THE GREEN-TAO THEOREM FOR PIATETSKI-SHAPIRO PRIMES

HONGZE LI AND HAO PAN

Abstract. Let $m \geq 3$. Suppose that

$$1 - 2^{-2^{m^2 m}} < \gamma < 1.$$ 

Then the set

$$\{ p \text{ prime : } p = [n^\gamma] \text{ for some } n \in \mathbb{N} \}$$

contains infinitely many non-trivial $m$-term arithmetic progressions.

1. Introduction

The well-known Green-Tao theorem [8] asserts that for each $m \geq 3$, the set of all primes $\mathcal{P}$ contains infinitely many non-trivial arithmetic progressions of length $m$. That is, for each $m \geq 3$, there exists infinitely many pairs of positive integers $a, d$ such that $a, a + d, \ldots, a + (m - 1)d$ are all primes. In fact, Green and Tao proved a Szemerédi-type theorem for primes: any subset $A$ of $\mathcal{P}$ with $\overline{d}_\mathcal{P}(A) > 0$ contains infinitely many non-trivial $m$-terms arithmetic progressions for each $m \geq 3$, where the relative upper density

$$\overline{d}_\mathcal{P}(A) := \limsup_{X \to \infty} \frac{|A \cap [1, X]|}{|\mathcal{P} \cap [1, X]|}. $$

The Green-Tao theorem has been generalized for the primes of some special forms, including the Chen primes (by Zhou [23]), the primes $p$ such that the interval $[p + 1, p + 7 \times 10^7]$ contains at least one prime (by Pintz [16]), the primes of the form $x^2 + y^2 + 1$ (by Sun and Pan [18]), etc.. For the further generalizations of the Green-Tao theorem, the reader may refer to [9, 12, 13, 19, 20, 21].

The Piatetski-Shapiro prime is another kind of primes of the special form. A well-known conjecture asserts that there exist infinitely many primes of the form $n^2 + 1$. This conjecture is far from solved under the current techniques, though Iwaniec [11] proved that there exist infinitely many $n$ such that $n^2 + 1$ has at most two prime factors. In 1953, Piatetski-Shapiro [15] considered another approximation to this conjecture. Suppose that $\gamma \in (0, 1)$ and $\gamma^{-1} \notin \mathbb{Z}$. Then $x^{\frac{1}{\gamma}}$ can be viewed

2010 Mathematics Subject Classification. Primary 11P32; Secondary 05D10, 11B25, 11L07, 11N36.

Key words and phrases. Piatetski-Shapiro prime, Green-Tao theorem, pseudorandom measure.
as a polynomial of degree $\gamma^{-1}$. Let

$$\mathbb{N}_{1}^{\gamma} := \{[n^{\frac{1}{\gamma}}] : n \in \mathbb{N}\}$$

where $[x]$ denotes the integer part of $x$, i.e., $\mathbb{N}_{1}^{\gamma}$ is the set of all non-negative integers of the form $[n^{\frac{1}{\gamma}}]$. Clearly $|\mathbb{N}_{1}^{\gamma} \cap [0, x]| = x^\gamma + O(1)$ for any sufficiently large $x$. Piatetski-Shapiro [15] proved that for any $\gamma \in (11/12, 1)$, there exist infinitely many primes lying in $\mathbb{N}_{1}^{\gamma}$. Explicitly, he obtained that

$$|\mathcal{P}_{\gamma} \cap [1, x]| = (1 + o(1)) \cdot \frac{x^\gamma}{\log x}$$

(1.1)

as $x \to +\infty$, where

$$\mathcal{P}_{\gamma} := \{p \text{ prime} : p \in \mathbb{N}_{1}^{\gamma}\},$$

is the set of those primes of the form $[n^{\frac{1}{\gamma}}]$. Nowadays, the primes in $\mathcal{P}_{\gamma}$ is also called Piatetski-Shapiro primes. The result of Piatetski-Shapiro has been improved many times. The current best known result on the distribution of the Piatetski-Shapiro primes is due to Rivat and Wu [17], they proved that

$$|\mathcal{P}_{\gamma} \cap [1, x]| \gg \frac{x^\gamma}{\log x}$$

(1.2)

for any $\gamma \in (205/243, 1)$.

On the other hand, with the help of Heath-Brown identity [10], Balog and Friedlander [1] proved that for any $\gamma \in (20/21, 1)$, every sufficiently large odd integer can be represented as the sum of three primes lying in $\mathcal{P}_{\gamma}$. The key ingredient of Balog and Friedlander’s proof is the following estimation of exponential sum:

$$\frac{1}{\gamma} \sum_{p \in \mathcal{P}_{\gamma} \cap [1, x]} p^{1-\gamma} \log p \cdot e(p\theta) = \sum_{p \leq [1, x]} \log p \cdot e(p\theta) + O(x^{1-%})$$

(1.3)

for any $\gamma \in (8/9, 1)$, where $e(\theta) = \exp(2\pi \sqrt{-1}\theta)$ and $\epsilon > 0$ is a constant only depending on $\gamma$. Clearly using the discussions of Balog and Friedlander, one can easily prove that $\mathcal{P}_{\gamma}$ contains infinitely many non-trivial three-term arithmetic progressions for each $\gamma \in (20/21, 1)$. Furthermore, with the help of the transference principle, Mirek [14] obtained a Roth-type theorem for the Piatetski-Shapiro primes and showed that for any $\gamma \in (71/72, 1)$ and any subset $\mathcal{A} \subseteq \mathcal{P}_{\gamma}$ with

$$\limsup_{x \to +\infty} \frac{\mathcal{A} \cap [1, x]}{x^{\gamma(\log x)^{-1}}} > 0,$$

$\mathcal{A}$ contains infinitely many non-trivial three-term arithmetic progressions.

It is natural to ask whether the Piatetski-Shapiro primes contain longer non-trivial arithmetic progressions. In this paper, we shall prove the following Szemerédi-type result for Piatetski-Shapiro primes.
Theorem 1.1. Let $m \geq 3$. Suppose that

$$1 - 2^{-2^{m-2}m^2} < \gamma < 1$$

and $A$ is a subset of $\mathcal{P}_\gamma$ with

$$\limsup_{x \to +\infty} \frac{|A \cap [1,x]|}{x^\gamma (\log x)^{-1}} > 0.$$ 

Then $A$ contains infinitely many non-trivial $m$-term arithmetic progressions.

Of course, the lower bound of $\gamma$ in (1.4) is not very optimal, and surely could be improved via some more accurate calculations. However, we believe that with the help of the current techniques, it is impossible to obtain a lower bound of $\gamma$ independent on $m$. In fact, by the heuristic discussions, (1.1) should be valid for any $\gamma \in (0,1)$ with $\gamma^{-1} \not\in \mathbb{Z}$, which is evidently very far from being proved. On the other hand, Green and Tao [9] introduced the complexity of linear equations and showed that the non-trivial $m$-term arithmetic progressions factly correspond to a linear equation with the complexity $m - 2$. Now for the simplest equation $p = \lfloor n^{\frac{1}{2}} \rfloor$, we only can prove the existence of solutions when $205/243 < \gamma < 1$. So for those equations with higher complexity, the lower bound of $\gamma$ would naturally become worse.

Let us introduce the strategy for the proof of Theorem 1.1. The key ingredient is to construct a suitable pseudorandom measure for Piatetski-Shapiro primes and verify the corresponding linear forms condition. By revising the construction of Green-Tao, it is easy to obtain such a pseudorandom measure. However, our main difficulty is how to verify the linear forms condition. One reason is that the Piatetski-Shapiro prime theorem (1.1) arises from the estimations of exponential sums, rather than the sieve method. Explicitly, in order to prove Theorem 1.1, we have to give a non-trivial upper bound for the exponential sum of the form

$$\sum_{X \leq n \leq X+Y} e\left(s_1\psi_1(n) + s_2\psi_2(n) + \cdots + s_h\psi_h(n)\right),$$

where $s_1, \ldots, s_h \in \mathbb{Z}$ and $\psi_1, \ldots, \psi_h$ are some linear functions.

The classical van der Corput theorem is a useful tools to estimate the exponential sum of the form

$$\sum_{X \leq n \leq X+Y} e(f(n)),$$

where $f$ is a smooth function over the interval $[X, X + Y]$. Suppose that

$$\lambda \leq |f''(x)| \leq \alpha \lambda$$

for each $x \in [X, X + Y]$, where $\alpha, \lambda > 0$ are independent on $x$. Then the van der Corput theorem asserts that

$$\sum_{X \leq n \leq X+Y} e(f(n)) \ll \alpha Y (\lambda^{\frac{1}{2}} + Y^{-1} \lambda^{-\frac{1}{2}}).$$
Unfortunately, since it is possible that some of $s_i$ in \[1.5\] are negative and the others are positive, for the function $f(x) = s_1\psi_1(x) + \cdots + s_h\psi_h(x)$, generally \[1.6\] doesn’t hold.

Our strategy is to apply the generalized van der Corput theorem concerning the derivatives of higher order. However, for $f(x) = \sum_i s_i\psi_i(x)$ and any given integer $r \geq 2$, we also don’t know whether $|f^{(r)}(x)|$ could be bounded by $\lambda$ and $\alpha\lambda$. So there are two key ingredients in our proof. First, with the help of some suitable linear transformations, the estimation of \[1.5\] can be reduced to a special case that $\psi_i(x) = a_i x + b_i$ with $1 \leq |a_1| < |a_2| < \ldots < |a_h| \leq M$. Next, we can show that for any $s_1, \ldots, s_h \in \mathbb{Z}$, there exists $r \in [2, c(M)]$, where $c(M)$ is a constant only depending on $M$, such that

$$\lambda \leq |f^{(r)}(x)| \leq \alpha\lambda$$

for any $x \in [X, X + Y]$, where $\alpha, \lambda > 0$ and $\alpha$ only depends on $M$. Thus by using a generalization of van der Corput’s theorem, we can get a desired upper bound for the exponential sum \[1.5\].

The whole paper will be organized as follows. First, in the next section, we shall introduce Green-Tao’s transference principle and give the definitions of the pseudorandom measure and the linear forms condition. Then in the third section, we shall construct a pseudorandom measure $\nu$ for Piatetski-Shapiro primes. In Section 4, in order to verify the linear forms condition for $\nu$, we shall reduce a Goldston-Yıldırım-type estimation to the estimation of some exponential sums. Section 5 is the core part of the proof of Theorem 1.1, which will contain two key auxiliary lemmas. Finally, in Section 6, we shall complete the proof of Theorem 1.1 by combining a generalized van der Corput theorem with the two lemmas in Section 5.

Throughout this paper, $f(x) \ll g(x)$ means $f(x) = O(g(x))$ as $x$ tends to $\infty$. Furthermore, without the additional mentions, the implied constants in $O$, $\ll$ and $\gg$ at most depends on $m$. As usual, let $\phi(\cdot)$ and $\mu(\cdot)$ denote the Euler totient function and the Möbius function respectively. Furthermore, let $\log_k$ denote the $k$-th iteration of the logarithm function.

### 2. Transference Principle

Let $N$ be a sufficiently large prime and let $\mathbb{Z}_N := \mathbb{Z}/NZ$ denote the cyclic group of order $N$. Suppose that

$$\nu(n) : \mathbb{Z}_N \to \mathbb{R}$$

is a non-negative function. First, we introduce the definition of the $(h_0, k_0, m_0)$-linear forms condition. For $1 \leq h \leq h_0$ and $1 \leq k \leq k_0$, suppose that

$$\psi_i(x_1, \ldots, x_k) = a_{i1}x_1 + \cdots + a_{ik}x_k + b_i, \quad 1 \leq i \leq h,$$

where $b_i, a_{ij} \in \mathbb{Q}$ and the numerators and denominators of those $a_{ij}$ all lie in $[-m_0, m_0]$. Furthermore, assume that $\psi_i(x_1, \ldots, x_k)$ is not a rational multiple of
ψ_j(x_1,\ldots,x_k) for any distinct i,j. Clearly, we may also view those ψ_i as the linear functions over \(\mathbb{Z}_N\) whenever \(N > m_0\). Then we say \(\nu\) obeys the \((h_0,k_0,m_0)\)-linear forms condition, provided that as \(N \to \infty\),
\[
\frac{1}{N^k} \sum_{x_1,\ldots,x_k \in \mathbb{Z}_N} \nu(\psi_1(x_1,\ldots,x_k)) \cdots \nu(\psi_h(x_1,\ldots,x_k)) = 1 + o_{h_0,k_0,m_0}(1),
\]
for any \(1 \leq h \leq h_0, 1 \leq k \leq k_0\) and those \(\psi_1,\ldots,\psi_h\).

Next, we say \(\nu\) satisfies the \(k_0\)-correlation condition, if for any \(1 \leq k \leq k_0\), there exists a non-negative weight function \(\tau_k: \mathbb{Z}_N \to \mathbb{R}\) such that
\[
\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \tau_k(x)^s = O_{k,s}(1)
\]
for any integer \(s \geq 1\), and
\[
\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \nu(x + b_1) \cdots \nu(x + b_k) \leq \sum_{1 \leq i < j \leq k} \tau_k(b_i - b_j)
\]
for any \(b_1,\ldots,b_k \in \mathbb{Z}_N\).

Call \(\nu(x)\) a \(m\)-pseudorandom measure, provided that \(\nu\) obeys \((2^{m-1}m,3m-4,m)\)-linear forms condition and \(2^{m-1}\)-correlation condition. The important transference principle of Green and Tao [8, Theorem 3.5] asserts that

**Theorem 2.1.** Suppose that \(\delta > 0\) and \(m \geq 3\). Let \(f(x)\) be a function over \(\mathbb{Z}_N\) such that
\[
\frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) \geq \delta,
\]
and
\[
0 \leq f(x) \leq \nu(x)
\]
for each \(x \in \mathbb{Z}_N\), where \(\nu\) is a \(m\)-pseudorandom measure over \(\mathbb{Z}_N\). Then as \(N \to \infty\),
\[
\frac{1}{N^2} \sum_{x,y \in \mathbb{Z}_N} f(x)f(x+y) \cdots f(x+(m-1)y) \geq c(m,\delta) + o_{k,\delta}(1),
\]
where \(c(m,\delta)\) is a constant only depending on \(m\) and \(\delta\).

With the help of the arguments of Goldston and Yıldırım [5], for any \(m \geq 3\), Green and Tao constructed a \(m\)-pseudorandom measure \(\nu(n)\) over \(\mathbb{Z}_N\) such that
\[
\nu(n) \geq c_0 \cdot \lambda_{W,b}(n)
\]
for any \(n \in [\epsilon_0 N, 2\epsilon_0 N]\), where \(W = \prod_{p \leq \log_4 N} p\),
\[
\lambda_{W,b}(n) = \begin{cases} 
\frac{\phi(W)}{W} \cdot \log(Wn+b), & \text{if } Wn+b \text{ is prime}, \\
0, & \text{otherwise},
\end{cases}
\]
where \(\phi(W)\) denotes the Euler's totient function of \(W\).
and $c_0, \epsilon_0 > 0$ are two small constants only depending on $m$. Let 

$$f(x) = \begin{cases} c_0 \lambda_{W,b}(x), & \text{if } x \in [\epsilon_0 N, 2\epsilon_0 N], \\ 0, & \text{otherwise}, \end{cases}$$

be a function over $\mathbb{Z}_N$. According to the Siegel-Walfisz theorem, we have 

$$\sum_{x \in \mathbb{Z}_N} f(x) = c_0 \epsilon_0 N + o(N)$$

as $N \to \infty$. It follows from (2.3) and (2.4) that 

$$\sum_{x,r \in \mathbb{Z}_N} f(x)f(x+r) \cdots f(x+(m-1)r) \geq c_1 N^2$$

(2.6)

for some constant $c_1 > 0$. Clearly $f(x)f(x+r) \cdots f(x+(m-1)r) > 0$ implies that $x, x+r, \ldots, x+(m-1)r$ modulo $N$ all lie in the interval $[\epsilon_0 N, 2\epsilon_0 N]$. Since $\epsilon_0 < 1/2$ and $x \in [\epsilon_0 N, 2\epsilon_0 N]$, if $1 \leq r < N/2$, it is impossible that $x + r - N \in [\epsilon_0 N, 2\epsilon_0 N]$. Hence we must have $x + r \in [\epsilon_0 N, 2\epsilon_0 N]$, as well as $x + 2r, \ldots, x + (m-1)r$. Suppose that $N/2 < r < N$. Letting $r' = N - r$ and $x' \in [\epsilon_0 N, 2\epsilon_0 N]$ with $x' \equiv x + (m-1)r \pmod{N}$, we also have $x' + r', \ldots, x' + (m-1)r' \in [\epsilon_0 N, 2\epsilon_0 N]$. Recall that $f(n) \leq 2c_0\phi(W)\log N/W$. We obtain that there exist at least 

$$\frac{c_1 N^2}{2} \left( \frac{W}{2c_0\phi(W)\log N} \right)^m - \epsilon_0 N$$

pairs of $(x, r)$ with $1 \leq x, r < N$ such that $Wx + b, \ldots, W(x + (m-1)r) + b$ form a non-trivial arithmetic progression in primes.

Green and Tao's transference principle becomes a powerful tools to prove the relative Szemerédi-type theorems nowadays. Furthermore, in [2], Conlon, Fox and Zhao weakened the requirements concerning the pseudorandom measures. In fact, they defined the notion of $k$-linear forms condition. Let $\nu$ be a non-negative function over $\mathbb{Z}_N$. Suppose that as $N \to \infty$, 

$$\frac{1}{N^{2k}} \sum_{x_1, y_1, \ldots, x_k, y_k \in \mathbb{Z}_N} \prod_{j=1}^{k} \prod_{I \subseteq \{1, \ldots, k\}, I \cap J = \emptyset} \nu \left( \sum_{s \in I} (s-j)x_s + \sum_{t \in J} (t-j)y_t \right)^{\delta_{j,I,J}} = 1 + o(1)$$

for any choice of $\delta_{j,I,J} \in \{0, 1\}$. Then we say $\nu$ obeys the $k$-linear forms condition. Conlon, Fox and Zhao proved that 

**Theorem 2.2.** Suppose that $\delta > 0$ and $m \geq 3$. Let $f(x)$ be a non-negative function over $\mathbb{Z}_N$ such that 

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) \geq \delta,$$

and 

$$0 \leq f(x) \leq \nu(x)$$
for some function $\nu$ obeys the $m$-linear forms condition. Then (2.3) is also valid.

Clearly the $m$-linear forms condition is weaker than the $(2^{m-1}m, 2m, m)$-linear forms condition. So in order to get a relative Szemerédi-type theorem for the Piatetski-Shapiro primes, we only need to construct a suitable pseudorandom measure $\nu$ over $\mathbb{Z}_N$ and verify the $(2^{m-1}m, 2m, m)$-linear forms condition for $\nu$. This is our main task in the remainder sections.

3. Pseudorandom Measure

Let $\gamma$ and $m$ be given in Theorem 1.1. In this section, we shall construct a pseudorandom measure $\nu$ for those primes in $P_\gamma$. Let

$$\kappa_1 = \frac{1}{2^m \cdot (m + 4)!}, \quad \kappa_2 = \kappa_1 \cdot \left(1 + \frac{1}{2^{m-2}m^5}\right).$$

Suppose that $X$ is a sufficiently large positive integer. Let

$$W = \prod_{p \leq \log_4 X \text{ \(p\) is prime}} p.$$

According to the Piatetski-Shapiro prime number theorem,

$$|P_\gamma \cap [(\kappa_1 + \kappa_0)X, (\kappa_2 - \kappa_0)X]| = (\kappa_2 - \kappa_1 - 2\kappa_0 + o(1)) \cdot \frac{X^\gamma}{\log X},$$

where $\kappa_0 = (\kappa_2 - \kappa_1)/100$. Recall that $A$ is a subset of $P_\gamma$ with a positive relatively upper density. Let

$$\delta_0 = \limsup_{x \to +\infty} \frac{|A \cap [1, x]|}{|P_\gamma \cap [1, x]|}$$

be the relatively upper density of $A$. Then we may choose a sufficiently large $X$ such that

$$|A \cap [(\kappa_1 + \kappa_0)X, (\kappa_2 - \kappa_0)X| \geq \frac{4\delta_0(\kappa_2 - \kappa_1)}{5} \cdot \frac{X^\gamma}{\log X}. \quad (3.1)$$

Let $N$ be a prime lying in $[(1 - \kappa_0)X/W, (1 + \kappa_0)X/W]$. By the prime number theorem, such prime $N$ always exists. Clearly

$$[\kappa_1WN, \kappa_2WN] \supseteq [(\kappa_1 + \kappa_0)X, (\kappa_2 - \kappa_0)X].$$

By the pigeonhole principle, there exists $1 \leq b_0 \leq W$ with $(b_0, W) = 1$ such that

$$|\{Wn + b_0 \in A : \kappa_1N \leq n \leq \kappa_2N\}| \geq \frac{1}{\varphi(W)} \cdot \frac{\delta_0(\kappa_2 - \kappa_1)}{2} \cdot \frac{(WN)^\gamma}{\log N}. \quad (3.1)$$

Let

$$\eta_0 = 2^\gamma \cdot (\kappa_1WN)^{\gamma - 1}.$$ 

For any $n \in [\kappa_1N, \kappa_2N]$, we have

$$(Wn + b_0 + 1)^\gamma - (Wn + b_0)^\gamma \leq \gamma \cdot (Wn + b_0)^{\gamma - 1} \leq \eta_0.$$
On the other hand, clearly $Wn + b_0 \in \mathbb{N}_{\frac{1}{\gamma}}^+$ if and only if
\[ [(Wn + b_0 + 1)\gamma] > [(Wn + b_0)\gamma] \quad \text{or} \quad (Wn + b_0)\gamma \in \mathbb{N}. \]
Clearly $[(Wn + b_0 + 1)\gamma] > [(Wn + b_0)\gamma]$ is equivalent to
\[ (Wn + b_0 + 1)\gamma - (Wn + b_0)\gamma \geq 1 - \{(Wn + b_0)\gamma\}. \]
So if $Wn + b_0 \in \mathbb{N}_{\frac{1}{\gamma}}^+$, then
\[ 1 - \{(Wn + b_0)\gamma\} \leq \eta_0 \quad \text{or} \quad \{(Wn + b_0)\gamma\} = 0. \]

Let
\[ h_0 = 2^{m-1}m, \quad k_0 = 2m, \]
\[ \delta_0 = \frac{1}{3}\left(\gamma + \frac{1}{2^{2m+2}m} - 1\right), \]
and $r_0$ be the least positive integer such that
\[ \left(r_0 - \frac{h + 2}{1 - \gamma + \delta_0}\right) \cdot (1 - \gamma + \delta_0) > r_0(1 - \gamma) \]
for each $1 \leq h \leq h_0$. According to [22, Lemma 12 of Chapter I], there exists a smooth function $\rho(t)$ with the period 1 such that
\begin{enumerate}[(i)]
  \item $0 \leq \rho(t) \leq 1$ for any $t$ and
  \[ \rho(t) = \begin{cases} 
  1, & \text{if } 1 - \eta_0 \leq t \leq 1, \\
  0, & \text{if } \eta_0 \leq t \leq 1 - 2\eta_0, 
  \end{cases} \]  
  \begin{equation} \tag{3.2} \end{equation}
\item
  \[ \rho(t) = 2\eta_0 + \sum_{|j| \geq 1} \alpha_j e(jt), \]  
  \begin{equation} \tag{3.3} \end{equation}
where
  \[ \alpha_j \ll r_0 \min \left\{ \eta_0, \frac{1}{|j|}, \frac{1}{\eta_0 |j|^{r_0+1}} \right\}. \]  
  \begin{equation} \tag{3.4} \end{equation}
\end{enumerate}
Thus for any $n \in [\kappa_1 N, \kappa_2 N]$, $Wn + b_0 \in \mathbb{N}_{\frac{1}{\gamma}}^+$ implies that
\[ \rho((Wn + b_0)\gamma) = 1. \]

Let
\[ R = N^{\delta_0}. \]
Define
\[ \Lambda_R(n) := \sum_{\substack{d|n \\text{d} \leq R \atop d \leq R}} \mu(d) \log \frac{R}{d}. \]  
\begin{equation} \tag{3.5} \end{equation}
Clearly if \( n > R \) is a prime, then \( \Lambda_R(n) = \log R \). Let
\[
\varpi(n) = \frac{\Lambda_R(n)^2}{\log R} \cdot \frac{\rho(n)}{2\eta_0},
\]
and define
\[
\nu(n) := \begin{cases} 
\frac{\phi(W)}{W} \cdot \varpi(Wn + b_0), & \text{if } n \in [\kappa_1 N, \kappa_2 N], \\
1, & \text{otherwise}.
\end{cases}
\]
Set
\[
f_0(n) = \begin{cases} 
\frac{\delta_0}{3\eta_0} \cdot \lambda_{W,b_0}(n), & \text{if } n \in [\kappa_1 N, \kappa_2 N] \text{ and } Wn + b_0 \in \mathcal{A}, \\
0, & \text{otherwise},
\end{cases}
\]
where \( \lambda_{W,b_0} \) is the one given in (2.5). If \( n \in [\kappa_1 N, \kappa_2 N] \) and \( Wn + b_0 \in P_\gamma \), then
\[
f_0(n) \leq \frac{\delta_0}{3\eta_0} \cdot \frac{\phi(W)}{W} \cdot \log(\kappa_2 WN + W) \leq \frac{\phi(W)}{2\eta_0 W} \cdot \log R = \nu(n).
\]
That is, we always have
\[
0 \leq f_0(n) \leq \nu(n)
\]
for every \( 1 \leq n \leq N \).

In view of (3.1),
\[
\sum_{n \in [\kappa_1 N, \kappa_2 N], Wn + b_0 \in \mathcal{A}} f_0(n) \geq \frac{\delta_0(\kappa_2 - \kappa_1)}{2\phi(W)} \cdot \frac{(WN)^\gamma}{\log N} \cdot \frac{\delta_0}{3\eta_0} \cdot \frac{\phi(W)}{W} \cdot \log(\kappa_2 WN + W) \geq \frac{\delta_0(\kappa_2 - \kappa_1)}{13\gamma\kappa_1^{-1}} \cdot N,
\]
by recalling that \( \eta_0 = 2\gamma \cdot (\kappa_1 WN)^{\gamma - 1} \). Hence by Theorem 2.2 if \( \nu \) obeys the \( m \)-linear forms condition, then
\[
\sum_{x, r \in [1, N]} f_0(x) f_0(x + r) \cdots f_0(x + (m - 1)r) \geq c_{m, \delta_0} N^2
\]
for some constant \( c_{m, \delta_0} > 0 \) only depending on \( m \) and \( \delta_0 \). By (3.9), we have \( f_0(x) \leq \log R \cdot \phi(W)/(2\eta_0 W) \). According to the discussions after (2.6), there exist at least
\[
\frac{c_{m, \delta_0} N^2}{2} \cdot \left( \frac{2\eta_0 W}{\phi(W) \log R} \right)^m - (\kappa_2 - \kappa_1) N
\]
pairs of \( (x, r) \) with \( 1 \leq x, r < N \) such that \( Wx + b_0, \ldots, W(x + (m - 1)r) + b_0 \) form a non-trivial arithmetic progression in \( \{ p \in \mathcal{A} : p \equiv b_0 \pmod{W} \} \). Thus Theorem 1.1 is concluded.

Our remainder task is to verify the \((2^{m-1}m, 2m, m)\)-linear forms condition for the measure \( \nu \). In the next section, we shall propose a Goldston-Yıldırım-type estimation for \( \nu \), which evidently implies the \((2^{m-1}m, 2m, m)\)-linear forms condition.
4. The Goldston-Yıldırım-type estimation

Suppose that $1 \leq h \leq h_0$ and $1 \leq k \leq k_0$. Let
\[
\psi_i(x_1, \ldots, x_k) = a_{i1}x_1 + \cdots + a_{ik}x_k, \quad 1 \leq i \leq h,
\]
where $a_{ij} \in \mathbb{Z}$ and $|a_{ij}| \leq m$. Further, suppose that $(a_{i1}, \ldots, a_{ik})$ and $(a_{j1}, \ldots, a_{jk})$ are linearly independent for any $1 \leq i < j \leq h$. Below, for convenience, we write $\vec{x} = (x_1, \ldots, x_k)$. Then we have the following Goldston-Yıldırım-type estimation.

**Proposition 4.1.**

\[
\frac{1}{Nk} \sum_{\vec{x} \in \mathbb{Z}_N^k} \prod_{i=1}^h \nu(\psi_i(\vec{x}) + b_i) = 1 + o(1), \tag{4.1}
\]

for any $b_1, \ldots, b_h \in \mathbb{Z}_N$.

In this section, we need to reduce the proof of Proposition 4.1 to the estimations of some exponential sums. We shall follow the the same way of Green and Tao in [8]. Let $Q = \lceil N/\log_4 N \rceil$ and $U = \lfloor N/Q \rfloor$. Let
\[
B_{u_1, \ldots, u_k} = \{(x_1, \ldots, x_k) : u_iQ < x_i \leq (u_i + 1)Q \text{ for each } 1 \leq i \leq k\}
\]
for each $0 \leq u_1, \ldots, u_k \leq U - 1$, and let $\mathcal{B}$ be the set of all those $B_{u_1, \ldots, u_k}$. For any $B \in \mathcal{B}$, we say $B$ is **good** provided that for any $1 \leq i \leq h$, either $\psi_i(B) \subseteq [\kappa_1N, \kappa_2N]$ or $\psi_i(B) \cap [\kappa_1N, \kappa_2N] = \emptyset$. Also, we call $B \in \mathcal{B}$ **bad** if $B$ is not good.

Suppose that $B \in \mathcal{B}$ is good. Let $\mathcal{J}_B = \{1 \leq j \leq h : \psi_j(B) \subseteq [\kappa_1N, \kappa_2N]\}$. In view of (3.7),
\[
\sum_{\vec{x} \in B} \prod_{i=1}^h \nu(\psi_i(\vec{x}) + b_i) = \frac{\phi(W)|\mathcal{J}_B|}{W|\mathcal{J}_B|} \sum_{\vec{x} \in B} \varpi(\psi_j(\vec{x}) + b_j)W + b_0). \tag{4.2}
\]

On the other hand, according to Green and Tao’s discussions in [8] Page 528, the number of all bad $B \in \mathcal{B}$ is $O(U^{k-1})$. Hence
\[
\sum_{B \text{ is bad}} \sum_{\vec{x} \in B} \prod_{i=1}^h \nu(\psi_i(\vec{x}) + b_i) + \sum_{\vec{x} \in \mathbb{Z}_N^k \setminus (\bigcup_{B \in \mathcal{B}} B)} \prod_{i=1}^h \nu(\psi_i(\vec{x}) + b_i)
\]
\[
= O\left(U^{k-1} \max_{I_1, \ldots, I_k \subseteq [1, N]} \sum_{\vec{x} \in I_1 \times \cdots \times I_k} \prod_{i=1}^h \nu(\psi_i(\vec{x})) \right).
\]

Note that
\[
\nu(n) \leq 1 + \frac{\phi(W)}{W} \cdot \varpi(Wn + b_0)
\]
for any \( n \in \mathbb{Z}_N \). Therefore, according to (4.2), Proposition 4.1 immediately follows from the estimation
\[
\phi(W)^{|J|} \frac{1}{W^{|J|}Q^k} \sum_{\vec{x} \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_k} \varpi \left( \psi_j(\vec{x})W + b_j^* \right) = 1 + o(1) \tag{4.3}
\]
for any \( J \subseteq \{1, \ldots, h\} \) and any \( \mathcal{I}_1, \ldots, \mathcal{I}_k \subseteq [1, N] \) with \( |\mathcal{I}_i| = Q + O(1) \), where
\[
b_j^* = b_jW + b_0.
\]
Since \( h \) is an arbitrary positive integer not greater than \( h_0 \), it suffices to show that
\[
\phi(W)^h \frac{1}{W^h Q^k} \sum_{\vec{x} \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_k} \prod_{j=1}^h \varpi \left( \psi_j(\vec{x})W + b_j^* \right) = 1 + o(1). \tag{4.4}
\]
Let us turn to the proof of (4.4). Let
\[
H_1 = N^{1-\gamma+\delta_0}.
\]
Recall that in view of (3.3),
\[
\varpi(n) = \frac{\Lambda_R(n)^2}{\log R^2} \cdot \frac{1}{2\eta_0} \left( 2\eta_0 + \sum_{|s| \geq 1} \alpha_s e(s n^\gamma) \right).
\]
And by (3.4),
\[
\sum_{|s| \geq H_1} |\alpha_s| \ll \tau_0 \sum_{|s| \geq H_1} \frac{1}{r_0^{|s|} |s|^{r_0+1}} \ll \eta_0^{-r_0} H_1^{-r_0 + \frac{h+2}{1-\gamma+\delta_0}} \sum_{|s| \geq H_1} |s|^{-\frac{h+2}{1-\gamma+\delta_0} - 1}
\ll \tau_0 H_1^{-\frac{h+2}{1-\gamma+\delta_0}} = N^{-h-2}.
\]
Hence for any intervals \( \mathcal{I}_1, \ldots, \mathcal{I}_k \subseteq [1, N] \) with \( \mathcal{I}_i = Q + O(1) \), we have
\[
\sum_{\vec{x} \in \mathcal{I}} \prod_{j=1}^h \varpi \left( \psi_j(\vec{x})W + b_j^* \right)
= \sum_{\vec{x} \in \mathcal{I}} \frac{\Lambda_R(\psi_j(\vec{x})W + b_j^*)^2}{\log R} \left( 1 + \frac{1}{2\eta_0} \sum_{1 \leq |s| \leq H_1} \alpha_s e(s \cdot (\psi_j(\vec{x})W + b_j^*)^\gamma) \right) + O(N^{-1}),
\]
for any \( n \in \mathbb{Z}_N \). Therefore, according to (4.2), Proposition 4.1 immediately follows from the estimation
\[
\phi(W)^{|J|} \frac{1}{W^{|J|}Q^k} \sum_{\vec{x} \in \mathcal{I}_1 \times \cdots \times \mathcal{I}_k} \varpi \left( \psi_j(\vec{x})W + b_j^* \right) = 1 + o(1) \tag{4.3}
\]
for any \( J \subseteq \{1, \ldots, h\} \) and any \( \mathcal{I}_1, \ldots, \mathcal{I}_k \subseteq [1, N] \) with \( |\mathcal{I}_i| = Q + O(1) \), where
\[
b_j^* = b_jW + b_0.
\]
where \( \mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_k \). Now

\[
\sum_{\bar{x} \in \mathcal{I}} \prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} \left(1 + \frac{1}{2\eta_0} \sum_{1 \leq |s| \leq H_1} \alpha_s e\left(s \cdot (\psi_j(\bar{x})W+b_j^*)^\gamma\right)\right)
\]

\[
= \sum_{\bar{x} \in \mathcal{I}} \prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} \sum_{I \subseteq \{1,\ldots,h\}} \prod_{i \in I} \left(\sum_{1 \leq |s_i| \leq H_1} \alpha_s e\left(s_i \cdot (\psi_i(\bar{x})W+b_i^*)^\gamma\right)\right)
\]

\[
= \sum_{I \subseteq \{1,\ldots,h\}} \prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} \sum_{I \subseteq \{1,\ldots,h\}} \prod_{i \in I} \frac{\alpha_{s_i}}{2\eta_0} \sum_{\bar{x} \in \mathcal{I}} e\left(\sum_{i \in I} s_i \cdot (\psi_i(\bar{x})W+b_i^*)^\gamma\right).
\]

Notice that Green and Tao [8, Proposition 9.5] had proven that

\[
\frac{\phi(W)^h}{W^h} \sum_{\bar{x} \in \mathcal{I}} \prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} = (1 + o(1)) \prod_{j=1}^{h} |\mathcal{I}_j|.
\]

(4.5)

And we have those \( \alpha_{s_i} \ll_{\eta_0} \eta_0 \) by (3.1). Hence it suffices to show that

\[
\prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} \sum_{\bar{x} \in \mathcal{I}} e\left(\sum_{i \in I} s_i \cdot (\psi_i(\bar{x})W+b_i^*)^\gamma\right) = o\left(\frac{Q^k}{H_1^{\ell I}} \cdot \frac{W^h}{\phi(W)^h}\right)
\]

(4.6)

for any \( \emptyset \neq I \subseteq \{1,\ldots,h\} \) and those \( s_i \) with \( 1 \leq |s_i| \leq H_1 \). Clearly in view of (3.3),

\[
\prod_{j=1}^{h} \frac{\Lambda_R(\psi_j(\bar{x})W+b_j^*)^2}{\log R} \sum_{\bar{x} \in \mathcal{I}} e\left(\sum_{i \in I} s_i \cdot (\psi_i(\bar{x})W+b_i^*)^\gamma\right)
\]

\[
= \frac{1}{(\log R)^{2h}} \sum_{d_1,\ldots,d_h,e_1,\ldots,e_h \leq R} \prod_{j=1}^{h} \mu(d_j)\mu(e_j) \log R \frac{R}{d_j} \frac{\log R}{e_j}
\]

\[
\cdot \sum_{[d_j,e_j]|\psi_j(\bar{x})W+b_j^*} e\left(\sum_{i \in I} s_i \cdot (\psi_i(\bar{x})W+b_i^*)^\gamma\right).
\]

Fix \( 1 \leq d_1,\ldots,d_h,e_1,\ldots,e_h \leq R \) with \( (d_je_j,W) = 1 \) for \( 1 \leq j \leq h \). Let \( D_j = [d_j,e_j] \) for \( 1 \leq j \leq h \) and let \( D = [D_1,\ldots,D_h] \). Let

\[
\mathcal{I}_W = \{(x_1,\ldots,x_k) : W a_i \leq x_i \leq W b_i \text{ for } 1 \leq i \leq k\}
\]
provided that \( T = [a_1, b_1] \times \cdots \times [a_k, b_k] \). Then since \( \psi_i(x_1, \ldots, x_k)W = \psi_i(x_1W, \ldots, x_kW) \),

\[
\sum_{\bar{x} \in \overline{T}} e\left( \sum_{i \in I} s_i \cdot (\psi_i(\bar{x})W + b_i^*)^\gamma \right)
\]

for any \( 1 \leq j \leq h \)

\[
= \sum_{\bar{y} = \overline{(y_1, \ldots, y_k)}} e\left( \sum_{i \in I} s_i \cdot (\psi_i(\bar{y}) + b_i^*)^\gamma \right)
\]

\[
\frac{1}{W^k D^h} \sum_{0 \leq u_1, \ldots, u_h < D} \sum_{\bar{y} \in \overline{T} \setminus \overline{W}} e\left( \sum_{i \in I} s_i (\psi_i(\bar{y}) + b_i^*)^\gamma + \sum_{j=1}^h \frac{(\psi_j(\bar{y}) + b_j^*)u_j}{D} + \sum_{j=1}^k \frac{y_j v_j}{W} \right).
\]

So we only need to show that

\[
\sum_{\bar{y} \in \overline{T} \setminus \overline{W}} e\left( \sum_{i \in I} s_i (\psi_i(\bar{y}) + b_i^*)^\gamma + \sum_{j=1}^h \frac{(\psi_j(\bar{y}) + b_j^*)u_j}{D} + \sum_{j=1}^k \frac{y_j v_j}{W} \right) = o\left( \frac{Q^k}{H^h R^{2h}} \right). \tag{4.7}
\]

However, in (4.7), it is difficult to give a suitable lower bound for the second derivatives of the sum in \( e(\cdot) \), since perhaps some \( s_i \) are positive and the other \( s_i \) are negative. That is, we can’t directly apply the classical van Corput theorem to (4.7). There are two auxiliary lemmas in the next section, which are the key ingredients of our proof.

5. Two auxiliary lemmas

Lemma 5.1. Let \( A = (a_{ij})_{1 \leq i \leq h} \) be a \( h \times k \) matrix with integral coefficients and let \( M = \max_{1 \leq j \leq k} |a_{ij}| \). Assume that each two lines of \( A \) are linearly independent.

Then there exists a non-singular matrix \( T \in \mathbb{Z}^{k \times k} \) such that \( |c_{11}|, |c_{21}|, \ldots, |c_{h1}| \) are distinct positive integers bounded by \( h^4 M \), where those \( c_{ij} \) are given by

\[
(c_{ij})_{1 \leq i \leq h} = (a_{ij})_{1 \leq i \leq h} T.
\]

Proof. We shall obtain the matrix \( (c_{ij}) \) via a sequence of column operations on \( (a_{ij}) \). First, we need to get a new matrix \( (b_{ij})_{1 \leq i \leq h} \) such that \( b_{11}, \ldots, b_{h1} \) are all non-zero. Assume that \( a_{i0} = 0 \). Since \( (a_{i1}, \ldots, a_{i0}) \) is not a zero vector, there exists \( 2 \leq f_0 \leq k \) such that \( a_{i0} \neq 0 \). Note that \( |\{1 \leq i \leq h : a_{i1} \neq 0\}| \leq h - 1 \) now. By the pigeonhole role, we may choose \( \theta \in \{1, 2, \ldots, h\} \) such that

\[
a_{i1} + \theta \cdot a_{ij0} \neq 0
\]
for each $1 \leq i \leq h$ with $a_{i1} \neq 0$. Then the first element of the $i_0$-th line of the new matrix

$$
\begin{pmatrix}
  a_{11} + \theta \cdot a_{1j_0} & a_{12} & \cdots & a_{1k} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{h1} + \theta \cdot a_{hj_0} & a_{h2} & \cdots & a_{hk}
\end{pmatrix}
$$

is non-zero. And for each $1 \leq i \leq h$ with $a_{i1} \neq 0$, the first element of the $i$-th line of the above new matrix is still non-zero. Continuing the above process at most $h$ times, we obtain a matrix $(b_{ij})$ with $b_{11}, \ldots, b_{h1} \neq 0$.

Let us turn to the matrix $(b_{ij})$. We shall apply an operation to $(b_{ij})$ such that the number

$$
|\{(s, t) : 1 \leq s < t \leq h, |b_{s1}| = |b_{t1}|\}|
$$

can be reduced by at least 1. Assume that $|b_{s01}| = |b_{t01}| \neq 0$ for some distinct $s_0, t_0$. Since $(b_{s01}, \ldots, b_{sok})$ and $(b_{t01}, \ldots, b_{t0k})$ are also linearly independent, we may choose $2 \leq t_0 \leq k$ such that

$$
\begin{cases}
  b_{s0t_0} \neq b_{t0t_0}, & \text{if } b_{s01} = b_{t01}, \\
  b_{s0t_0} \neq -b_{t0t_0}, & \text{if } b_{s01} = -b_{t01}.
\end{cases}
$$

Clearly

$$
|b_{s01} + \theta \cdot b_{s0t_0}| = |b_{t01} + \theta \cdot b_{t0t_0}|
$$

implies that either

$$
b_{s01} - b_{t01} = \theta \cdot (b_{t0t_0} - b_{s0t_0}),
$$

or

$$
b_{s01} + b_{t01} = -\theta \cdot (b_{t0t_0} + b_{s0t_0}).
$$

So there exists at most one $\theta \neq 0$ such that

$$
|b_{s01} + \theta \cdot b_{s0t_0}| = |b_{t01} + \theta \cdot b_{t0t_0}|.
$$

Hence, we may choose one $\theta$ such that

$$
|b_{s01} + \theta \cdot b_{s0t_0}| \cdot |b_{t01} + \theta \cdot b_{t0t_0}| > 0, \quad |b_{s01} + \theta \cdot b_{s0t_0}| \neq |b_{t01} + \theta \cdot b_{t0t_0}|.
$$

On the other hand, if $s \neq t$ and $|b_{s1}| \neq |b_{t1}|$, then

$$
|b_{s1} + \theta \cdot b_{sl0}| = |b_{t1} + \theta \cdot b_{tl0}|
$$

implies that either

$$
b_{s1} - b_{t1} = \theta \cdot (b_{tl0} - b_{sl0}),
$$

or

$$
b_{s1} + b_{t1} = -\theta \cdot (b_{tl0} + b_{sl0}).
$$

Hence, there exist at most four integers $\theta$ such that either

$$
|b_{s1} + \theta \cdot b_{sl0}| = |b_{t1} + \theta \cdot b_{tl0}|,
$$

or

$$
b_{s1} + \theta \cdot b_{sl0} = 0,$$
or
\[ b_{t_1} + \theta \cdot b_{t_0} = 0. \]

Thus by the pigeonhole role, we may choose \( \theta \in \{1, 2, \ldots, 2h(h - 1)\} \) such that

(i) \( |b_{s_1} + \theta \cdot b_{s_0}| \) and \( |b_{t_1} + \theta \cdot b_{t_0}| \) are distinct positive integers;

(ii) \( |b_{s_1} + \theta \cdot b_{s_0}| \) and \( |b_{t_1} + \theta \cdot b_{t_0}| \) are distinct positive integers, provided that \( |b_{s_1}| \) and \( |b_{t_1}| \) are distinct positive integers.

Transfer the matrix \((b_{ij})\) to a new matrix
\[
\left(\begin{array}{cccc}
  b_{11} + \theta \cdot b_{10} & b_{12} & \cdots & b_{1k} \\
  \vdots & \vdots & \ddots & \vdots \\
  b_{h1} + \theta \cdot b_{h0} & b_{h2} & \cdots & b_{hk}
  \end{array}\right).
\]

Repeating such a process at most \( h(h - 1)/2 \) times, we may obtain the expected new matrix \((c_{ij})\).

**Lemma 5.2.** Suppose that \( M \geq 2 \) and let \( \psi_i(x) = \alpha_i x + \beta_i, \ 1 \leq i \leq h \), be some linear functions with \( \alpha_i \in \mathbb{Z} \) satisfying that
\[ 1 \leq |\alpha_1| < |\alpha_2| < \cdots < |\alpha_h| \leq M. \]

Let \( \mathcal{I} \subseteq \mathbb{N} \) be an interval of integers. Suppose that
\[ N \leq \psi_i(x) \leq \omega N \]
for any \( x \in \mathcal{I} \) and any \( 1 \leq i \leq h \), where
\[ \omega = 1 + \frac{1}{2(M - 1)}. \]

Let \( \theta_1, \ldots, \theta_h \in \mathbb{R} \) and let
\[ F(x) = \theta_1 \psi_1(x)^\gamma + \cdots + \theta_h \psi_h(x)^\gamma. \]

Then there exists
\[ 1 \leq r \leq 3M^2 \log M \log h \cdot (1 + M \log M)^{h-1} \]
such that for every \( x \in \mathcal{I} \),
\[ \frac{\omega^{\gamma - r} |\alpha_h|^r}{2h^{3(1+M \log M)^r}} \cdot \max_{1 \leq i \leq h} \{|\theta_i|\} \leq \frac{|F^{(r)}(x)|}{|\gamma| r \gamma^{r \gamma - r}} \leq \frac{3|\alpha_h|^r}{2} \cdot \max_{1 \leq i \leq h} \{|\theta_i|\}, \]
where
\[ (\gamma)_r = \gamma(\gamma - 1) \cdots (\gamma - r + 1). \]
Proof. First, we shall give some $L_0, L_1, \ldots, L_{h-1} > 0$ and $R_1, R_2 \ldots, R_h \in \mathbb{N}$. Let $L_0 = 1$ and $R_1 = 1$. For $1 \leq j \leq h$, assume that $L_0, \ldots, L_{j-1}$ and $R_1, \ldots, R_j$ have been given. Since $\alpha_1, \ldots, \alpha_h$ are all non-zero integers lying in $[-M, M]$, evidently $|\omega \alpha_i| < |\alpha_{i+1}|$ for each $1 \leq i < h$. Let

$$L_j = \frac{2h}{\omega_\gamma} \left( \frac{|\omega| \alpha_h}{|\alpha_j|} \right)^{R_j}$$

(5.1)

and let $R_{j+1}$ be the least positive integer such that

$$\left( \frac{|\alpha_{j+1}|}{\omega |\alpha_j|} \right)^{R_{j+1}} \geq \frac{2h}{\omega_\gamma} \cdot L_1 \cdots L_j.$$  

(5.2)

Let

$$J = \{1 \leq j < h : |\theta_j| \geq L_j \cdot \max_{i > j} |\theta_i| \}.$$  

First, suppose that $J$ is non-empty. Let $j_0$ be the least element of $J$ and let $r = R_{j_0}$.

For any $x \in \mathcal{I}$, evidently

$$\left| \frac{F^{(r)}(x)}{|(\gamma)_r|} \right| = \left| \sum_{j=1}^{h} \alpha_j^r \theta_j \cdot \psi_j(x)^{\gamma-r} \right| \geq |\alpha_{j_0}|^r |\theta_{j_0}| \cdot \omega^{\gamma-r} N^{\gamma-r} - N^{\gamma-r} \sum_{1 \leq i \leq h, i \neq j_0} |\alpha_i|^r |\theta_i|,$$

(5.3)

by noting that $\psi_i(x) \in [N, \omega N]$ and $\gamma - r < 0$.

We claim that for each $i \neq j_0$,

$$|\alpha_{j_0}|^r |\theta_{j_0}| \geq 2 \omega^{r-\gamma} h \cdot |\alpha_i|^r |\theta_i|.$$  

(5.4)

Since $j_0 = \min J$, we have

$$|\theta_i| < L_i \max_{t > i} |\theta_t|$$

for each $1 \leq i < j_0$, i.e.,

$$\max_{t \geq i} |\theta_i| < L_i \max_{t \geq i+1} |\theta_t|.$$  

If $i < j_0$, then

$$|\theta_i| < L_i \max_{t \geq i+1} |\theta_t| < L_i L_{i+1} \max_{t \geq i+2} |\theta_t| < \cdots < L_i L_{i+1} \cdots L_{j_0-1} \max_{t \geq j_0} |\theta_t|$$

$$= L_i L_{i+1} \cdots L_{j_0-1} |\theta_{j_0}|.$$  

(5.5)

In view of (5.2) and $r = R_{j_0}$,

$$\frac{|\alpha_{j_0}|^r |\theta_{j_0}|}{|\alpha_i|^r |\theta_i|} \geq \frac{|\alpha_{j_0}|^r}{|\alpha_{j_0-1}|^r} \cdot \frac{|\theta_{j_0}|}{|\theta_i|} > 2 \omega^{r-\gamma} h L_1 \cdots L_{j_0-1} \cdot \frac{1}{L_i \cdots L_{j_0-1}} \geq 2 \omega^{r-\gamma} h,$$

by noting that $L_j \geq 1$ for any $j$.  

Suppose that $i > j_0$. Then by (5.1),
\[
\left| \frac{\alpha_i}{\alpha_{j_0}} \right|^r \leq \left| \frac{\alpha_h}{\alpha_{j_0}} \right|^r = \frac{L_{j_0}}{2\omega^{r-\gamma} h}.
\]
So
\[
\left| \frac{\alpha_{j_0}}{\alpha_i} \right|^r \cdot \left| \frac{\theta_{j_0}}{\theta_i} \right| \geq \frac{2\omega^{r-\gamma} h}{L_{j_0}} \cdot L_{j_0} = 2\omega^{r-\gamma} h.
\]
Thus (5.4) is always valid.

Combining (5.3) with (5.4), we get
\[
\left| F(r)(x) \right| \geq \left( \left| \frac{\alpha_h}{\alpha_{j_0}} \right|^r \right) \cdot \left| \frac{\theta_{j_0}}{\theta_i} \right| \cdot 2\omega^{r-\gamma} h \cdot \max_{1 \leq i \leq h} \{ \left| \theta_i \right| \}.
\]

On the other hand, by (5.5), clearly we have
\[
\left| \theta_{j_0} \right| \geq \frac{1}{L_1 \cdots L_{h-1}} \cdot \max_{1 \leq i \leq h} \{ \left| \theta_i \right| \}.
\]
Hence
\[
\frac{\left| F(r)(x) \right|}{\left( \left| \gamma \right| \right)_r} \leq \frac{3\left| \alpha_{j_0} \right|^r \left| \theta_{j_0} \right|}{2} \cdot N^{\gamma-r}.
\]

Next, suppose that $J$ is empty. For any $1 \leq i \leq h - 1$, in view of (5.5), similarly we have
\[
\left| \theta_i \right| < L_i \max_{1 \leq i+1} \left| \theta_i \right| < L_i L_{i+1} \max_{1 \leq i+2} \left| \theta_i \right| < \cdots < L_i L_{i+1} \cdots L_{h-1} \left| \theta_h \right|.
\]
Letting $r = R_h$, we get
\[
\frac{\left| \alpha_i \right|^r \left| \theta_h \right|}{\left| \alpha_i \right|^r \left| \theta_i \right|} \geq \frac{2\omega^{r-\gamma} h L_1 \cdots L_{h-1} \cdot \frac{1}{L_1 \cdots L_{h-1}}}{\left| \alpha_{j_0} \right|^r \left| \theta_{j_0} \right|} \geq 2\omega^{r-\gamma} h.
\]
It follows that
\[
\frac{\left| F(r)(x) \right|}{\left( \left| \gamma \right| \right)_r} \geq \frac{\left| \alpha_h \right|^r \left| \theta_h \right|}{\left| \alpha_h \right|^r \left| \theta_i \right|} \cdot \omega^{r-\gamma} h \cdot N^{\gamma-r} - N^{\gamma-r} \sum_{i=1}^{h-1} \left| \alpha_i \right|^r \left| \theta_i \right| \geq \frac{\left| \alpha_h \right|^r \left| \theta_h \right|}{2} \cdot \omega^{r-\gamma} h \cdot N^{\gamma-r} \geq \frac{\left| \alpha_h \right|^r \omega^{r-\gamma} h \cdot N^{\gamma-r}}{2L_1 \cdots L_{h-1}} \cdot \max_{1 \leq i \leq h} \{ \left| \theta_i \right| \} \cdot N^{\gamma-r},
\]
and
\[
\frac{\left| F(r)(x) \right|}{\left( \left| \gamma \right| \right)_r} \leq \frac{3\left| \alpha_h \right|^r \left| \theta_h \right|}{2} \cdot N^{\gamma-r} + N^{\gamma-r} \sum_{i=1}^{h-1} \left| \alpha_i \right|^r \left| \theta_i \right| \leq \frac{3\left| \alpha_h \right|^r \left| \theta_h \right|}{2} \cdot N^{\gamma-r}.
\]
Finally, we need to give an upper bound for $R_h$. Clearly
\[
\log L_j = \log(2h^{-\gamma}) + R_j \log \frac{\omega|\alpha_h|}{|\alpha_j|}
\]
\[
\leq \log(2h^{-\gamma}) + \log \frac{\omega|\alpha_h|}{|\alpha_j|} \cdot \left( \log \frac{\omega|\alpha_h|}{|\alpha_{j-1}|} \right)^{-1} \cdot \left( \log(3h^{-\gamma}) + \sum_{i=1}^{j-1} \log L_i \right)
\]
\[
\leq \log(2h^{-\gamma}) + \log M \cdot M \cdot \left( \log(3h^{-\gamma}) + \sum_{i=1}^{j-1} \log L_i \right). \quad (5.6)
\]
We claim that
\[
\log L_j \leq 3M \log M \log h \cdot (1 + M \log M)^j \quad (5.7)
\]
for each $0 \leq j \leq h$. In fact, assume that (5.7) holds for $L_1, \ldots, L_{j-1}$. Then by (5.6),
\[
\log L_j \leq \log(2h^{-\gamma}) + M \log M \left( \log(3h^{-\gamma}) + 3M \log M \log h \cdot (1 + M \log M)^{j-1} \right)
\]
\[
\leq M \log M \cdot \log h (1 + M \log M)^j,
\]
since it is easy to verify
\[
\log(2h^{-\gamma}) + M \log M \log(3h^{-\gamma}) \leq 3M \log M \log h.
\]
So
\[
R_h \leq \log L_{h-1} \cdot \left( \log \frac{\omega|\alpha_h|}{|\alpha_{h-1}|} \right)^{-1} \leq 3M^2 \log M \log h \cdot (1 + M \log M)^{h-1},
\]
and
\[
\log L_1 + \cdots + \log L_{h-1} \leq 3 \log h \cdot (1 + M \log M)^h.
\]

6. Proof of Theorem 1.1

In this section, we shall complete the proof Theorem 1.1. Let $\nu$ be the pseudorandom measure constructed in (3.7). According to Theorem 2.2, we only need to verify that $\nu$ obeys the $(2^{m-1}m, 2m, m)$-linear forms condition.

Recall that $h_0 = 2^{m-1}m$ and $k_0 = 2m$. Suppose that $1 \leq h \leq h_0$ and $1 \leq k \leq k_0$. Suppose that
\[
\psi_i(\vec{x}) = a_{i1}x_1 + \cdots + a_{ik_0}x_k, \quad 1 \leq i \leq h
\]
with $|a_{ij}| \leq m$, and $b_1, \ldots, b_i \in \mathbb{Z}_N$. As we has mentioned, it suffices to show that
\[
\frac{1}{N^k} \sum_{\vec{x} \in \mathbb{Z}_N^k} \prod_{i=1}^{h} \nu(\psi_i(\vec{x}) + b_i) = 1 + o(1). \quad (6.1)
\]
By Lemma 5.1 there exists a non-singular matrix $T \in \mathbb{Z}_N^{k \times k}$ such that $(a_{ij}^*)_{1 \leq i < h} = (a_{ij})_{1 \leq i < h} T$ satisfies $1 \leq |a_{11}^*| < \ldots < |a_{h1}^*| \leq h^4 m$. Since $T$ is non-singular,

$$
\frac{1}{N_k} \sum_{\vec{x} \in \mathbb{Z}_N^k} \prod_{i=1}^h \nu(\psi_i(\vec{x})) = \frac{1}{N_k} \sum_{\vec{x} \in \mathbb{Z}_N^k} \prod_{i=1}^h \nu(\psi_i(\vec{x}T)) = \frac{1}{N_k} \sum_{\vec{x} \in \mathbb{Z}_N^k} \prod_{i=1}^h \nu(a_{i1}^* x_1 + \cdots + a_{ih}^* x_h + b_i).
$$

Therefore without loss of generality, we may assume that those linear functions $\psi_i(\vec{x}) = a_{i1} x_1 + \cdots + a_{ih} x_h$ in (6.1) satisfy

$$
1 \leq |a_{11}| < |a_{21}| < \ldots < |a_{h1}| \leq h^4 m.
$$

According to our discussions in Section 3, (6.1) follows from (4.7), i.e.,

$$
\sum_{\vec{x} \in \mathcal{I}} e\left(\sum_{i \in I} s_i (\psi_i(\vec{x}) + b_i^*) \gamma + \sum_{j=1}^h \frac{(\psi_j(\vec{x}) + b_j^*) u_j}{D} + \sum_{j=1}^k \frac{x_j v_j}{W}\right) = o(Q^k H^{-h} R^{-2h}),
$$

where

$$
1 \leq |s_i| \leq H_1, \quad 1 \leq u_j \leq D, \quad 1 \leq v_j \leq W
$$

and $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_k$ with

$$|\mathcal{I}_1|, \ldots, |\mathcal{I}_k| = W Q + O(W).
$$

Below we need the following general form of van der Corput’s theorem [3 Satz 4]:

**Lemma 6.1.** Let $f(x)$ be a smooth function on the interval $[X, X + Y]$. Suppose that $r \geq 2$ and

$$
0 < \lambda \leq |f^{(r)}(x)| \leq \alpha \lambda
$$

on $[X, X + Y]$. Then

$$
\sum_{X \leq n \leq X + Y} e(f(n)) \ll \alpha Y \left(\lambda^{\frac{1}{r-2}} + Y^{-\frac{1}{r-1}} + (Y^r \lambda)^{-\frac{1}{r-1}}\right). \quad (6.3)
$$

Let

$$
F_{x_1, \ldots, x_k}(y) = \sum_{i \in I} s_i (\psi_i(y, x_2, \ldots, x_k) + b_i^*) \gamma + \sum_{j=1}^h \frac{(\psi_j(y, x_2, \ldots, x_k) + b_j^*) u_j}{D} + \sum_{j=1}^k \frac{y_j v_j}{W}.
$$

Clearly

$$
\left| \sum_{\vec{x} \in \mathcal{I}} e\left(F_{x_1, \ldots, x_k}(x_1)\right) \right| \leq \sum_{x_2 \in \mathcal{I}_2, \ldots, x_k \in \mathcal{I}_k} \left| \sum_{y \in \mathcal{I}} e\left(F_{x_1, \ldots, x_k}(y)\right) \right|.
$$
Let

\[ M_0 = h_0^4m. \]

Applying Lemma 5.2 to \( F'_{x_2, \ldots, x_k} \), for any given \( x_2 \in I_2, \ldots, x_k \in I_k \), there exists

\[ 2 \leq r \leq 3M_0^2 \log M_0 \log h_0 \cdot (1 + M_0 \log M_0)^{h_0-1} + 1, \]

such that for any \( y \in I_1 \)

\[ c_1 \Psi N^{\gamma-r} \leq |F'_{x_2, \ldots, x_k}(y)| \leq c_2 \Psi N^{\gamma-r}, \]

where \( c_1, c_2 > 0 \) are two constants only depending on \( m \) and

\[ \Psi = \max_{i \in I} |s_i|. \]

Let \( \lambda = \Psi N^{\gamma-r} \). Since \( \Psi \leq H_1 = N^{1-\gamma+\delta_0} \),

\[ \lambda^{-\frac{1}{2}} \leq \left( H_1 N^{\gamma-r} \right)^{-\frac{1}{2}} = N^{\frac{1}{2} + h_0-2}. \]

Recalling that \( W \gg \log_3 N \) and \( Q = \lceil N/ \log_4 N \rceil \), Let \( Y = WQ \) we have and

\[ (Y^r \lambda)^{-\frac{1}{2}} \leq (N^r \cdot N^{\gamma-r})^{-\frac{1}{2}} = N^{\frac{\gamma-r}{2}}. \]

Using Lemma 6.3 we get that

\[
\sum_{y \in I_1} e \left( F_{x_2, \ldots, x_k}(y) \right) \ll Y \left( \lambda^{-\frac{1}{2}} + Y^{-\frac{1}{2}} + (Y^r \lambda)^{-\frac{1}{2}} \right)
\]

\[ \ll WQ \cdot N^{-2^{m-2} - 3h_0^2 \log M_0 \log h_0 \cdot (1 + M_0 \log M_0)^{h_0-1}}. \]

It is not difficult to check that

\[ 3M_0^2 \log M_0 \log h_0 \cdot (1 + M_0 \log M_0)^{h_0-1} + 3 \leq 2^{4m^2} - \frac{\log h_0}{\log 2} \]

for each \( m \geq 3 \). Hence recalling that \( 1 - r + 3\delta_0 = 2^{-2^{4m^2}} \), we have

\[
\sum_{\substack{x \in I \\ x = (x_1, \ldots, x_k)}} e \left( F_{x_2, \ldots, x_k}(x_1) \right) \ll W^k Q^k \cdot N^{-h_0 2^{1 - 2^{4m^2}}}
\]

\[ \ll Q^k \cdot N^{-h_0(1-\gamma+3\delta_0)} = Q^k \cdot H_1^{-h_0 R^{-2h_0}}. \]

Thus (6.2) is concluded, i.e., the function \( \nu \) really obeys the \((2^{m-1}m, 2m, m)\)-linear forms condition.


References

[1] A. Balog and J. Friedlander, A hybrid of theorems of Vinogradov and Piatetski-Shapiro. Pacific J. Math. 156 (1992), 45–62.
[2] D. Conlon, J. Fox and Y. Zhao, A relative Szemerédi theorem, Geom. Funct. Anal. 25 (2015), 733-762.
[3] J. G. van der Corput, Neue zahlentheoretische Abschätzungen II, Math. Z., 29(1929), 397-426.
[4] D. A. Goldston, J. Pintz and C. Y. Yıldırım, Primes in tuples. I, Ann. of Math, 170(2009), 819-862.
[5] D. A. Goldston and C. Y. Yıldırım, Higher correlations of divisor sums related to primes, I: Triple correlations, Integers 3 (2003), 66pp.
[6] S. W. Graham and G. Kolesnik, Van der Corput’s method of exponential sums. London Mathematical Society Lecture Note Series, 126, Cambridge University Press, Cambridge, 1991.
[7] B. Green, Roth’s theorem in the primes, Ann. of Math. (2) 161 (2005), 1609-1636.
[8] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, Ann. of Math. (2) 167 (2008), 481–547.
[9] B. Green and T. Tao, Linear equations in primes, Ann. of Math. (2) 171 (2010), 1753–1850.
[10] D. R. Heath-Brown, The Piatetski-Shapiro prime number theorem. J. Number Theory, 16(1983), 242-266.
[11] H. Iwaniec, Almost-primes represented by quadratic polynomials, Invent. Math., 47(1978), 171-188.
[12] T.-H. Lê, Green-Tao theorem in function fields, Acta Arith. 147 (2011), 129–152.
[13] T.-H. Lê and J. Wolf, Polynomial configurations in the primes, Int. Math. Res. Not. IMRN 2014, 6448-6473.
[14] M. Mirek, Roth’s theorem in the Piatetski-Shapiro primes, Rev. Mat. Iberoam., 31(2015), 617-656.
[15] I. I. Piatetski-Shapiro, On the distribution of prime numbers in sequences of the form \([f(n)]\), Mat. Sb., 33(1953), 559-566.
[16] J. Pintz, Polignac numbers, conjectures of Erdős on gaps between primes, arithmetic progressions in primes, and the bounded gap conjecture, From arithmetic to zeta-functions, 367–384, Springer, 2016.
[17] J. Rivat and J. Wu, Prime numbers of the form \([n^c]\). Glasg. Math. J. 43 (2001), 237-254.
[18] Y.-C. Sun and H. Pan, The Green-Tao theorem for primes of the form \(x^2+y^2+1\), Monatsh. Math., to appear.
[19] T. Tao, The Gaussian primes contain arbitrarily shaped constellations, J. Anal. Math. 99 (2006), 109-176.
[20] T. Tao and T. Ziegler, The primes contain arbitrarily long polynomial progressions, Acta Math. 201 (2008), 213–305.
[21] T. Tao and T. Ziegler, A multi-dimensional Szemerédi theorem for the primes via a correspondence principle, Israel J. Math. 207 (2015), 203–228.
[22] I. M. Vinogradov, The method of trigonometrical sums in the theory of numbers, translated, revised and annotated by K. F. Roth and Anne Davenport, Interscience Publishers, London and New York, 1954.
[23] B.-B. Zhou, The Chen primes contain arbitrarily long arithmetic progressions, Acta Arith. 138 (2009), 301–315.
School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China
E-mail address: lihz@sjtu.edu.cn

School of Applied Mathematics, Nanjing University of Finance and Economics, Nanjing 210046, People’s Republic of China
E-mail address: haopan79@zoho.com