $\mathbb{Z}_N^d = (\mathbb{Z}/N\mathbb{Z})^d$, N being a large prime. In this settings we need the following, more quantitative version:

Theorem A (Furstenberg-Katznelson [1]). *Let $\alpha > 0$, $d \in \mathbb{N}$ and let $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}_N^d$ be a fixed set of vectors. If $f : \mathbb{Z}_N^d \to [0,1]$ is a given function such that $\mathbb{E}(f(x) : x \in \mathbb{Z}_N^d) \geq \alpha$, then one has*

$$\mathbb{E}(f(x)f(x+te_1)\ldots f(x+te_d) :\ x \in \mathbb{Z}_N^d,\ t \in \mathbb{Z}_N) \geq c(\alpha, e) \tag{1.1}$$

where $c(\alpha, e) > 0$ is a constant depending only on α and the set e.

Here we used the "expectation" notation: $\mathbb{E}(f(x) :\ x \in A) = \frac{1}{|A|}\sum_{x\in A} f(x)$.

In the relative setting, when $A \subseteq \mathbb{P}^d$, the condition: $\mathbb{E}(f(x) : x \in \mathbb{Z}_N^d) \geq \alpha$ (after identifying $[1,N]^d$ with $\mathbb{Z}_N^d$) does not hold for the indicator function $f = \mathbf{1}_A$, however it holds for $f = \mathbf{1}_A \Lambda^d$ where Λ^d is the d-fold tensor product of the von Mangoldt function Λ. The price one pays is that the function f is no longer bounded uniformly in N. Following the strategy of [4] we will show that the d-fold tensor product $\otimes^d \nu$ of the pseudo-random measure ν used in [4] is sufficiently random in our settings in order to apply the transference principle of [3]; we will refer to such measures ν as *d-pseudo-random* measures. We postpone the definition of d-pseudo-random measures to the next section, but state our main result in the finitary settings below:

Theorem 1.2. *Let $\alpha > 0$ be given, and d be fixed. There exists a constant $c(\alpha, e) > 0$ such that the following holds. If $0 \leq f \leq \mu$ is a given function on $\mathbb{Z}_N^d$ such that $\mu = \otimes^d \nu$ where ν is d-pseudo-random, and $\mathbb{E}(f(x) :\ x \in \mathbb{Z}_N^d) \geq \alpha$, then for any basis $e = \{e_1, ..., e_d\}$ in general position, we have that*

$$\mathbb{E}(f(x)f(x+te_1)...f(x+te_d) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) \geq c(\alpha, e) \tag{1.2}$$

1.3. **Norms, Transference, and Pseudo-random Measures.** First we introduce the d-dimensional box norms. We actually introduce one norm for each linearly independent set of vectors $\{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$.

For a function $f : \mathbb{Z}_N^d \to \mathbb{C}$ this norm with respect to a basis e is given by

$$||f||_{\square(e)^d}^{2^d} = \mathbb{E}\Big(\prod_{\omega\in\{0,1\}^d} f(x+\omega te) : x \in \mathbb{Z}_N^d,\ t \in \mathbb{Z}_N^d\Big)$$

with the notation $\omega te = \omega_1 t_1 e_1 + ... + \omega_d t_d e_d$.

That this is actually a norm is not immediate, but for the standard basis it can be shown by repeated applications of the Cauchy-Schwarz inequality, similarly as for the Gowers norms (see for example [2]). For a different basis, note that we have $||f||_{\Box(e)^d} = ||f \circ T||_{\Box^d}$ for an appropriate linear transformation T, where $||f||_{\Box^d}$ is the norm with respect to the standard basis. The same way one shows [2] that the analogue of the so-called Gowers-Cauchy-Schwarz inequality holds

Proposition 1.1. *($\Box^d(e)$-Cauchy-Schwarz inequality)*
Given 2^d functions, indexed by elements of $\{0,1\}^d$, we have

$$\langle f_\omega : \omega \in \{0,1\}^d \rangle = \mathbb{E}(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega t e) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N^d) \leq \prod_{\omega \in \{0,1\}^d} ||f_\omega||_{\Box(e)^d}$$

Gowers presents an alternative approach to the Green-Tao Transference Theorem from a more functional analytic point of view, making use of the Hahn-Banach Theorem. The specific version he provides will be presented below after we recall some definitions. First we note that $||\cdot||^*$ is the defined to be the dual norm of $||\cdot||$.

Definition 1.2. *Let $||\cdot||$ be a norm on $\mathcal{H} = L^2(\mathbb{Z}_n)$ such that $||f||_{L^\infty} \leq ||f||^*$, and let $X \subseteq \mathcal{H}$ be bounded. Then $||\cdot||$ is a quasi algebra predual (QAP) norm with respect to X if there exists an operator $\mathcal{D} : \mathcal{H} \to \mathcal{H}$, a positive function c on $\mathbb{R}$ and an increasing positive function C on $\mathbb{R}$ satisfying:*
(i) $\langle f, \mathcal{D}f \rangle \leq 1$ for all $f \in X$,
(ii) $\langle f, \mathcal{D}f \rangle \geq c(\epsilon)$ for every $f \in X$ with $||f|| \geq \epsilon$, and
(iii) $||\mathcal{D}f_1 ... \mathcal{D}f_K||^* \leq C(K)$ for any $f_1, ..., f_K \in X$.

This definition in enough to state the transference principle.

Theorem B. *(Gowers [3]) Let μ and ω be non-negative functions on Y, Y finite, with $||\mu||_{L^1}, ||\omega||_{L^1} \leq 1$, and $\eta, \delta > 0$ be given parameters. Also let $||\cdot||$ be a QAP norm with respect to X, the set of all functions bounded above by $\max\{\mu, \omega\}$ in absolute value. There exists an $\epsilon > 0$ such that the following holds: If we have that $||\mu - \omega|| < \epsilon$, then for every function with $0 \leq f \leq \mu$ there exists a function g with $0 \leq g \leq \omega/(1-\delta)$ and $||f - g|| \leq \eta$.*

Remarks: By a simple re-scaling of the norms the constants 1 in Definition 1.2 and Theorem B can be replaced by any other fixed constants. The actual form given by Gowers is more explicit, in fact giving a specific choice of ϵ. However, for our purposes, we only need such an ϵ that is independent of the size of Y. Also, for our purpose one may choose $\omega \equiv 1$ and $\delta = 1/2$.

The definition of a pseudo-random measure in this paper will be slightly stronger than that of Green and Tao, adapted to the higher dimensional settings. Let us begin with the one dimensional case. Following [4], we define a *measure* to be a function $\nu : \mathbb{Z}_N \to \mathbb{R}$ to be a non-negative function such that

$$\mathbb{E}(\nu(x) : x \in \mathbb{Z}_N) = 1 + o(1).$$

where the $o(1)$ notation means a quantity which tends to 0 as $N \to \infty$. A measure will be deemed pseudo-random if it satisfies two properties at a specific level. The first of these is known as the linear forms condition, as we will use only forms with integer coefficients we need a slightly simplified version.

Definition 1.3. *(Green-Tao [4]) Let ν be a measure, and $m_0, t_0 \in \mathbb{N}$ be small parameters. Then ν satisfies the (m_0, t_0)-linear forms condition if the following holds. For $m \leq m_0$ and $t \leq t_0$ arbitrary, suppose that $\{L_{i,j}\}_{1 \leq i \leq m, 1 \leq j \leq t}$ are integers, and that b_i are arbitrary elements of $\mathbb{Z}_N$. Given m linear forms $\phi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$ with*

$$\phi_i(x) = \sum_{j=1}^{t} L_{i,j} x_j + b_i,$$

$x = (x_1, ..., x_t)$ and $b = (b_1, ..., b_t)$, if we have that each ψ_i is nonzero and that they are pairwise linearly independent, then

$$\mathbb{E}\left(\prod_{i=1}^{m} \nu(\phi_i(x)) : x \in \mathbb{Z}_N^t\right) = 1 + o(1), \tag{1.3}$$

where the $o(1)$ term is independent of the choice of the b_i's.

The next condition is referred to as the correlation condition.

Definition 1.4. *Let ν be a measure. Then ν satisfies the (m_0, m_1) correlation condition if for every $1 \leq m \leq m_0$ there exists a function $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}_+$ such that for all $k \in \mathbb{N}$*

$$\mathbb{E}(\tau^k(x) : x \in \mathbb{Z}_N) = O_{m,k}(1)$$

and also

$$\mathbb{E}\left(\prod_{i=1}^{m_1}\prod_{j=1}^{m_0} \nu(\phi_i(y) + h_{i,j}) : y \in \mathbb{Z}_N^r\right) \leq \prod_{i=1}^{m_0}\left(\sum_{1\leq j<j'\leq m_0} \tau(h_{i,j} - h_{i,j'})\right) \tag{1.4}$$

where the functions $\phi_i : \mathbb{Z}_N^r \to \mathbb{Z}_N$ are pairwise independent linear forms.

Remarks:
This is a stronger condition that what is used in [4], in fact they used the special case when $m_1 = 1$, and ϕ is the identity. We define below a d-pseudo-random measure to be a measure satisfying these conditions at specific levels.

Definition 1.5. *We call a measure ν a d-pseudo-random if if ν satisfies the $((d^2 + 2d)2^{d-1}, 2d^2 + d)$-linear forms condition and the $(d, 2^d)$-correlation condition*

We will deal with d-fold tensor product of measures, $\mu = \otimes_{i=1}^{d}\nu$ and call them d-measures. We will call such a d-measure μ to be pseudo-random if the corresponding measure ν is d-pseudo-random. Finally, note that for a d-measure

$$\mathbb{E}(\mu(x) : x \in \mathbb{Z}_N^d) = \prod_{i=1}^{d} \mathbb{E}(\nu(x_i) : x_i \in \mathbb{Z}_N) = 1 + o(1).$$

1.4. **Outline of the Paper.** In Sections 2-3 we prove two key propositions, the so-called *generalized von Neumann inequality* and the *dual function estimate.* The first roughly says that the number of constellations defined by a set e is controlled by the appropriate box norm. The second is the essential step in showing that the box norms are QAP norms.

In Section 4, we prove our main results assuming that the measure exhibited in [4] is also d-pseudo-random in the sense defined above. First we show Theorem 1.2, which follows then easily from the Transference Principle, that is from Theorem B. Next, we prove Theorem 1.1 by a standard argument passing from $\mathbb{Z}_N$ to $\mathbb{Z}$.

Finally, in an Appendix, we prove d-pseudo-randomness of the measure ν used by Green and Tao, slightly modifying their arguments of Sec.10 in [4] based on earlier work of Goldston and Yıldırım [5] [6].

2. The Generalized von Neumann inequality.

Let $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}_N^d$ be a base of $\mathbb{Z}_N^d$ which is also in general position, which in this settings means that $|\pi_i(e \cup \{0\})| = d + 1$ for each i where $\pi_i : \mathbb{Z}_N^d \to \mathbb{Z}_N$ is the orthogonal projection to the i-th coordinate axis.

Proposition 2.1. *(Generalized von Neumann Inequality)*
Let $w = \otimes^d \nu$ be a pseudo-random d-measure. Given a function $0 \le f \le w$, we have that

$$\Lambda f := \mathbb{E}\left(f(x)f(x+te_1)...f(x+te_d) : \, x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N\right) = O(||f||_{\square(e')^d}) \tag{2.1}$$

where $e' = \{e_d, e_d - e_1, ..., e_d - e_{d-1}\}$.

Proof. We shall apply the Cauchy-Schwarz inequality several times. Begin by writing

$$\Lambda f \equiv \Lambda = \mathbb{E}(f(x)\prod_{i=1}^{d} f(x+t_1e_i) : x \in \mathbb{Z}_N^d, t_1 \in \mathbb{Z}_N).$$

Push through the summation on t_1 and split the f to write this as

$$\mathbb{E}(\sqrt{f(x)}\,\mathbb{E}(\sqrt{f(x)}\prod_{i=1}^{d} f(x+t_1e_i) : t_1 \in \mathbb{Z}_N) : x \in \mathbb{Z}_N^d).$$

Applying Cauchy-Schwarz to get

$$\Lambda^2 \le \mathbb{E}(w(x)\prod_{i=1}^{d} f(x+t_1e_i)\prod_{j=1}^{d} f(x+t_1e_j+t_2e_j) : t_1, t_2 \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have made the substitution $t_2 \mapsto t_1 + t_2$ for the new variable. Note that there should be a $\mathbb{E}(w(x)) = 1 + o(1)$ multiplier, following from the fact that $f \le w$ and from the linear forms condition, but for convenience we suppress it and will continue to do so (this is a big O result, so this is not of any consequence). We make one further substitution, $x \mapsto x - t_1e_1$, yielding

$$\Lambda^2 \le \mathbb{E}(w(x-t_1e_1)\prod_{i=2}^{d}\prod_{\omega\in\{0,1\}} f(x+t_1e_i^{(1)}+\omega t^{(1)}e_i)\prod_{\omega'\in\{0,1\}} f(x+\omega' t^{(1)}e_1) : t_1, t_2 \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have introduced the notations $e_i^{(j)} = e_i - e_j$, and $t^{(i)} = \{t_{1+j}\}_{j=1}^{i}$. Note that the final product of this expression is independent of t_1.

We now repeat this procedure exactly, pushing through the t_1 sum and splitting the terms independent of t_1, followed by a change of variables. After l applications of Cauchy-Schwarz inequality, we claim to have

$$\Lambda^{2^l} \le \mathbb{E}(W_l(x, t_1, ..., t_{l+1})\prod_{i=l+1}^{d}\prod_{\omega\in\{0,1\}^l} f(x+t_1e_i^{(l)}+\omega t^{(l)}e_{i;l}))$$
$$\times \prod_{\omega'\in\{0,1\}^l} f(x+\omega' t^{(l)}e_{l;l-1}) : t_1, ..., t_{l+1} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d), \tag{2.2}$$

for an appropriate weight function W_l which is a product of w's, evaluated on linear forms which are pairwise linearly independent.

The notations introduced here are $e_{i;l} = \{e_i, e_i^{(1)}, ..., e_i^{(l-1)}\}$ (note that $l > 1$), and $\omega t^{(l)}e_{i;l} = \omega_1t_2e_i + \omega_2t_3e_i^{(1)} + ... + \omega_lt_{l+1}e_i^{(l-1)}$.

To check this form, using induction, apply the Cauchy-Schwarz inequality one more time with the new variable $t_1 + t_{l+2}$ to get

$$\Lambda^{2^{l+1}} \le \mathbb{E}(W_l(x, t_1, ..., t_{l+1})W_l(x, t_1+t_{l+2}, ..., t_{l+1})$$
$$\times \prod_{i=l+1}^{d}\prod_{\omega\in\{0,1\}^l} f(x+t_1e_i^{(l)}+\omega t^{(l)}e_{i;l})f(x+t_1e_i^{(l)}+t_{l+2}e_i^{(l)}+\omega t^{(l)}e_{i;l})$$

$$\times \prod_{\omega'\in\{0,1\}^l} w(x+\omega' t^{(l)} e_{l;l-1}) : t_1,...,t_{l+2} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

Write

$$W'_{l+1}(x,t_1,...,t_{l+2}) = W_l(x,t_1,...,t_{l+1}) W_l(x,t_1+t_{l+2},...,t_{l+1}) \prod_{\omega'\in\{0,1\}^l} w(x+\omega' t^{(l)} e_{l;l-1}). \tag{2.3}$$

We now apply the substitution $x \mapsto x - t_1 e_{l+1}^{(l)}$, note that $e_i^{(l)} - e_{l+1}^{(l)} = e_i^{(l+1)}$, and set

$$W_{l+1}(x,t_1,...,t_{l+2}) = W'_{l+1}(x - t_1 e_{l+1}^{(l)}, t_1,...,t_{l+2}), \tag{2.4}$$

This gives

$$\Lambda^{2^{l+1}} \le \mathbb{E}(W_{l+1}(x,t_1,...,t_{l+2}) \times \prod_{i=l+2}^{d} \prod_{\omega\in\{0,1\}^{l+1}} f(x+t_1 e_i^{(l+1)} + \omega t^{(l+1)} e_{i;l+1})$$

$$\times \prod_{\omega'\in\{0,1\}^{l+1}} f(x+\omega' t^{(l+1)} e_{l+1;l}) : t_1,...,t_{l+2} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

and this is the form we wanted to obtain.

After $d-1$ iterations, one arrives at the form

$$\Lambda^{2^{d-1}} \le \mathbb{E}(W_{d-1}(x,t_1,...,t_d) \prod_{\omega'\in\{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_1,,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

This may be written as

$$\Lambda^{2^{d-1}} \le \mathbb{E}(\prod_{\omega'\in\{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_2,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d) + E,$$

where

$$E = \mathbb{E}((W_{d-1}(x,t_1,...,t_d) - 1) \prod_{\omega'\in\{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_1,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

To see that the main term is in fact an appropriate box norm, notice that

$$e_{d;d-1} = \{e_d, e_d - e_1, ..., e_d - e_{d-1}\}$$

is also in general position.

To deal with the error term E, we apply the Cauchy-Schwarz inequality one more time to get

$$E \le \mathbb{E}((W(x,t_2,...,t_d) - 1)^2 \prod_{\omega'\in\{0,1\}^d} w(x+\omega' t^{(d)} e_{d;d-1}) : t_2,,...,t_{d+1} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have set

$$W(x,t_2,...,t_d) = \mathbb{E}(W_{d-1}(x,t_1,t_2,...,t_d) : t_1 \in \mathbb{Z}_N)$$

and again used the fact that $f \le w$. Now to show that $E = o(1)$, it is enough to show that the linear forms defining W are pairwise independent, after of course expanding $(W-1)^2$ and applying the linear forms condition. By following the construction of W, this amounts to showing that at each step W_l satisfies pairwise independence, which itself reduces to showing that the coefficient of x is 1 in each form and each form has a nonzero coefficient in t_1 (in each coordinate).

To be more precise, the case $l = 1$ is immediate. Assuming this is so for l fixed, then

$$W'_{l+1}(x,t_1,...,t_{l+2}) = W_l(x,t_1,...,t_{l+1}) W_l(x,t_1+t_{l+2},...,t_{l+1}) \prod_{\omega'\in\{0,1\}^l} w(x+\omega' t^{(l)} e_{l;l-1}).$$

certainly satisfies this, as the the forms in $W_l(x, t_1, ..., t_{l+1})$ and $W_l(x, t_1 + t_{l+2}, ..., t_{l+1})$ are pairwise independent because the t_1 coefficient is non-zero, and $\prod_{\omega' \in \{0,1\}^l} w(x + \omega' t^{(l)} e_{l;l-1})$ is independent of t_1. The statement about the coefficient of x is obvious. Also, it not hard to see that the vector multiple of t_1 is either e_{l+1} or $e^{(i)}_{l+1}$ (for forms appearing after i applications of Cauchy-Schwarz). Thus the statement is true for $l + 1$.

The fact that $E = o(1)$ then follows directly from the $(d(d + 2)2^{d-1}, d(2d + 1))$ linear forms condition. □

3. The dual function estimate.

As before we assume that a basis $e = \{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$ is given which is in general position. We will use the notation $\omega y e = \omega_1 y_1 e_1 + ... + \omega_d y_d e_d$, for $\omega \in \{0, 1\}^d$ and $y \in \mathbb{Z}_N^d$. First we define the dual of a function $f : \mathbb{Z}_N^d \to \mathbb{R}$ with respect to the $\| \, \|_{\square(e)^d}$ norm.

Definition 3.1. . *Let $f : \mathbb{Z}_N^d \to \mathbb{R}$ be a given function and let $e = \{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$ be a basis of $\mathbb{Z}_N^d$. The dual of the the function f is the function*

$$\mathcal{D}f(x) = \mathbb{E}\,(\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} f(x + \omega t e) : t \in \mathbb{Z}_N^d) \tag{3.1}$$

Proposition 3.1. *With X and $\mathcal{D}$ as above, and e in general position, we have*

$$||\mathcal{D}f_1 ... \mathcal{D}f_K||^*_{\square(e)^d} \leq C(K)$$

for any $f_1, ..., f_K \in X$.

Proof. We must show that

$$\langle f, \mathcal{D}f_1 ... \mathcal{D}f_K \rangle \leq C_K ||f||_{\square(e)^d}$$

by the definition of the dual norm. By applying the definition of $\mathcal{D}f$, the LHS gives

$$\langle f, \mathcal{D}f_1 ... \mathcal{D}f_K \rangle = \mathbb{E}(f(x) \prod_{i=1}^{K} \mathbb{E}(\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} f_i(x + \omega t^i e) : t^i \in \mathbb{Z}_N^d) : x \in \mathbb{Z}_N^d).$$

Expanding out the products then gives the RHS as

$$\mathbb{E}(\mathbb{E}(f(x) \prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} \prod_{i=1}^{K} f_i(x + \omega t^i e + \omega t e) : x, t \in \mathbb{Z}_N^d) : T = (t^1, ..., t^K) \in (\mathbb{Z}_N^d)^K)$$

after a substitution $t^i \mapsto t + t^i$ for each i for some fixed t, and adding a redundant summation in t. Now we call $F_{(\omega,T)}(x) = \prod_{i=1}^{K} f_i(x + \omega t^i e)$ for non-zero ω, and $F_{(0^d,T)}(x) = f(x)$. The last expression then becomes

$$\mathbb{E}(\langle F_{(\omega,T)} : \omega \in \{0, 1\}^d \rangle : T \in \mathbb{Z}_N^d).$$

By applying the $\square(e)$-Cauchy-Schwarz inequality, we have arrived at

$$||\mathcal{D}f_1 ... \mathcal{D}f_K||^*_{\square(e)^d} \leq \mathbb{E}(\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0^d} ||F_{(\omega,T)}||_{\square(e)^d} : T \in \mathbb{Z}_N^d).$$

An application of the Holder inequality gives that the RHS is bounded above by

$$\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0^d} \mathbb{E}(||F_{(\omega,T)}||^{2^d}_{\square(e)^d} : T \in (\mathbb{Z}_N^d)^K),$$

where we added one factor of the constant 1 function, which has L^q-norm one for each q. Thus, we now just need to show that for a fixed $\omega \neq 0^d$ we have

$$\mathbb{E}(||F_{(\omega,T)}||^{2^d}_{\square(e)^d} : T \in (\mathbb{Z}_N^d)^K) = O(K)$$

for $T = (t^1, ..., t^K)$.

We continue by expanding the last expression for a fixed $\omega \neq 0^d$,

$$||F_{(\omega,T)}||^{2^d}_{\square(e)^d} : T \in (\mathbb{Z}_N^d)^K) = O(K) = \mathbb{E}(\prod_{\omega' \in \{0,1\}^d} \prod_{i=1}^{K} f_i(x+\omega t^i e+\omega' te) : x, t, t^1, ..., t^K \in \mathbb{Z}_N^d).$$

The RHS factorizes as

$$\mathbb{E}(\prod_{i=1}^{K} \mathbb{E}(\prod_{\omega' \in \{0,1\}^d} f_i(x + \omega ye + \omega' te) : y \in \mathbb{Z}_N^d) : x, t \in \mathbb{Z}_N^d).$$

Applying the bound $f \leq \nu$ gives

$$\mathbb{E}(\mathbb{E}^K(\prod_{\omega' \in \{0,1\}^d} \nu(x + \omega ye + \omega' te) : y \in \mathbb{Z}_N^d) : x, t \in \mathbb{Z}_N^d).$$

The inner sum is now split component wise

$$\mathbb{E}(\prod_{j=1}^{d} \prod_{\omega' \in \{0,1\}^d} \mu((\omega ye)_j + (\omega' te + x)_j) : y \in \mathbb{Z}_N^d),$$

where the notation $(x)_j$ denotes the j^{th} coordinate. The terms $(\omega ye)_j$ represent the linear forms $\sum_{s=1}^d \omega_s y_s (e_s)_j$, which satisfy the hypothesis in the $(d, 2^d)$ correlation condition by the assumptions on e. Hence we have

$$\mathbb{E}(\prod_{j=1}^{d} \prod_{\omega' \in \{0,1\}^d} \mu((\omega ye)_j + (\omega' te + x)_j) : y \in \mathbb{Z}_N^d) \leq \prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \tau(((\omega' - \omega'')te)_j),$$

as the $(x)_j$ terms drop out in the subtraction.

Plugging this bound back in gives

$$\mathbb{E}((\prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \tau(((\omega' - \omega'')te)_j))^K : t \in \mathbb{Z}_N^d).$$

Making use of the triangle inequality in $\mathcal{L}^{dK}$, after another application of Holder, reduces our task to bounding

$$\prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \mathbb{E}(\tau^{dK}(((\omega' - \omega'')te)_j) : t \in \mathbb{Z}_N^d).$$

By the assumptions on e and the fact that $\omega' - \omega'' \neq 0^d$, $((\omega' - \omega'')te)_j$ provides a uniform cover of $\mathbb{Z}_N$, and we may reduce this to

$$\mathbb{E}(\tau^{dK}(t) : t \in \mathbb{Z}_N).$$

This expression is $O_K(1)$. $\square$

4. Proof of the main results.

In this section we prove our main results under the assumption that the measure exhibited in [4] is d-pseudo-random, i.e. it satisfies Definition 1.5.

4.1. **Proof of Theorem 1.2.** Let $e = \{e_1, \dots, e_d\} \subseteq \mathbb{Z}_N^d$ be a basis which is in general position. For a function $f : \mathbb{Z}_N^d \to \mathbb{R}$ we define its dual by

$$\mathcal{D}f(x) = \mathbb{E}(\prod_{\omega \in \{0,1\}^d, \omega \neq 0} f(x + \omega t e) : t \in \mathbb{Z}_N^d). \tag{4.1}$$

Then clearly

$$\langle f, Df \rangle = \|f\|_{\square(e)^d}^{2^d} \tag{4.2}$$

Let $\mu = \otimes^d \nu$ be a pseudo-random d-measure, and let X be the set of functions f on $\mathbb{Z}_N^d$ such that $|f| \leq \mu$ pointwise.

Lemma 4.1. *The norm $\| \ \|_{\square(e)^d}$ is a quasi algebra predual (QAP) norm, with respect to the set X and the operator D.*

Proof. We have already shown part (iii) of Definition 1.2, which was the content of Proposition 3.1. If $\|f\|_{\square(e)^d}^d \leq \varepsilon$ then

$$\langle f, Df \rangle = \|f\|_{\square(e)^d}^{2^d} \leq \varepsilon^{2^d}$$

and part (ii) follows. Finally, since $|f| \leq \mu$ it follows

$$\langle f, Df \rangle \leq \|\mu\|_{\square(e)^d}^{2^d} = 1 + o(1)$$

as the linear forms $(x + \omega t e)_j$ are pairwise linearly independent (for each j) and ν satisfies the linear forms condition. □

We are in the position to apply the transference principle to decompose a function $0 \leq f \leq \mu$ into the sum of a bounded function g and a function h which has small contribution to the expression in (1.2).

Proof of Theorem 1.2. Let $\alpha > 0$ and let $0 \leq f \leq \mu$ be function such that $\mathbb{E}f \geq \alpha$, where μ is a pseudo-random d-measure on $\mathbb{Z}_N^d$. We apply Theorem B, with $Y = \mathbb{Z}_N^d$, $\delta = 1/2$ and $\eta > 0$. Note that since μ is a measure one has that $\|\mu\|_{L^1} = \mathbb{E}\mu = 1 + o(1)$. Since $\| \ \|_{\square(e)^d}$ is a QAP norm with respect to the set $X = \{f : Y \to \mathbb{R},\ |f| \leq \mu\}$, it follows that there is an $\varepsilon > 0$ such that if

$$\|\mu - 1\|_{\square(e)^d} < \varepsilon \tag{4.3}$$

then there is a decomposition $f = g + h$ such that

$$0 \leq g \leq 2 \qquad \text{and} \qquad \|h\|_{\square(e)^d} < \eta. \tag{4.4}$$

Since μ is pseudo-random $\|\mu - 1\|_{\square(e)^d} = o(1)$ thus (4.3) holds for large enough N. Using this decomposition together with Theorem A and Proposition 2.1 one may write

$$\mathbb{E}(f(x)f(x + te_1)...f(x + te_d) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) =$$

$$= \mathbb{E}(g(x)g(x+te_1)...g(x+te_d) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) + O(\|h\|_{\square(e)^d}) \geq c'(\alpha, e) - C_d\eta \geq c'(\alpha, e)/2$$

by choosing η sufficiently small with respect to α and e. This proves Theorem 1.2. □

4.2. **Proof of Theorem 1.1.** Let us identify $[1,N]^d$ with $\mathbb{Z}_N^d$. First we show that constellations in $\mathbb{Z}_N^d$ defined by e which are contained in a box $B \subseteq [1,N]^d$ of size εN, are in fact genuine constellations contained in B. We say that $e = \{e_1, \dots, e_d\} \in \mathbb{Z}^{d^2}$ is *primitive* if the segment $[0,e]$ does not contain any other lattice points other than its endpoints in $\mathbb{Z}^{d^2}$ considered as a lattice point in $\mathbb{Z}^{d^2}$. Let us also define the positive quantity $\tau(e)$ by

$$\tau(e) = \inf_{m \notin \{0,e\},\, x \in [0,e]} |m - x|_\infty \quad \text{where} \quad |x|_\infty = \max_{1 \le j \le d^2} |x_j|$$

m is running through the lattice points $\mathbb{Z}^{d^2}$ other than 0 and e.

Lemma 4.2. *Let $0 < \varepsilon < \tau(e)$. Let N be sufficiently large, and let $B = I^d$ be a box of size εN contained in $[1,N]^d \simeq \mathbb{Z}_N^d$. If there exist $x \in \mathbb{Z}_N^d$ and $t \in \mathbb{Z}_N \backslash \{0\}$ such that $x \in B$ and $x + te \subseteq B$ as a subset on $\mathbb{Z}_N^d$, then there exists a scalar $t' \neq 0$ such that $x + t'e \subseteq B$ also as a subset of $\mathbb{Z}^d$. Moreover if e is primitive (and $1 \le t < N$) then one may take $t' = t$ or $t' = t - N$.*

Proof. First, note that one can assume e is primitive as $x + te = x + tse'$ for a fixed primitive e' and $s \in \mathbb{N}$. By our assumption, there is an $x \in [1,N]^d$ and $t \in [1, N-1]$ such that $x \in B$ and $x + te_j \in B + (N\mathbb{Z})^d$ for all $1 \le j \le d$. Thus for each j there exits $m_j \in \mathbb{Z}^d$ such that $|te_j - Nm_j|_\infty \le \varepsilon N$ and hence $|\lambda e - m|_\infty \le \varepsilon$, where $m = \{m_1, \dots, m_d\} \in \mathbb{Z}^{d^2}$ and $\lambda = t/N$. Since $0 < \lambda < 1$ and $\varepsilon < \tau(e)$ this implies that $m = 0$ or $m = e$. If $m = 0$ then $|te|_\infty \le \varepsilon N$ and since $x \in B$ it follows that $x + te \subseteq B \subseteq \mathbb{Z}^d$. If $m = e$ then $|(t-N)e_j|_\infty \le \varepsilon N$ thus $x + (t-N)e \subseteq B \subseteq \mathbb{Z}^d$, so $x + t'e \subseteq B$ as a subset of $\mathbb{Z}^d$. This proves the lemma. $\square$

Let us briefly recall the pseudo-random measure ν defined in Sec.9 [4]. Let $w = w(N)$ be a sufficiently slowly growing function (choosing $w(N) \ll \log\log N$ is sufficient as in [4]) and let $W = \prod_{p \le w} p$ be the product of primes up to w. For given b relative prime to W define the modified von Mangoldt function $\bar{\Lambda}_b$ by

$$\bar{\Lambda}_b(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn + b) & \text{if } Wn + b \text{ is a prime;} \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

where ϕ is the Euler function. Note that by Dirichlet's theorem on the distribution of primes in residue classes one has that $\sum_{n \le N} \bar{\Lambda}_b(n) = N(1 + o(1))$. Also, if $A \subseteq \mathbb{P}^d$ is of positive relative α and if $\bar{\Lambda}_b^d := \otimes^d \bar{\Lambda}_b$ is the d-fold tensor product of $\bar{\Lambda}_b$ the it is easy to see that there exists a b such that

$$\limsup_{N \to \infty} N^{-d} \sum_{x \in [1,N]^d} \mathbf{1}_A(x) \bar{\Lambda}_b^d(x) \; > \; \alpha/2 \tag{4.6}$$

We will fix such b and choose N sufficiently large N for which the expression in (4.6) is at least $\alpha/2$. Let $R = N^{d^{-1}2^{-d-5}}$ and recall the Goldston-Yildirim divisor sum [4], [5]

$$\Lambda_R(n) = \sum_{d | n, d \le R} \mu(d) \log(R/d)$$

μ being the Mobius function. For given small parameters $0 < \varepsilon_1 < \varepsilon_2 < 1$ (whose exact values will be specified later) recall the Green-Tao measure

$$\nu(n) = \begin{cases} \frac{\phi(W)}{W} \frac{\Lambda_R(Wn+b)^2}{\log R} & \text{if } \varepsilon_1 N \le n \le \varepsilon_2 N; \\ 1 & \text{otherwise.} \end{cases} \tag{4.7}$$

Note that $\nu(n) \ge 0$ for all n, and also it is easy to see that for N sufficiently large, one has that

$$\nu(n) \ge d^{-1} 2^{-d-6} \, \bar{\Lambda}_b(n) \tag{4.8}$$

for all $\varepsilon_1 N \le n \le \varepsilon_2 N$. Indeed, this is trivial unless $Wn + b$ is a prime. In that case, since $\varepsilon_1 N > R$, $\Lambda_R(Wn + b) = \log R = d^{-1} 2^{-d-5} \log N$ and (4.8) follows.

Proof of Theorem 1.1. Set $\mu = \otimes^d \nu$, and let

$$g(x) := c_d\, \bar{\Lambda}_b^d(x)\, \mathbf{1}_A(x)\, \mathbf{1}_{[\varepsilon_1 N, \varepsilon_2 N]^d}(x) \qquad (c_d = d^{-d} 2^{-d^2-6d}) \tag{4.9}$$

Then by (4.8) one has that $g(x) \leq \mu(x)$ for all $x \in \mathbb{Z}_+^d$. By (4.6) one may choose a sufficiently large number N' for which

$$(N')^{-d} \sum_{x \in [1,N']^d} \mathbf{1}_A(x) \bar{\Lambda}_b^d(x) \;>\; \alpha/2 \tag{4.10}$$

and a prime N such that

$$(1 - \frac{\alpha}{100d}) N' \leq \varepsilon_2 N \leq N'$$

If ε_1 is such that $\varepsilon_1/\varepsilon_2 \leq \alpha/100d$, then by the Prime Number Theorem in arithmetic progressions

$$(N')^{-d} \sum_{x \in [1,N']^d \setminus [\varepsilon_1 N, \varepsilon_2 N]^d} \bar{\Lambda}_b^d(x) \;\leq\; \alpha/10 \tag{4.11}$$

It follows from (4.10) and (4.11)

$$N^{-d} \sum_{x \in [1,N']^d} g(x) \;\geq\; c_d\, N^{-d} \sum_{x \in [\varepsilon_1 N, \varepsilon_2 N]^d} \mathbf{1}_A(x) \bar{\Lambda}_b^d(x) \;\geq\; c_d \varepsilon_2^d \alpha/4 \tag{4.12}$$

Using the identification $[1,N]^d \simeq \mathbb{Z}_N^d$, one has that $\mathbb{E}(g(x) :\, x \in \mathbb{Z}_N^d) \geq \alpha'$ (with $\alpha' = c_d^d \varepsilon_2^d \alpha/4$), and $0 \leq g(x) \leq \mu(x)$ for all x. Thus, save for proving that the measure ν is d-pseudo-random, Theorem 1.2 implies that

$$\mathbb{E}(g(x) g(x+te_1) \ldots g(x+te_d) :\; x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) \;\geq\; c'(\alpha, e) > 0.$$

Note that the contribution of trivial constellations, corresponding to $t = 0$, is at most $O(N^{-1} \log^d N)$, as $|\bar{\Lambda}_b^d| \leq \log^d N$ uniformly on $[1,N]^d$. Since the support of g is contained in $A \cap [\varepsilon_1 N, \varepsilon_2 N]^d$, Lemma 4.2 implies that $A \cap [\varepsilon_1 N, \varepsilon_2 N]^d$ must contain genuine constellations of the form $\{x, x+te_1, \ldots, x+te_d\}$ as a subset of $\mathbb{Z}^d$. Choosing an infinite sequence of N's it follows that A contains infinitely many constellations defined by e. $\square$

5. Appendix: The correlation condition.

To complete the proof of Theorem 1.1, one needs to show that the measure ν defined in (4.7) satisfies both the linear forms conditions and the $(d, 2^d)$ correlation conditions given in (1.4). Since the measure ν is the same (apart from the slight change in the interval where $\nu \equiv 1$) is the one given in [4] (see Definition 9.3, there), the linear forms condition is already established in Prop. 9.8 in [4]. It turns out that the arguments given in [4] (see Prop. 9.6, Lemma 9.9 and Prop.9.10) generalize in a straightforward manner to obtain the more general (m_0, m_1) correlation condition for any given specific values of m_0 and m_1.

Proposition 5.1. *For a fixed m_0, m_1, there exists a function τ such that*

$$\mathbb{E}\tau^k = O_k(1)$$

and also

$$\mathbb{E}(\prod_{i=1}^{m_1} \prod_{j=1}^{m_0} \nu(\phi_i(y) + h_{i,j}) : y \in \mathbb{Z}_N^r) \leq \prod_{i=1}^{m_0} (\sum_{1 \leq j < j' \leq m_0} \tau(h_{i,j} - h_{i,j'})) \tag{5.1}$$

where the $\phi_i : \mathbb{Z}_N^r \to \mathbb{Z}_N$ are pairwise linearly independent linear forms.

Let us first note that the arguments of Lemma 9.9 and Prop. 9.10 of [4] applies to our case and it is enough to establish the following inequality (see Prop. 9.6 [4])

$$\mathbb{E}\,(\prod_{i=1}^{m_1}\prod_{j=1}^{m_0}\Lambda_R^2(W(\phi_i(y)+h_{i,j})+b):\ y\in B)$$
$$\leq\ C_M\left(\frac{W\log R}{\phi(W)}\right)^M\prod_{i=1}^{m_1}\prod_{p|\triangle_i}(1+O_M(p^{-1/2})) \tag{5.2}$$

where $M=m_1m_0$ and B is a box of size at most R^{10M}. Moreover one can assume that $h_{i,j}\neq h_{i,j'}$ for all i, $j\neq j'$.

The next step is, following [4], to write the the expression

$$\mathbb{E}(\prod_{i=1}^{M}\Lambda_R^2(\theta_i(y)):y\in B),$$

where $\theta_i=W(\phi_{\lfloor i/m_1\rfloor}(y)+h_{\lfloor i/m_1\rfloor,\,(i\,(p)})+b$ ($\lfloor x\rfloor$ is the floor function, $i\,(m_1)$ is i modulo m_1), as a contour integral of the the following form plus a small error

$$(2\pi i)^{-M}\int_{\Gamma_1}\dots\int_{\Gamma_1}F(z,z')\prod_{j=1}^{M}\frac{R^{z_j+z'_j}}{z_j^2z_j'^2}dz_jdz_j', \tag{5.3}$$

where $z=(z_1,...,z_M)$, $z'=(z_1',...,z_M')$, and function $F(z,z')$ is taking form of an Euler product

$$F(z,z')=\prod_p E_p(z,,z'),$$

where

$$E_p(z,z')=\sum_{X,X'\subseteq[M]}\frac{(-1)^{|X|+|X'|}\omega_{X\bigcup X'}(p)}{p^{\sum_{j\in X}z_j+\sum_{j\in X'}z_j'}}.$$

The function ω relates this expression to the particular forms. Specifically

$$\omega_X(p)=\mathbb{E}(\prod_{i\in X}\mathbf{1}_{\theta_i\equiv 0\,(p)}:x\in\mathbb{Z}_N^r).$$

Lemma 5.1. *(Local factor estimate). Set the intervals $I_i=[(i-1)m_1+1,im_1]$ as a partition of $[M]$. For $\alpha\in I_i$, the homogeneous part of θ_α is $W\phi_i$. Also, set $\Delta_i=\prod_{j<j;\,j,j'\in I_i}|h_{i,j}-h_{i,j'}|$.The following estimates hold: $\omega_X(p)$:*

1. *If $p\leq w(N)$, then $\omega_X(p)=1$ if $|X|=0$, and is 0 otherwise.*
2. *If $p>w(N)$ and $|X|=0$, then $w_X(p)=1$.*
3. *If $p>w(N)$ and $X\subseteq I_i$ is nonempty, we have $w_X(p)=p^{-1}$ when $|X|=1$, and $w_X(p)\leq p^{-1}$ when $|X|>1$. In the latter case, if $p\nmid\Delta_\alpha$, we have that $\omega_X(p)=0$.*
4. *If $p>w(N)$ and $X\cap I_i\neq\emptyset$ and $X\cap I_{i'}\neq\emptyset$ for some $i\neq i'$, we have $\omega_X(p)\leq p^{-2}$.*

Proof. When $p\leq w(N)$, then $W\phi_i+b\equiv b\,(p)$, giving the first result. The second statement is trivial.

For the third statement, let us start with $X\subseteq I_i$ with $|X|=1$. Then we have

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)}:y\in\mathbb{Z}_N^r)=p^{-1}$$

for any fixed j, proving the first part. The second part requires an estimate of

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)}\mathbf{1}_{W(\phi_i(y)+h_{i,\,j'})+b\equiv 0\,(p)}:y\in\mathbb{Z}_N^r),$$

with $j\neq j'$. If $p|\,|h_{\alpha,\,j}-h_{\alpha,\,j'}|$, then the we are left with simply a single equation ($p\nmid W$), and may refer to the first part. When $p\nmid\Delta_\alpha$, $\omega_X(p)=0$ as $h_{i,\,j}$ is not congruent to $h_{i,\,j'}$, modulo p.

For the last statement, we have the upper bound

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)}\mathbf{1}_{W(\phi_i'(y)+h_{i',\,j'})+b\equiv 0\,(p)} : y\in\mathbb{Z}_N^r)$$

for some $i\neq i'$ and j,j'. The forms ϕ_i and $\phi_{i'}$ are linearly independent modulo p (see the proof of Lemma 10.1 in [4]), hence we have the intersection of two distinct linear algebraic sets, which has size at most p^{r-2}. □

The terms E_p in the Euler product can be separated as

$$E_p(z,z') = 1-\mathbf{1}_{p>w(N)}\sum_{j=1}^{M}(p^{-1-z_j}+p^{-1-z_j'}-p^{-1-z_j-z_j'})$$

$$+\sum_{i=1}^{m_1}\mathbf{1}_{p>w(N);\,p|\Delta_i}\lambda_p^{(i)}(z,z')+\sum_{X\bigcup X'\not\subset I_\alpha,\,\alpha\in[m_1];\,|X\bigcup X'|>1}\frac{O_M(p^{-2})}{p^{\sum_X z_j+\sum_{X'}z_j'}},$$

where

$$\lambda_p^{(i)}(z,z') = \sum_{X\bigcup X'\subset I_i;\,|X\bigcup X'|>1}\frac{O_M(p^{-1})}{p^{\sum_X z_j+\sum_{X'}z_j'}}.$$

We define the terms

$$E_p^{(0)} = 1+\sum_{i=1}^{m_1}\mathbf{1}_{p>w(N);\,p|\Delta_i}\lambda_p^{(i)}(z,z'),$$

and factorize $E_p = E_p^{(0)}E_p^{(1)}E_p^{(2)}E_p^{(3)}$ as follows:

$$E_p^{(1)} = \frac{E_p}{E_p^{(0)}\prod_{j=1}^M(1-\mathbf{1}_{p>w(N)}p^{-1-z_j})(1-\mathbf{1}_{p>w(N)}p^{-1-z_j'})(1-\mathbf{1}_{p>w(N)}p^{-1-z_j-z_j'})^{-1}}$$

$$E_p^{(2)} = \prod_{j=1}^{M}(1-\mathbf{1}_{p\le w(N)}p^{-1-z_j})^{-1}(1-\mathbf{1}_{p\le w(N)}p^{-1-z_j'})^{-1}(1-\mathbf{1}_{p\le w(N)}p^{-1-z_j-z_j'})$$

$$E_p^{(3)} = \prod_{j=1}^{M}(1-p^{-1-z_j})(1-p^{-1-z_j'})(1-p^{-1-z_j-z_j'})^{-1},$$

and set $G_i=\prod_p E_p^{(i)}$, noting that

$$G_3 = \prod_{j=1}^{M}\frac{\zeta(1+z_j+z_j')}{\zeta(1+z_j)\zeta(1+z_j')}.$$

The the following is the analogue of lemma 10.6 in [4]. To state it, Let us recall the domain $\mathcal{D}_\sigma^M$ to be the set

$$\{z_j,z_j' : \Re z_j,\Re z_j'\in(-\sigma,100)\,,\,1\le j\le M\}.$$

We also have the norms on for f analytic on $\mathcal{D}_\sigma^M$, denoted $||f||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)}$, given by

$$||f||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)} = \sup||(\frac{\partial}{\partial z_1})^{\alpha_1}...(\frac{\partial}{\partial z_M})^{\alpha_1}(\frac{\partial}{\partial z_1'})^{\alpha_1}...(\frac{\partial}{\partial z_M'})^{\alpha_1}f||_{\mathcal{L}^\infty(\mathcal{D}_\sigma^M)},$$

where the supremum is taken over all $\alpha_1,...,\alpha_M,\alpha_1',...,\alpha_M'$ whose sum is at most k.

Lemma 5.2. *Let $0 < \sigma = 1/(6M)$. Then the Euler products G_i are absolutely convergent for $i = 0, 1, 2$ in the domain $\mathcal{D}_\sigma^M$, and hence represent analytic functions on this domain. We also have the estimates*

$$||G_0||_{\mathcal{C}^r(\mathcal{D}_\sigma^M)} = O_M(\log(R)/\log\log(R))^r \prod_{p|\prod_{i=1}^{m_1}\Delta_i} (1 + O_M(p^{2M\sigma-1}))$$

$$||G_0||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq \exp(O_M(\log^{1/3}(R)))$$

$$||G_1||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq O_M(1)$$

$$||G_2||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq O_{M,w(N)}(1)$$

$$G_0(0,0) = \prod_{i=1}^{m_1}\prod_{p|\Delta_i}(1 + O_M(p^{-1/2}))$$

$$G_1(0,0) = 1 + o_M(1)$$

$$G_2(0,0) = (W/\phi(W))^M,$$

where the first bound is for all $0 \leq r \leq M$.

Proof. The estimates proceed exatctly as in Lemma 10.3 and Lemma 10.6 in [4] with $\Delta = \prod_{i=1}^{m_1}\Delta_i$, barring the statement about $G_0(0,0)$. To see this, we have

$$G_0(0,0) = \prod_{p|\Delta} E_p^{(0)} = \prod_{p|\Delta}(1 + \sum_{i=1}^{m_1}\lambda_p^{(i)}(0,0)) \leq \prod_{i=1}^{m_1}\prod_{p|\Delta_i}(1 + |\lambda_p^{(i)}(0,0)|)$$

and we crudely have $|\lambda_p^{(i)}(0,0)| = 1 + O_M(p^{-1/2})$. □

The expression in (5.3) takes the form

$$(2\pi i)^{-M}\int_{\Gamma_1}\cdots\int_{\Gamma_1} G(z,z')\prod_{j=1}^{M}\frac{\zeta(1+z_j+z_j')R^{z_j+z_j'}}{\zeta(1+z_j)\zeta(1+z_j')z_j^2 z_j'^2}dz_j dz_j'$$

with $G = G_0G_1G_2$. To estimate it let us recall the following general result on contour integration from [4], see Lemma 10.4 there.

Lemma 5.3. *(Goldston-Yıldırım [4][6]) Let R be a positive number. If $G(z,z')$ is analytic in the $2M$ variables on $\mathcal{D}_\sigma^M$ for some $\sigma > 0$, and we have the estimate*

$$||G||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)} = \exp(O_{M,\sigma}(\log^{1/3}(R))),$$

then

$$(2\pi i)^{-M}\int_{\Gamma_1}\cdots\int_{\Gamma_1} G(z,z')\prod_{j=1}^{M}\frac{\zeta(1+z_j+z_j')R^{z_j+z_j'}}{\zeta(1+z_j)\zeta(1+z_j')z_j^2 z_j'^2}dz_j dz_j'$$

$$= G(0,...,0)\log^M(R) + \sum_{j=1}^{M} O_{M,\sigma}(||G||_{\mathcal{C}^j(\mathcal{D}_\sigma^M)})\log^{M-j}(R) + O_{M,\sigma}(\exp(-\delta\sqrt{\log(R)}))$$

for some $\delta > 0$.

Estimate (5.2) follows easily applying Lemma 5 (with $\sigma = 1/6M$) to $G = G_0G_1G_2$ using Lemma 4, which in turn implies Proposition 5.1, where the function τ is defined precisely as in [4]. This finishes the proof of Theorem 1.1.

References

[1] H. Furstenberg, Y. Katznelson, *An ergodic Szemerédi theorem for commuting trnasformations*, J. Analyse Math. 31 (1978), 275-291

[2] W.T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Annals of Math. 166/3 (2007), 897-946

[3] W.T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, Bull. London Math. Soc. 42 (4) (2010), 573-606

[4] B. Green and T. Tao, *The primes contain arbitrary long arithmetic progressions*, Annals of Math. 167 (2008), 481-547

[5] D. Goldston, C. Yildirim, *Higher correlations of divisor sums related to primes I: triple correlations*, Integers: Electronic Journal of Combinatorial Number theory, 3 (2003), 1-66

[6] D. Goldston, C. Yildirim, *Higher correlations of divisor sums related to primes III: small gaps between primes* , Proc. London Math. Soc. 95 (2007), 653-686

[7] O. Reingold, L. Trevisan, M. Tulsiani, S. Vadham, *Dense subsets of pseudorandom sets* Electronic Colloquium of Computational Complexity, Report TR08-045 (2008)

[8] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 299-345

[9] T. Tao, *The Gaussian primes contain arbitrarily shaped constellations*, J. Analyse Math., 99/1 (2006), 109-176

[10] T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Math., 201 (2008), 213-305

Department of Mathematics, The University of British Columbia, Vancouver, BC, V6T1Z2, Canada
E-mail address: bcook@math.ubc.ca

Department of Mathematics, University of British Columbia, Vancouver, B.C. V6T 1Z2, Canada
E-mail address: magyar@math.ubc.ca