MULTIDIMENSIONAL CONFIGURATIONS IN THE PRIMES WITH SHIFTED PRIME STEPS

A. N. LE[1,*] and T. H. LÊ[2,†]

[1]Department of Mathematics, The Ohio State University, 458 Math Tower,
231 West 18th Avenue, Columbus, OH 43210, USA
e-mail: le.286@osu.edu

[2]Department of Mathematics, University of Mississippi, University, MS 38677, USA
e-mail: leth@olemiss.edu

Abstract. Let $\mathcal{P}$ denote the set of primes. For a fixed dimension d, Cook–Magyar–Titichetrakun, Tao–Ziegler and Fox–Zhao independently proved that any subset of positive relative density of $\mathcal{P}^d$ contains an arbitrary linear configuration. In this paper, we prove that there exists such configuration with the step being a shifted prime (prime minus 1 or plus 1).

1. Introduction

1.1. History and statement of the result. For a set $A \subset \mathbb{Z}^d$, we define its upper density by $\overline{d}(A) = \limsup_{N \to \infty} \frac{|A \cap [1,N]^d|}{N^d}$. We say A is dense if $\overline{d}(A) > 0$. The celebrated Szemerédi theorem [16] states that any dense subset of the integers must contain arbitrarily long arithmetic progressions. After Furstenberg's [8] ergodic theoretic proof of this result appeared, many far-reaching generalizations have been obtained. For example, Furstenberg–Katznelson [9] proved a multidimensional generalization of Szemerédi's theorem, namely that for any vectors $h_1, \ldots, h_t \in \mathbb{Z}^d$, any dense subset of $\mathbb{Z}^d$ contains a configuration of the form $\{a + rh_1, \ldots, a + rh_t\}$ for some $a \in \mathbb{Z}^d$ and $r \in \mathbb{Z}$ nonzero. Bergelson–Leibman [2] generalized Furstenberg–Katznelson's result to polynomial configurations. That is to say: Given vectors $h_1, \ldots, h_t \in \mathbb{Z}^d$ and polynomials $P_1, \ldots, P_t \in \mathbb{Z}[x]$ without constant terms, any dense subset of $\mathbb{Z}^d$ must contain a configuration of the form $\{a + P_1(r)h_1, \ldots, a + P_t(r)h_t\}$ for some $a \in \mathbb{Z}^d$ and $r \in \mathbb{Z}$ nonzero.

In [6] and [7], Frantzikinakis–Host–Kra showed that the "step" r in the theorems of Furstenberg–Katznelson and Bergelson–Leibman can be taken to be a shifted prime, i.e. a number of the form $p - 1$ where p is a prime[1], using what we now term the comparison method. These results were also proved independently by Wooley–Ziegler [20] (in the polynomial case) and Bergelson–Leibman–Ziegler [3] (in the multidimensional case).

Let $\mathcal{P}$ denote the set of primes. For a set $A \subset \mathcal{P}^d$, we define its relative upper density by $\overline{d}_{\mathcal{P}^d}(A) = \limsup_{N\to\infty} \frac{|A \cap [1,N]^d|}{|\mathcal{P}^d \cap [1,N]^d|}$. We say A is dense in $\mathcal{P}^d$ if $\overline{d}_{\mathcal{P}^d}(A) > 0$. The Green–Tao theorem [10] states that any dense subset of $\mathcal{P}$ must contain an arbitrarily long arithmetic progression. Tao–Ziegler [17] generalized this to polynomial configurations, namely any dense subset of $\mathcal{P}$ must contain a configuration $\{a + P_1(r), \ldots, a + P_t(r)\}$ for some $a, r \in \mathbb{Z}, r \neq 0$, where $P_1, \ldots, P_t \in \mathbb{Z}[x]$ are given polynomials without constant terms. Regarding multidimensional configurations, Cook–Magyar–Titichetrakun [4] and Tao–Ziegler [18] independently proved a prime version of the Furstenberg–Katznelson theorem. That is to say for any $h_1, h_2, \ldots, h_k \in \mathbb{Z}^d$, any dense subset of $\mathcal{P}^d$ must contain a configuration $\{a + rh_1, a + rh_2, \ldots, a + rh_k\}$ for some $a \in \mathbb{Z}^d$ and $r \in \mathbb{Z}$ nonzero. Shortly after, Fox–Zhao [5] came up with a very short proof of the same result. Tao–Ziegler and Fox–Zhao's proofs rely on the program of counting linear configurations in the primes by Green, Tao and Ziegler [11,13,14]. In contrast, Cook–Magyar–Titichetrakun's [4] proof uses a hypergraph removal lemma for weighted hypergraphs.

In [15], Wolf and the second author "completed the square" by proving a hybrid of Tao–Ziegler [17] and Wooley–Ziegler's [20] results, namely that in polynomial configurations $\{a + P_1(r), \ldots, a + P_t(r)\}$ in dense subsets of the primes, one can require the step r to be of the form $p - 1$ (or $p + 1$), where p is prime. Their proof relies on the comparison method of Frantzikinakis–Host–Kra. In this paper, we complete another square by showing that the step r in the multidimensional result of Cook–Magyar–Titichetrakun, Tao–Ziegler and Fox–Zhao can also be taken to be of the form $p - 1$ or $p + 1$. To be precise, we will prove the following:

THEOREM 1.1. *Let $A \subseteq \mathcal{P}^d$ be of positive relative upper density. Then for every $h_1, \ldots, h_k \in \mathbb{Z}^d$, there exists $a \in \mathbb{Z}^d$ and $p \in \mathcal{P}$ such that $a + (p-1)h_1$, $\ldots$, $a + (p-1)h_k \in A$. The same is true if we replaced $p - 1$ by $p + 1$.*

In fact, we are able to find a lower bound for the number of such configurations. Again, simple examples show that $\mathcal{P} - 1$ and $\mathcal{P} + 1$ are the only translates of the primes having this property.

---

[1] The same result is true with $p + 1$ in place of $p - 1$. Simple counterexamples show that these are the only translates of the primes enjoying this property.

1.2. Proof strategy. In [15], the authors used Tao–Ziegler's theorem on polynomial configurations in the primes as a black box and employed Frantzikinakis–Host–Kra's comparison method. In the current situation, our implementation of the comparison method is less straightforward due to different kinds of averages being considered. Our proof relies on Tao–Ziegler's [18] weighted Furstenberg correspondence principle, which we now recall. With the setup as in Theorem 1.1, Tao and Ziegler constructed a measure preserving system $(X, \mathcal{B}, \mu, (T_h)_{h \in \mathbb{Z}^d})$ and a set $E \in \mathcal{B}$ such that for $m \in \mathbb{N}$, the expression

$$\mu\big(T_{mh_1}E \cap \cdots \cap T_{mh_k}E\big) \tag{1}$$

essentially captures the density of homothetic copies of $\{h_1, \ldots, h_k\}$ in A. The existence of m that makes (1) positive is then guaranteed by the Furstenberg–Katznelson theorem [9]. The whole idea is reminiscent of that of Furstenberg [8] when proving Szemerédi's theorem, in which he constructed a system and a set arising from a dense subset of $\mathbb{Z}$ (instead of $\mathcal{P}^d$).

To show the step of the configurations found in [18] can be restricted to shifted primes, one is tempted to use the Franzikinakis–Host–Kra's comparison method. However, this approach does not produce configurations as desired. In contrast to the measure built in [8], the measure in [18] contains an extra average on a parameter r (see Section 2.4). Hence with this approach, we end up with configurations of the form

$$a + mrh_1, \quad \ldots, \quad a + mrh_k \tag{2}$$

with $m = p - 1$ for some prime p and some $r \in \mathbb{Z}$ which we have no control of. The appearance of r arises from the use of Varnavides' trick [19].

Instead, to prove Theorem 1.1, we first use Furstenberg and Katznelson's theorem to fix an m that makes (1) positive. Only after that, the comparison method is implemented, taking advantage of the additional average. As a result, we can obtain configurations as (2) with $r = (p - 1)/m$ for some prime p. This produces the desired result.

REMARK 1.2. We remark that Fox–Zhao's proof [5] also uses Varnavides' trick. While their proof is very simple, the number of configurations it provides is too small for our purpose. Therefore, we have to resort to the methods of Tao–Ziegler [18], and this shows that the results in [18] are not completely superseded by [5].

The detailed proof will be presented in Section 3 contingent upon Proposition 3.1, which is proved in Section 4. In the next section, we set up the notation and necessary background.

2. Preliminaries

2.1. Notations. For $N \in \mathbb{N}$, let $[N]$ denote the set $\{1, 2, \ldots, N\}$. For a finite set S and a function $f\colon S \to \mathbb{R}$, let $\mathbb{E}_{n \in S} f(n)$ denote the average $\frac{1}{|S|} \sum_{n \in S} f(n)$.

For two functions $f, g\colon \mathbb{N} \to \mathbb{R}$, by writing $f(n) \ll g(n)$, we mean there is a constant $c > 0$ such that $|f(n)| \leq cg(n)$ for all $n \in \mathbb{N}$.

Let $o_{n \to \infty}(1)$ or $o(1)$ denote a function that goes to 0 as n tends to infinity.

Let ω_n be a function growing slowly with n (to be determined later). Then we define

$$W_n = \prod_{p \leq \omega_n, p \in \mathcal{P}} p.$$

For $-W_n < b < W_n$ coprime to W_n, define the W-tricked von Mangoldt function as

$$\Lambda_{W_n, b}(m) = \frac{\phi(W_n)}{W_n} \Lambda'(W_n m + b)$$

where Λ' is the modified von Mangoldt function

$$\Lambda'(m) = \begin{cases} \log m & \text{if } m \in \mathcal{P} \\ 0 & \text{otherwise.} \end{cases}$$

2.2. Furstenberg $\mathbb{Z}^d$-system. Let $X = 2^{\mathbb{Z}^d} = \{B : B \subseteq \mathbb{Z}^d\}$. Let $\mathcal{B}$ be the σ-algebra on X which is generated by basic cylinder sets $\{B \in X : b \in B\}$ for some $b \in \mathbb{Z}^d$. This space has an $\mathbb{Z}^d$-action $(T_h)_{h \in \mathbb{Z}^d}$ defined by $T_h B = B + h$ for all $h \in \mathbb{Z}^d$ and $B \in X$. Then $(X, \mathcal{B}, (T_h)_{h \in \mathbb{Z}^d})$ is a topological $\mathbb{Z}^d$-system, which we call the Furstenberg $\mathbb{Z}^d$-system.

2.3. The Furstenberg–Katznelson theorem. To provide a uniform lower bound for the number of desired configurations, we will utilize the following uniform version of the Furstenberg–Katznelson theorem by Bergelson, Host, McCutcheon and Parreau.

THEOREM 2.1 [1, Theorem 2.1(ii)]. *For any k and for any $\delta > 0$, there exists a constant $c(\delta, k) > 0$ depending only on δ and k such that the following holds. For any k commuting measure preserving transformations T_1, ..., T_k of a probability measure space $(X, \mathcal{B}, \mu)$ and any $A \in \mathcal{B}$ with $\mu(A) = \delta$, there exists a positive integer n such that $\mu(T_1^{-n} A \cap \cdots \cap T_k^{-n} A) > c(\delta, k)$.*

2.4. Summary of Tao and Ziegler's construction. Our proof relies heavily on the proof of Tao and Ziegler [18]. Since many detailed constructions in that proof will be used, we summarize them here for the convenience.

From now on we fix a dimension d, as well as a dense subset $A \subseteq \mathcal{P}^d$. There exists a $\delta > 0$ and a sequence N_n going to infinity as $n \to \infty$ such that

$$|A_n| \geq \delta |\mathcal{P} \cap [N_n]|^d$$

where $A_n = A \cap [N_n]^d$. By deleting a small number of elements from A_n (and reducing δ accordingly), we may assume $A_n \in [\delta' N_n, (1 - \delta') N_n]^d$ for some $0 < \delta' < 1/2$. Note that in [18], the authors only removed the primes in $[0, \delta' N_n]$. The reason for this removal is to make sure when n is large enough, all primes in A_n are coprime to W_n. Then they used the pigeonhole principle to choose $b_1, \ldots, b_d$ satisfying (6) later. Here, for our purpose, we also delete the primes in $[(1 - \delta')N_n, N_n]$. It is because later, we will need to shift the variable a in A_n. This deletion makes sure the shift does not move a out of $[N_n]$.

Let $\omega_n \ll \log \log \log N_n$ be a sufficiently slowly growing function which will be chosen later. Define $W_n = \prod_{p \leq \omega_n, p \in \mathcal{P}} p$ as in Section 2.1. For $-W_n < b_1, \ldots, b_d < W_n$ coprime to W_n, denote $b = (b_1, \ldots, b_d) \in \mathbb{Z}^d$. Define $N'_n = \lfloor N_n / W_n \rfloor$ and

$$(3) \qquad A'_n = \left\{ a \in [N'_n]^d : W_n a + b \in A_n \right\}$$

Let $M_n = o(N'_n)$ be a sequence of natural numbers. For $a \in A'_n$ and $r \in [M_n]$, let $B_{a,r,n} = \{b \in \mathbb{Z}^d : a + rb \in A'_n\}$. For finite set $\Omega \subseteq \mathbb{Z}^d$ and $n \in \mathbb{N}$, define a measure $\mu_{\Omega,n}$ on the Furstenberg $\mathbb{Z}^d$-system X by

$$\mu_{\Omega,n} = \mathbb{E}_{a \in [N'_n]^d} \mathbb{E}_{r \in [M_n]} \delta_{B_{a,r,n}} \prod_{i=1}^{d} \prod_{c_i \in \Omega_i} \Lambda_{W_n, b_i}(a_i + c_i r)$$

where Ω_i is the projection of Ω on the i-th coordinate, $a = (a_1, a_2, \ldots, a_d)$ and δ is the delta mass, i.e.

$$(4) \qquad \delta_B(S) = \begin{cases} 1 & \text{if } B \in S \\ 0 & \text{otherwise} \end{cases}$$

for any $B \subset \mathbb{Z}^d$ and $S \subset X$. Let E be the basic cylinder set

$$E = \left\{ B \in X : 0_{\mathbb{Z}^d} \in B \right\}$$

Then for every $h_1, \ldots, h_k \in \mathbb{Z}^d$, by definition of the measure $\mu_{\Omega,n}$, one has

$$(5) \qquad \mu_{\Omega,n}(T_{h_1} E \cap \cdots \cap T_{h_k} E) =$$

$$\mathbb{E}_{a \in [N'_n]^d} \mathbb{E}_{r \in [M_n]} \prod_{j=1}^{k} 1_{A'_n}(a + r h_j) \prod_{i=1}^{d} \prod_{c_i \in \Omega_i} \Lambda_{W_n, b_i}(a_i + c_i r)$$

In [18], it is shown that we can choose ω_n, b_1, ..., b_d and M_n such that Λ_{W_n,b_i} for $1 \leq i \leq d$ satisfy the linear forms condition (see Proposition 2.2 below) and

$$(6) \qquad \lim_{n\to\infty} \mathbb{E}_{a\in[N'_n]^d} \, 1_{A'_n}(a) \prod_{i=1}^{d} \Lambda_{W_n,b_i}(a_i) \geq \delta.$$

Then we define a measure μ on X by

$$(7) \qquad \mu = \text{p-lim}_{l\to\infty}\big(\text{p-lim}_{n\to\infty} \mu_{[-l,l]^d,n}\big)$$

where p-lim is a fixed Banach limit functional, i.e. a linear functional extending the standard limit functional lim on convergent sequences such that

$$(8) \qquad \liminf_{n\to\infty} x_n \leq \text{p-lim}_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

It is easy to see that the measure μ is well-defined and is a probability measure on $(X, \mathcal{B})$. The crucial point is that Tao and Ziegler used the linear forms condition to prove μ is invariant under T_h for all $h \in \mathbb{Z}^d$ and to prove the compatibility property ([18, Proposition 2.7]). For $h_1,\ldots,h_k \in \mathbb{Z}^d$, let Ω_j be the projection of $\{h_1,\ldots,h_k\}$ to j-th coordinate and $\Omega = \prod_{j=1}^{d} \Omega_j$. By definition,

$$\mu_{\Omega,n} = \mu_{\{h_1,\ldots,h_k\},n}.$$

The compatibility property says that if we have $\Omega' = \prod_{j=1}^{d} \Omega'_j$ such that $\Omega_j \subseteq \Omega'_j$ for all $1 \leq j \leq d$, then

$$\mu_{\Omega,n}(T_{h_1}E \cap \cdots \cap T_{h_k}E) = \mu_{\Omega',n}(T_{h_1}E \cap \cdots \cap T_{h_k}E) + o(1).$$

In other words,

$$(9) \qquad \mu_{\Omega',n}(T_{h_1}E \cap \cdots \cap T_{h_k}E) = \mu_{\{h_1,\ldots,h_k\},n}(T_{h_1}E \cap \cdots \cap T_{h_k}E) + o(1)$$

as long as Ω' of the form $\prod_{j=1}^{d} \Omega'_j$ and $\{h_1,\ldots,h_k\} \subseteq \Omega'$.

With the measure μ, one gets $(X, \mathcal{B}, \mu, (T_h)_{h\in\mathbb{Z}^d})$ is a measure preserving $\mathbb{Z}^d$-system. From (6) and (9),

$$(10) \qquad \mu(E) = \text{p-lim}_{n\to\infty} \mu_{\{0\},n}(E) = \text{p-lim}_{n\to\infty} \mathbb{E}_{a\in[N'_n]} \, 1_{A'_n}(a) \prod_{i=1}^{d} \Lambda_{W_n,b_i}(a_i) \geq \delta.$$

2.5. Linear forms condition. As mentioned earlier, the construction of the measure μ relies on the following property of W-tricked von Mangoldt functions called the linear forms condition.

PROPOSITION 2.2. *Let $d \geq 1$ be a fixed dimension. Let N'_n be a sequence of natural numbers going to infinity as $n \to \infty$. Then we can choose sequences $M_n = o(N'_n)$, $H_n = o(M_n)$ and ω_n such that the following is true:*

Let $m, k_1, \ldots, k_d \geq 0$ be natural numbers. For $1 \leq i \leq d$, $1 \leq j \leq k_i$, let $\phi_{i,j} \colon \mathbb{Z}^{d+m} \to \mathbb{Z}$ be linear forms with integer coefficients and pairwise linearly independent. One has

$$(11) \qquad \mathbb{E}_{a \in [N'_n]^d} \mathbb{E}_{r \in \prod_{j=1}^m L_{j,n}} \prod_{i=1}^d \prod_{j=1}^{k_i} \nu_{i,j,n}(\phi_{i,j}(a,r)) = 1 + o(1)$$

for all subintervals $L_{j,n}$ of $[-M_n, M_n]$ of length greater than H_n and $\nu_{i,j,n} = \Lambda_{W_n,b}$ for some $(b, W_n) = 1$.

REMARK 2.3. The linear forms condition found in [18, Proposition 2.5] is slightly weaker than the version presented here in the sense that it is restricted to linear forms of the form $\phi_{i,j}(a,r) = a_i + \psi_{i,j}(r)$ for some linear forms $\psi_{i,j} \colon \mathbb{Z}^m \to \mathbb{Z}$ where $a = (a_1, \ldots, a_d)$. The generalization here is needed for Proposition 3.1 and its proof is identical to [18, Proof of Theorem 1.4]. We summarize below for completeness.

PROOF. First note that the collection of $(m, k_1, \ldots, k_d, \phi_{i,j})$ is countable. Fix an arbitrary enumeration of them. By [12, Theorem 5.1], for $M \in \mathbb{N}$, there exists $\omega = \omega^{(M)}$ and $n^{(M)}$ such that

$$(12) \qquad \left| \mathbb{E}_{a \in [N'_n]^d} \mathbb{E}_{r \in \prod_{j=1}^m L_{j,n}} \prod_{i=1}^d \prod_{j=1}^{k_i} \nu_{i,j,n}(\phi_{i,j}(a,r)) - 1 \right| < \frac{1}{M}$$

for the first M choices of m, k_1, $\ldots$, k_d, $\phi_{i,j}$ as long as $L_{j,n}$ are subintervals of $[-N'_n/M, N'_n/M]$ of length greater than N'_n/M^2 and $n \geq n^{(M)}$.

We can choose $n^{(M)}$ so that the sequence $n^{(M)}$ is increasing in M. For $M \in \mathbb{N}$ and $n^{(M)} \leq n < n^{(M+1)}$, let $\omega_n = \omega^{(M)}$, $M_n = N'_n/M$ and $H_n = N'_n/M^2$. These choices satisfy the conclusion of Proposition 2.2. $\square$

2.6. Other tools. Similar to [15], we need the following elementary lemmas.

LEMMA 2.4 (Cauchy–Schwarz). *Let A, B be finite sets, f, F be functions on A and g be a function on $A \times B$. If $|f| \leq F$ pointwise. Then*

$$\left| \mathbb{E}_{a \in A, b \in B} f(a) g(a,b) \right|^2 \leq \mathbb{E}_{a \in A} F(a) \times \mathbb{E}_{a \in A} F(a) |\mathbb{E}_{b \in B} g(a,b)|^2$$

LEMMA 2.5 (van der Corput). *Let* $(x_n)_{n \in \mathbb{Z}}$ *be complex-valued sequence satisfying* $x_n = 0$ *outside of interval* $[N]$. *Then*

$$|\mathbb{E}_{n \in [N]} x_n|^2 \ll \mathbb{E}_{|h| < N} \mathbb{E}_{n \in [N]} x_n x_{n+h}.$$

3. Proof of Theorem 1.1

In this section, we prove Theorem 1.1 conditional on a proposition whose proof will be presented in the next section. Let the setup be as in Section 2.4. Fix $h_1, h_2, \ldots, h_k \in \mathbb{Z}^d$. By (7), for every $m \in \mathbb{N}$, one has

$$(13) \quad \mu\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big) = \text{p-lim}_{l \to \infty} \text{p-lim}_{n \to \infty} \mu_{[-l,l]^d, n}(T_{mh_1} E \cap \cdots \cap T_{mh_k} E).$$

By (9) (compatibility property), if $\{mh_1, \ldots, mh_k\} \subseteq [-l,l]^d$, then

$$\text{p-lim}_{n \to \infty} \mu_{[-l,l]^d, n}\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big)$$

$$= \text{p-lim}_{n \to \infty} \mu_{\{mh_1, \ldots, mh_k\}, n}\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big).$$

Therefore, we have
$$(14)$$
$$\mu\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big) = \text{p-lim}_{n \to \infty} \mu_{\{mh_1, \ldots, mh_k\}, n}\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big).$$

Since $(X, \mathcal{B}, \mu, (T_h)_{h \in \mathbb{Z}^d})$ is a measure preserving $\mathbb{Z}^d$-system and $\mu(E) = \delta > 0$, Theorem 2.1 implies that there exist $c(\delta, k) > 0$ and $m \in \mathbb{N}$ such that

$$(15) \qquad \mu(T_{mh_1} E \cap \cdots \cap T_{mh_k} E) > c(\delta, k).$$

By (14),

$$\text{p-lim}_{n \to \infty} \mu_{\{mh_1, \ldots, mh_k\}, n}\big(T_{mh_1} E \cap \cdots T_{mh_k} E\big) > c(\delta, k)$$

This means

$$(16) \qquad \mu_{\{mh_1, \ldots, mh_k\}, n}\big(T_{mh_1} E \cap \cdots \cap T_{mh_k} E\big) > c(\delta, k) - o(1)$$

By (5), the inequality (16) implies
$$(17)$$
$$\mathbb{E}_{a \in [N_n']^d} \mathbb{E}_{r \in [M_n]} \prod_{j=1}^{k} 1_{A_n'}(a + rmh_j) \prod_{i=1}^{d} \prod_{c_i \in \Omega_i} \Lambda_{W_n, b_i}(a_i + c_i r) > c(\delta, k) - o(1)$$

where $\Omega_i \subset \mathbb{Z}$ is the projection of the set $\Omega = \{mh_1, \ldots, mh_k\}$ on the i-th coordinate.

The following proposition is our version of the comparison method.

PROPOSITION 3.1. *Let everything be as before. Then*

$$(18) \qquad \mathbb{E}_{r\in[M_n]}\big(\Lambda_{W_n,1}(mr)-1\big)\mathbb{E}_{a\in[N_n']^d}\prod_{j=1}^{k}1_{A_n'}(a+rmh_j)$$

$$\times\prod_{i=1}^{d}\prod_{c_i\in\Omega_i}\Lambda_{W_n,b_i}(a_i+c_ir)=o(1).$$

Assume we already had Proposition 3.1. Then (18) together with (17) imply that

$$(19)\ \ \mathbb{E}_{r\in[M_n]}\mathbb{E}_{a\in[N_n']^d}\Lambda_{W_n,1}(mr)\prod_{j=1}^{k}1_{A_n'}(a+rmh_j)\prod_{i=1}^{d}\prod_{c_i\in\Omega_i}\Lambda_{W_n,b_i}(a_i+c_ir)$$

$$> c(\delta,k)-o(1).$$

Hence for any sufficiently large n, the set A_n' contains a configuration of the form $a+rmh_1,\ldots,a+rmh_k$ with $W_nmr+1\in\mathcal{P}$. This is equivalent to saying that A_n contains $W_na+b+W_nmrh_j$ for $1\le j\le k$ with $W_nmr\in\mathcal{P}-1$. This finishes our proof of Theorem 1.1 for the step being $p-1$ for some $p\in\mathcal{P}$. For the case $p+1$, simply replace $\Lambda_{W_n,1}(mr)-1$ by $\Lambda_{W_n,-1}(mr)-1$ in Proposition 3.1. The rest of the proof remains the same.

In fact, (19) gives us a lower bound for number of the pairs $(\tilde{a},p)\in[N_n]^d\times[W_nmM_n+1]$ satisfying $p\in\mathcal{P}$ and $\tilde{a}+(p-1)h_j\in A$ for all $1\le j\le k$. This number is not less than the number of pairs $(a,r)\in[N_n']^d\times[M_n]$ such that $W_nmr+1\in\mathcal{P}$ and $a+mrh_j\in A_n'$. By (19), the latter is greater than

$$(20)$$
$$(c(\delta,k)-o(1))\times\frac{M_nN_n'^d}{\log(W_nmM_n+1)\log^{\sum_{i=1}^{d}|\Omega_i|}N_n}\times\left(\frac{W_n}{\phi(W_n)}\right)^{1+\sum_{i=1}^{d}|\Omega_i|}.$$

Since $N_n'=\lfloor N_n/W_n\rfloor$ and $W_n=o(\log(M_n))$, we get (20) is equal to

$$(21)\qquad (c(\delta,k)-o(1))\times\frac{M_nN_n^d}{\log M_n\log^{\sum_{i=1}^{d}|\Omega_i|}N_n}\times\frac{W_n^{1-d+\sum_{i=1}^{d}|\Omega_i|}}{\phi(W_n)^{1+\sum_{i=1}^{d}|\Omega_i|}}.$$

4. Proof of Proposition 3.1

The rest of our paper is devoted to proving Proposition 3.1. The strategy is similar to [15, Proposition 1] which in turn is inspired by the method in [6,7]. The idea here is that after a finite number of applications of Cauchy–Schwarz and van der Corput, the left hand side of (18) is bounded by an expression consisting entirely of W-tricked von Mangoldt functions. This expression then tends to zero by the linear forms condition.

4.1. A toy example. To illustrate the ideas of the proof, we first work with the following special case. The proof of the general case is not much different.

In this model case, we take $d = 2$, $h_1 = (1, 1)$ and $h_2 = (2, 1)$. Then the projection of $\Omega = \{h_1, h_2\}$ onto the first coordinate is $\Omega_1 = \{1, 2\}$ and onto the second coordinate is $\Omega_2 = \{1\}$. We need to show:

$$(22) \qquad \mathbb{E}_{r \in [M_n]}(\Lambda_{W,1}(mr) - 1) \mathbb{E}_{(a_1, a_2) \in [N_n']^2} 1_{A_n'}((a_1 + mr, a_2 + mr))$$

$$\times 1_{A_n'}((a_1 + 2mr, a_2 + mr))$$

$$\times \Lambda_{W,b_1}(a_1 + mr)\Lambda_{W,b_1}(a_1 + 2mr)\Lambda_{W,b_2}(a_2 + mr) = o(1).$$

From now on, for simplicity, we write W instead of W_n. Likewise, we will assume $m = 1$ since doing so significantly simplifies the notation while not affecting the proof.

For $n \in \mathbb{N}$ and $a = (a_1, a_2) \in \mathbb{Z}^2$, define following functions:

- $\nu(n) = \Lambda_{W,1}(n)$.
- $\nu_{1,1}(a) = \Lambda_{W,b_1}(a_1)$ and $\nu_{1,2}(a) = \Lambda_{W,b_2}(a_2)$.
- $\nu_{2,1}(a) = \Lambda_{W,b_1}(a_1)$ and $\nu_{2,2}(a) = 1$.
- $\theta_1(a) = \nu_{1,1}(a)\nu_{1,2}(a)$.
- $\theta_2(a) = \nu_{2,1}(a)\nu_{2,2}(a)$.

Note that the way we define $\nu_{i,j}$ $(i, j = 1, 2)$ depends on the set $\{h_1, h_2\}$. Since the second coordinate of h_2 coincides with the second coordinate of h_1, we define $\nu_{2,2}(a) = 1$. On the other hand, if they did not coincide, we would define $\nu_{2,2}(a) = \Lambda_{W,b_2}(a_2)$.

The left hand side of (22) becomes

$$(23) \qquad \mathbb{E}_{r \in [M_n]}(\nu(r) - 1) \mathbb{E}_{a \in [N_n']^2} \prod_{j=1}^{2} 1_{A_n'}(a + rh_j)\theta_j(a + rh_j).$$

As discussed in Section 2.4, the way we truncate A_n (hence A_n') allows us to shift $a + rh_1$ to a. Hence (23) is equal to

$$(24)$$
$$\mathbb{E}_{a \in [N_n']^2} 1_{A_n'}(a)\theta_1(a) \mathbb{E}_{r \in [M_n]}(\nu(r) - 1) 1_{A_n'}(a + (h_2 - h_1)r)\theta_2(a + (h_2 - h_1)r).$$

By Cauchy–Schwarz inequality, the square of previous expression is at most

$$(25) \qquad \mathbb{E}_{a \in [N_n']^2} 1_{A_n'}(a)\theta_1(a) \mathbb{E}_{a \in [N_n]62} 1_{A_n'}(a)\theta_1(a)$$

$$\times \left| \mathbb{E}_{r \in [M_n]}(\nu(r) - 1) 1_{A_n'}(a + (h_2 - h_1)r)\theta_2(a + (h_2 - h_1)r) \right|^2$$

Observe that

$$\mathbb{E}_{a \in [N_n']^2} 1_{A_n'}(a)\theta_1(a) \leq \mathbb{E}_{a \in [N_n']^2}\theta_1(a) = 1 + o(1)$$

where the last equality follows from the linear forms condition. Hence up to a factor of $1 + o(1)$, (25) is at most

$$(26) \quad \mathbb{E}_{a \in [N_n']^2} \theta_1(a) \big| \mathbb{E}_{r \in [M_n]} (\nu(r) - 1) 1_{A_n'} (a + (h_2 - h_1)r) \theta_2(a + (h_2 - h_1)r) \big|^2$$

By van der Corput's lemma, the previous expression is at most (up to a multiplicative constant)

$$(27) \qquad \mathbb{E}_{a \in [N_n']^2} \theta_1(a) \, \mathbb{E}_{r \in [M_n], |s_1| < M_n} (\nu(r) - 1)(\nu(r + s_1) - 1)$$
$$\times 1_{A_n'} (a + (h_2 - h_1)r) \theta_2(a + (h_2 - h_1)r)$$
$$\times 1_{A_n'} (a + (h_2 - h_1)r + (h_2 - h_1)s_1) \theta_2(a + (h_2 - h_1)r + (h_2 - h_1)s_1).$$

By shifting $a + (h_2 - h_1)r$ to a, the previous expression is equal to

$$(28) \qquad \mathbb{E}_{a \in [N_n']^2} 1_{A_n'}(a) \theta_2(a) \, 1_{A_n'} (a + (h_2 - h_1)s_1) \theta_2(a + (h_2 - h_1)s_1)$$
$$\times \, \mathbb{E}_{r \in [M_n], |s_1| < M_n} (\nu(r) - 1)(\nu(r + s_1) - 1) \theta_1(a + (h_1 - h_2)r).$$

We now move everything that does not depend on r outside of the average on r. Extra caution should be made here. At first glance, we may leave $\theta_1(a + (h_1 - h_2)r)$ inside since it seemingly depends on r. However, as $h_1 - h_2 = (-1, 0)$

$$(29) \qquad \qquad \theta_1(a + (h_1 - h_2)r) = \nu_{1,1}(a_1 - r) \nu_{1,2}(a_2).$$

Because $\nu_{1,2}(a_2)$ does not depend on r, we move it outside of the average on r. Thus (28) is equal to

$$(30) \quad \mathbb{E}_{a \in [N_n']^2} 1_{A_n'}(a) \theta_2(a) \, 1_{A_n'} (a + (h_2 - h_1)s_1) \theta_2(a + (h_2 - h_1)s_1) \nu_{1,2}(a_2)$$
$$\times \, \mathbb{E}_{r \in [M_n], |s_1| < M_n} (\nu(r) - 1)(\nu(r + s_1) - 1) \nu_{1,1}(a_1 - r).$$

By Cauchy–Schwarz inequality and van der Corput's lemma again, the square of above equation is at most (up to a multiplicative constant)

$$(31) \quad \mathbb{E}_{a \in [N_n']^2, r \in [M_n], |s_1| < M_n, |s_2| < |M_n|} (\nu(r) - 1)(\nu(r + s_1) - 1)(\nu(r + s_2) - 1)$$
$$\times \, (\nu(r + s_1 + s_2) - 1) \nu_{2,1}(a_1) \nu_{2,1}(a_1 + s_1) \nu_{1,2}(a_2) \nu_{1,1}(a_1 - r) \nu_{1,1}(a_1 - r - s_2).$$

Expanding out the last expression, we see that it is equal to a sum and difference of 16 averages of the form (11). Since no two forms in (31) are linearly dependent, each of these 16 averages is equal to $1 + o(1)$ by the linear forms condition. Therefore (31) is equal to $o(1)$. This finishes the proof of the special case.

4.2. The general case. We now prove Proposition 3.1 in its full generality. Define $\nu = \Lambda_{W,1}$ as before. For $1 \le j \le k$ and $1 \le i \le d$, define $\nu_{j,i} : \mathbb{Z}^d \to \mathbb{R}$ by

$$(32) \qquad \nu_{j,i}(a) = \begin{cases} \Lambda_{W,b_i}(a_i) & \text{if } h_{j,i} \ne h_{l,i} \ \forall 1 \le l < j \\ 1 & \text{otherwise.} \end{cases}$$

Then define $\theta_j = \prod_{i=1}^d \nu_{j,i}$.
   We need to show

$$(33) \qquad \mathbb{E}_{r \in [M_n]}(\nu(r) - 1)\, \mathbb{E}_{a \in [N_n']^d} \prod_{j=1}^k 1_{A_n'}(a + rh_j)\theta_j(a + rh_j) = o(1).$$

As in the toy example, we perform the following steps:
   (1) Shift a.
   (2) Move everything that does not depend on r outside of the average on r.
   (3) Apply Cauchy–Schwarz inequality.
   (4) Bound $1_{A_n'}\theta_j$ by θ_j.
   (5) Apply van der Corput's lemma.
   After k iterations, we get the following expression:

$$(34) \qquad \mathbb{E}_{a \in [N_n']^d, r \in [M_n], |s_1| < M_n, \ldots, |s_k| < M_n} \prod_{R \subseteq [k]} \left(\nu\left(r + \sum_{l \in R} s_l\right) - 1 \right)$$

$$\times \prod_{j=1}^k \prod_{i=1}^d \prod_{R \subseteq R_{j,i}} \nu_{j,i}\left(a_i + (h_{j,i} - h_{k,i})r + \sum_{l \in R}(h_{j,i} - h_{l,i})s_l \right)$$

where $R_{j,i} = \{1 \le l \le k : h_{l,i} \ne h_{j,i}\}$. Again, this expression is a sum and difference of 2^{2^k} averages of the form (11). In order to invoke the linear forms condition, it suffices to verify that no two forms appearing in (34) are linearly dependent.
   Firstly, those forms appearing in ν are independent from one another because the appearance of $s_1, s_2, \ldots, s_k$ in each form is corresponding to a subset R of $[k]$. They are also independent from the forms in $\nu_{j,i}$ because a_i appears in $\nu_{j,i}$, but not in ν.
   Secondly, for $i_1 \ne i_2$ and any j_1, j_2 (not necessarily distinct), the forms appearing in ν_{j_1,i_1} and ν_{j_2,i_2} are independent because a_{i_1} appears in ν_{j_1,i_1} while a_{i_2} appears in ν_{j_2,i_2}.
   For a fixed i, if $(h_{l,i} - h_{k,i})r = (h_{j,i} - h_{k,i})r$ for some $l < j$, then $h_{l,i} = h_{j,i}$. By the way we define $\nu_{j,i}$, this would force $\nu_{j,i} = 1$. Hence if $\nu_{j,i} \ne 1$,

$(h_{l,i} - h_{k,i})r \neq (h_{j,i} - h_{k,i})r$. This implies the forms appearing in $\nu_{j,i}$ are independent from the forms appearing in $\nu_{l,i}$ for all $l < j$.

And lastly, for fixed i and j, the forms appearing in $\nu_{j,i}$ are independent from one another because each form is in one-to-one correspondence with a subset of $R_{j,i}$. $\square$

References

[1] V. Bergelson, B. Host, R. Mccutcheon, and F. Parreau, Aspects of uniformity in recurrence, *Colloq. Math.*, **84/85** (2000), 549–576.

[2] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. Amer. Math. Soc.*, **9** (1996), 725–753.

[3] V. Bergelson, A. Leibman, and T. Ziegler, The shifted primes and the multidimensional Szemerédi and polynomial van der Waerden theorems, *C. R. Math. Acad. Sci. Paris*, **349** (2011), 123–125.

[4] B. Cook, A. Magyar, and T. Titichetrakun, A multidimensional Szemerédi theorem in the primes via combinatorics, *Ann. Comb.*, **22** (2018), 711–768.

[5] J. Fox and Y. Zhao, A short proof of the multidimensional Szemerédi theorem in the primes, *Amer. J. Math.*, **4** (2015), 1139–1145.

[6] N. Frantzikinakis, B. Host, and B. Kra, Multiple recurrence and convergence for sequences related to the prime numbers, *J. Reine Angew. Math.*, **611** (2007), 131–144.

[7] N. Frantzikinakis, B. Host, and B. Kra, The polynomial multidimensional Szemerédi theorem along shifted primes, *Israel J. Math.*, **194** (2013), 331–348.

[8] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.*, **31** (1977), 204–256.

[9] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. d'Analyse Math.*, **34** (1978), 275–291.

[10] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math.*, **167** (2008), 481–547.

[11] B. Green and T. Tao, Linear equations in primes, *Ann. of Math. (2)*, **171** (2010), 1753–1850.

[12] B. Green and T. Tao, Linear equations in primes, *Ann. of Math.*, **171** (2010), 1753–1850.

[13] B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences, *Ann. of Math. (2)*, **175** (2012), 541–566.

[14] B. Green, T. Tao, and T. Ziegler, An inverse theorem for the Gowers $U^{s+1}[N]$-norm, *Ann. of Math. (2)*, **176** (2012), 1231–1372.

[15] T. Lê and J. Wolf, Polynomial configurations in the primes, *Int. Math. Res. Not.*, **2014** (2014), 6448–6473.

[16] E. Szemerédi, On the sets of integers containing no k elements in arithmetic progressions, *Acta Arith.*, **27** (1975), 299–345.

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

[18] T. Tao and T. Ziegler, A multi-dimensional Szemerédi theorem for the primes via a correspondence principle, *Israel J. Math.*, **207** (2015), 203–228.

[19] P. Varnavides, On certain sets of positive density, *J. London Math. Soc.*, **34** (1959), 358–360.

[20] T. Wooley and T. Ziegler, Multiple recurrence and convergence along the primes, *Amer. J. Math.*, **134** (2012), 1705–1732.