ARITHMETIC PROGRESSIONS OF PRIMES IN SHORT INTERVALS

CHUNLEI LIU

Abstract. Green and Tao proved that the primes contains arbitrarily long arithmetic progressions. We show that, essentially the same proof leads to the following result: If $N$ is sufficiently large and $M$ is not too small compared with $N$, then the primes in the interval $[N, N + M]$ contains many arithmetic progressions of length $k$.

1. Introduction

Let $N$ be a positive integer going to infinity. We write $o(1)$ for any quantity which tends to zero as $N$ goes to infinity, and write $O(1)$ for any quantity which has a bound independent of $N$. Let $w = w(N) \leq \frac{N}{2} \log \log N$ be any function which tends to infinity with $N$, and let $W := \prod_{p \leq w} p$ be the product of the primes up to $w$. Let $\tilde{\Lambda}$ be the $W$-tricked von Mangoldt function defined by

$$\tilde{\Lambda}(n) := \begin{cases} \frac{\phi(W)}{W} \log(Wn + 1) & \text{when } Wn + 1 \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

Let $M$ be a large prime number. Define $\mathbb{Z}_M := \mathbb{Z}/M\mathbb{Z}$ to be the finite field consisting of residue classes modulo $M$. We always identify $\mathbb{Z}_M$ with the set

$$\{N - M, N - M + 1, \ldots, N - 1\},$$

which is a complete system of representatives modulo $M$.

If $A$ is a finite non-empty set and $f : A \rightarrow \mathbb{R}$ is a function, we write

$$\mathbb{E}(f) := \mathbb{E}(f(x) | x \in A)$$

for the average value of $f$, that is to say

$$\mathbb{E}(f) := \frac{1}{|A|} \sum_{x \in A} f(x).$$

Here, as is usual, we write $|A|$ for the cardinality of the set $A$. More generally, if $P(x)$ is any statement concerning an element of $A$ which is true for at least one $x \in A$, we define

$$\mathbb{E}(f(x) | P(x)) := \frac{\sum_{x \in A : P(x)} f(x)}{|\{x \in A : P(x)\}|}.$$
A famous theorem of Green-Tao in [4] asserts that the prime numbers contain arbitrarily long arithmetic progressions. In this paper we show that the proof of Green-Tao really yields the following theorem.

**Theorem 1.1.** Let \( M \) be a function of \( N \) with values in the set of prime numbers which satisfies \( N^\varepsilon < M \leq N \) for some positive number \( \varepsilon \). Suppose that on the interval \([N + \varepsilon_k M, N + 2\varepsilon_k M]\) the mean value of the \( W \)-tricked von Mongoldt function tends to 1 as \( N \) goes to infinity. Define the function \( f \) on \( \mathbb{Z}_M \) by setting

\[
  f(n) := \begin{cases} 
    k^{-1}2^{-k-5} \Lambda(n) & \text{when } \varepsilon_k M \leq n - N \leq 2\varepsilon_k M \\
    0 & \text{otherwise.}
  \end{cases}
\]

Then there is a positive constant \( c_k \) depending only on \( k \) such that

\[
  \mathbb{E}(f(x)f(x+r)\ldots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_M) \geq c_k - o(1).
\]

From that theorem we see that, for sufficiently large \( N \), there are at least \( b_k M^2 / \log^k N \) arithmetic progressions of length \( k \) consisting of primes in the interval \((WN, W(N+M)]\), where \( b_k \) is a positive constant \( c_k \) depending only on \( k \). According to Green-Tao, we can in fact take \( w \) to be a sufficiently large number independent of \( N \), depending only on \( k \). Then \( W \) will be a constant depending only on \( k \).

**Acknowledgements** The author would like to thank the Morningside Center of Mathematics, Chinese Academy of Sciences for support over several years.

2. THE LINEAR FORMS PROPERTY

In this section we construct a majorant \( \nu \) for \( f \) and prove that \( \nu \) satisfies the linear forms condition.

**Definition 2.1.** Let \( R \) be a parameter (in applications it will be a small power of \( N \)). Define

\[
  \Lambda_R(n) := \sum_{d \mid n, \quad d \leq R} \mu(d) \log(R/d) = \sum_{d \mid n} \mu(d) \log(R/d)_+.
\]

These truncated divisor sums have been studied in several papers, most notably the works of Goldston and Yıldırım [1, 2, 3] concerning the problem of finding small gaps between primes.

**Definition 2.2.** Let \( R := M^{k^{-1}2^{-k-4}} \). We define the function \( \nu : \mathbb{Z}_M \rightarrow \mathbb{R}^+ \) by

\[
  \nu(n) := \begin{cases} 
    \frac{\phi(W) \Lambda_R(Wn+1)^2}{\log R} & \text{when } \varepsilon_k M \leq n - N \leq 2\varepsilon_k M \\
    1 & \text{otherwise}
  \end{cases}
\]

for all \( N + M \leq n < N + M \).

**Lemma 2.3.** Let \( N \) be a sufficiently large integer depending on \( k \). Then the function \( \nu \) is majorant for \( f \) in Theorem 1.1. That is, \( \nu(n) \geq 0 \) for all \( n \in \mathbb{Z}_M \), and \( \nu(n) \geq k^{-1}2^{-k-5} \Lambda(n) \) for all \( N + \varepsilon_k M \leq n \leq N + 2\varepsilon_k M \).

**Proof.** The first claim is trivial. The second claim is also trivial unless \( WN + 1 \) is prime. From definition of \( R \), we see that \( WN + 1 > R \) if \( N \) is sufficiently large. Then the sum over \( d \mid WN + 1, \quad d \leq R \) in (2.1) in fact consists of just the one term \( d = 1 \).
Therefore \( \Lambda_R(Wn + 1) = \log R \), which means that \( \nu(n) = \frac{\phi(W)}{W} \log R \geq k^{-12-k^{-5}}\Lambda(n) \) by construction of \( R \) and \( N \).

**Definition 2.4** (Linear forms condition). Let \( m_0, t_0 \) and \( L_0 \) be small positive integer parameters. Then we say that \( \nu : \mathbb{Z}_M \to \mathbb{R}^+ \) satisfies the \((m_0, t_0, L_0)\)-linear forms condition if the following holds. Let \( m \leq m_0 \) and \( t \leq t_0 \) be arbitrary, and suppose that \((L_{ij})_{1 \leq i \leq m, 1 \leq j \leq t}\) are arbitrary rational numbers with numerator and denominator at most \( L_0 \) in absolute value, and that \( b_i, 1 \leq i \leq m \), are arbitrary elements of \( \mathbb{Z}_M \). For \( 1 \leq i \leq m \), let \( \psi_i : \mathbb{Z}_M^t \to \mathbb{Z}_M \) be the linear forms \( \psi_i(x) = \sum_{j=1}^t L_{ij}x_j + b_i \), where \( x = (x_1, \ldots, x_t) \in \mathbb{Z}_M^t \), and where the rational numbers \( L_{ij} \) are interpreted as elements of \( \mathbb{Z}_M \) in the usual manner (assuming \( M \) is prime and larger than \( L_0 \)). Suppose that as \( i \) ranges over \( 1, \ldots, m \), the \( t \)-tuples \((L_{ij})_{1 \leq j \leq t} \in \mathbb{Q}^t \) are non-zero, and no \( t \)-tuple is a rational multiple of any other. Then we have

\[
\mathbb{E} \left( \nu(\psi_1(x)) \cdots \nu(\psi_m(x)) \mid x \in \mathbb{Z}_M^t \right) = 1 + o_{L_0,m_0,t_0}(1).
\] (2.1)

Note that the rate of decay in the \( o(1) \) term is assumed to be uniform in the choice of \( b_1, \ldots, b_m \).

The following propositions plays a crucial role in proving that \( \nu \) satisfies the linear forms condition.

**Proposition 2.5** (Goldston-Yıldırım). Let \( m, t \) be positive integers. For each \( 1 \leq i \leq m \), let \( \psi_i(x) := \sum_{j=1}^t L_{ij}x_j + b_i \), be linear forms with integer coefficients \( L_{ij} \) such that \(|L_{ij}| \leq \sqrt{w(N)}/2 \) for all \( i = 1, \ldots, m \) and \( j = 1, \ldots, t \). We assume that the \( t \)-tuples \((L_{ij})_{i=1}^t \) are never identically zero, and that no two \( t \)-tuples are rational multiples of each other. Write \( \theta_i := W\psi_i + 1 \). Suppose that \( B \) is a product \( \prod_{i=1}^t I_i \subset \mathbb{R}^t \) of \( t \) intervals \( I_i \), each of which having length at least \( R^{10m} \). Then

\[
\mathbb{E}(\Lambda_R(\theta_1(x))^2 \cdots \Lambda_R(\theta_m(x))^2 \mid x \in B) = (1 + o_{m,t}(1)) \left( \frac{W \log R}{\phi(W)} \right)^m.
\]

**Remarks.** That proposition was stated and proved by Green-Tao in [4], however, according to Green-Tao, it is a straightforward generalisation of [3] Proposition 2.

Before proving the linear forms condition, We show that \( \mathbb{E}(\nu) = 1 + o(1) \).

**Lemma 2.6.** We have \( \mathbb{E}(\nu) = 1 + o(1) \).

**Proof.** Apply Proposition 2.5 with \( m := t := 1 \), \( \psi_1(x_1) := x_1 \) and \( B := [N + \epsilon_k M, N + 2\epsilon_k M] \) (taking \( N \) sufficiently large depending on \( k \), of course). Comparing with Definition 2.2 we thus have

\[
\mathbb{E}(\nu(x) \mid x \in [N + \epsilon_k M, N + 2\epsilon_k M]) = 1 + o(1).
\]

But from the same definition we clearly have

\[
\mathbb{E}(\nu(x) \mid x \in \mathbb{Z}_M \setminus [N + \epsilon_k M, N + 2\epsilon_k M]) = 1;
\]

Combining these two results confirms the lemma.

Now we verify the linear forms condition, which is proven in a similar spirit to the above lemma.
**Proposition 2.7.** The function \( \nu \) satisfies the \((k \cdot 2^{k-1}, 3k - 4, k)\)-linear forms condition.

**Proof.** Let \( \psi_i(x) = \sum_{j=1}^k L_{ij} x_j + b_i \) be linear forms of the type which feature in Definition 2.4. That is to say, we have \( m \leq k \cdot 2^{k-1}, \) \( t \leq 3k - 4, \) the \( L_{ij} \) are rational numbers with numerator and denominator at most \( k \) in absolute value, and none of the \( t \)-tuples \((L_{ij})_{j=1}^t\) is zero or is equal to a rational multiple of any other. We wish to show that

\[
\mathbb{E}(\nu(\psi_1(x)) \ldots \nu(\psi_m(x)) \mid x \in \mathbb{Z}_M^m) = 1 + o(1). \tag{2.2}
\]

We may clear denominators and assume that all the \( L_{ij} \) are integers, at the expense of increasing the bound on \( L_{ij} \) to \(|L_{ij}| \leq (k + 1)!\). Since \( w(N) \) is growing to infinity in \( N \), we may assume that \((k + 1)! < \sqrt{w(N)/2}\) by taking \( N \) sufficiently large. This is required in order to apply Proposition 2.5 as we have stated it.

The two-piece definition of \( \nu \) in Definition 2.2 means that we cannot apply Proposition 2.5 immediately, and we need the following localization argument.

We chop the range of summation in (2.2) into \( Q^t \) almost equal-sized boxes, where \( Q = Q(N) \) is a slowly growing function of \( N \) to be chosen later. Thus let

\[ B_{u_1, \ldots, u_t} = \{ x \in \mathbb{Z}_M^m : x_j \in [N + [u_j M/Q], N + [(u_j + 1) M/Q]], j = 1, \ldots, t \}, \]

where the \( u_j \) are to be considered \((\text{mod} \ Q)\). Observe that up to negligible multiplicative errors of \( 1 + o(1) \) (arising because the boxes do not quite have equal sizes) the left-hand side of (2.2) can be rewritten as

\[
\mathbb{E}(\mathbb{E}(\nu(\psi_1(x)) \ldots \nu(\psi_m(x)) \mid x \in B_{u_1, \ldots, u_t}) | u_1, \ldots, u_t \in \mathbb{Z}_Q).
\]

Call a \( t \)-tuple \((u_1, \ldots, u_t) \in \mathbb{Z}_Q^t \) nice if for every \( 1 \leq i \leq m \), the sets \( \psi_i(B_{u_1, \ldots, u_t}) \) are either completely contained in the interval \([N + \epsilon_i M, N + 2\epsilon_i M]\) or are completely disjoint from this interval. From Proposition 2.5 and Definition 2.2 we observe that

\[
\mathbb{E}(\nu(\psi_1(x)) \ldots \nu(\psi_m(x)) \mid x \in B_{u_1, \ldots, u_t}) = 1 + o_{m, t}(1)
\]

whenever \((u_1, \ldots, u_t) \) is nice, since we can replace each of the \( \nu(\psi_i(x)) \) factors by either \( \frac{\phi(W)}{W \log R} \Lambda_R^2(\theta_i(x)) \) or \( 1 \), and \( M/Q \) will exceed \( R^{10m} \) for \( Q \) sufficiently slowly growing in \( N \), by definition of \( R \) and the upper bound on \( m \). When \((u_1, \ldots, u_t) \) is not nice, then we can crudely bound \( \nu \) by \( 1 + \frac{\phi(W)}{W \log R} \Lambda_R^2(\theta_i(x)) \), multiply out, and apply Proposition 2.5 again to obtain

\[
\mathbb{E}(\nu(\psi_1(x)) \ldots \nu(\psi_m(x)) \mid x \in B_{u_1, \ldots, u_t}) = O_{m, t}(1) + o_{m, t}(1)
\]

We shall shortly show that the proportion of non-nice \( t \)-tuples \((u_1, \ldots, u_t) \) in \( \mathbb{Z}_Q^t \) is at most \( O_{m, t}(1/Q) \), and thus the left-hand side of (2.2) is \( 1 + o_{m, t}(1) + O_{m, t}(1/Q) \), and the claim follows by choosing \( Q \) sufficiently slowly growing in \( N \).

It remains to verify the claim about the proportion of non-nice \( t \)-tuples. Suppose \((u_1, \ldots, u_t) \) is not nice. Then there exists \( 1 \leq i \leq m \) and \( x, x' \in B_{u_1, \ldots, u_t} \) such that \( \psi_i(x) \) lies in the interval \([N + \epsilon_i M, N + 2\epsilon_i M]\), but \( \psi_i(x') \) does not. But from definition of \( B_{u_1, \ldots, u_t} \) and the boundedness of the \( L_{ij} \) we have

\[
\psi_i(x), \psi_i(x') = \sum_{j=1}^t L_{ij} (N + [Mu_j/Q]) + b_i + O_{m, t}(M/Q).
\]
Thus we must have
\[ N + a\epsilon_k M = \sum_{j=1}^{t} L_{ij}(N + \lfloor Mu_j/Q \rfloor) + b_i + O_{m,t}(M/Q) \]
for either \( a = 1 \) or \( a = 2 \). Dividing by \( M/Q \), we obtain
\[ \sum_{j=1}^{t} L_{ij} u_j = (1 - \sum_{j=1}^{t} L_{ij} - b_i)Q/N + a\epsilon_k Q + O_{m,t}(1) \pmod{Q}. \]
Since \((L_{ij})_{j=1}^{t}\) is non-zero, the number of \( t \)-tuples \((u_1, \ldots, u_t)\) which satisfy this equation is at most \( O_{m,t}(Q^{-1}) \). Letting \( a \) and \( i \) vary we thus see that the proportion of non-nice \( t \)-tuples is at most \( O_{m,t}(1/Q) \) as desired (the \( m \) and \( t \) dependence is irrelevant since both are functions of \( k \)).

3. THE CORRELATION PROPERTY

In this section we show that \( \nu \) satisfies the correlation condition.

**Definition 3.1 (Correlation condition).** Let \( m_0 \) be a positive integer parameter. We say that \( \nu : \mathbb{Z}_M \to \mathbb{R}^+ \) satisfies the \( m_0 \)-correlation condition if for every \( 1 < m \leq m_0 \) there exists a weight function \( \tau = \tau_m : \mathbb{Z}_M \to \mathbb{R}^+ \) which obeys the moment conditions
\[ \mathbb{E}(\tau^q) = O_{m,q}(1) \] (3.1)
for all \( 1 \leq q < \infty \) and such that
\[ \mathbb{E}(\nu(x+h_1)\nu(x+h_2)\ldots\nu(x+h_m) \mid x \in \mathbb{Z}_M) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j) \] (3.2)
for all \( h_1, \ldots, h_m \in \mathbb{Z}_M \) (not necessarily distinct).

The following proposition plays a crucial role in proving that \( \nu \) satisfies the correlation condition.

**Proposition 3.2 (Goldston-Yıldırım).** Let \( m \geq 1 \) be an integer, and let \( B \) be an interval of length at least \( R^{10m} \). Suppose that \( h_1, \ldots, h_m \) are distinct integers satisfying \( |h_i| \leq N^2 \) for all \( 1 \leq i \leq m \), and let \( \Delta \) denote the integer
\[ \Delta := \prod_{1 \leq i < j \leq m} |h_i - h_j|. \]
Then
\[ \mathbb{E}(\Lambda_R(W(x_1 + h_1) + 1)^2 \ldots \Lambda_R(W(x_m + h_m) + 1)^2 \mid x \in B) \leq (1 + o_m(1)) \left( \frac{W \log R}{\phi(W)} \right)^m \prod_{p \mid \Delta} (1 + O_m(p^{-1/2})). \] (3.3)

Here and in the sequel, \( p \) is always understood to be prime.

**Remarks.** That proposition was stated and proved by Green-Tao in [4], however, Green-Tao attributed it to Goldston-Yıldırım for reasons similar to Proposition 2.5.

In a short while we will use Proposition 3.2 to show that \( \nu \) satisfies the correlation condition. Prior to that, however, we must look at the average size of the “arithmetic” factor \( \prod_{p \mid \Delta} (1 + O_m(p^{-1/2})) \) appearing in that proposition.
Lemma 3.3. Let \( m \geq 1 \) be a parameter. There is a weight function \( \tau = \tau_m : \mathbb{Z} \to \mathbb{R}^+ \) such that \( \tau(n) \geq 1 \) for all \( n \neq 0 \), and such that for all distinct \( h_1, \ldots, h_j \in [N + \epsilon_k M, N + 2\epsilon_k M] \) we have

\[
\prod_{p \mid \Delta} (1 + O_m(p^{-1/2})) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j),
\]

where \( \Delta \) is defined in Proposition 3.2 and such that \( E(\tau^q(n) \mid 0 < |n| \leq M) = O_{m,q}(1) \) for all \( 0 < q < \infty \).

Proof. We observe that

\[
\prod_{p \mid \Delta} (1 + O_m(p^{-1/2})) \leq \prod_{1 \leq i < j \leq m} \left( \prod_{p \mid h_i - h_j} (1 + p^{-1/2}) \right)^{O_m(1)}.
\]

By the arithmetic mean-geometric mean inequality (absorbing all constants into the \( O_m(1) \) factor) we can thus take \( \tau_m(n) := O_m(1) \prod_{p \mid n} (1 + p^{-1/2})^{O_m(1)} \) for all \( n \neq 0 \). (The value of \( \tau \) at 0 is irrelevant for this lemma since we are taking all the \( h_i \) to be distinct).

To prove the claim, it thus suffices to show that

\[
E\left( \prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq M \right) = O_{m,q}(1) \text{ for all } 0 < q < \infty.
\]

Since \( (1 + p^{-1/2})^{O_m(q)} \) is bounded by \( 1 + p^{-1/4} \) for all but \( O_{m,q}(1) \) many primes \( p \), we have

\[
E\left( \prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq M \right) \leq O_{m,q}(1) E\left( \prod_{p \mid n} (1 + p^{-1/4}) \mid 0 < n \leq M \right).
\]

But \( \prod_{p \mid n} (1 + p^{-1/4}) \leq \sum_{d \mid n} d^{-1/4} \), and hence

\[
E\left( \prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq M \right) \leq O_{m,q}(1) \frac{1}{2M} \sum_{1 \leq |n| \leq M} \sum_{d \mid n} d^{-1/4}
\]

\[\leq O_{m,q}(1) \frac{1}{2M} \sum_{d=1}^{M} \frac{M}{d} d^{-1/4},\]

which is \( O_{m,q}(1) \) as desired. \( \square \)

We are now ready to verify the correlation condition.

Proposition 3.4. The measure \( \nu \) satisfies the \( 2^{k-1} \)-correlation condition.

Proof. Let us begin by recalling what it is we wish to prove. For any \( 1 \leq m \leq 2^{k-1} \) and \( h_1, \ldots, h_m \in \mathbb{Z}_N \) we must show a bound

\[
E(\nu(x + h_1)\nu(x + h_2) \ldots \nu(x + h_m) \mid x \in \mathbb{Z}_N) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j), \quad (3.4)
\]

where the weight function \( \tau = \tau_m \) is bounded in \( L^q \) for all \( q \).

Fix \( m, h_1, \ldots, h_m \). We shall take the weight function constructed in Lemma 3.3 (identifying \( \mathbb{Z}_M \) with the integers between \(-M/2\) and \(+M/2\), and set

\[
\tau(0) := \exp(C m \log N / \log \log N)
\]
for some large absolute constant $C$. From the previous lemma we see that $\mathbb{E}(\tau q) = O(m,q)(1)$ for all $q$, since the addition of the weight $\tau(0)$ at 0 only contributes $o(m,q)(1)$ at most.

We first dispose of the easy case when at least two of the $h_i$ are equal. In this case we bound the left-hand side of (2.2) crudely by $\|\nu\|_L^\infty$. But from Definitions 2.1, 2.2 and by standard estimates for the maximal order of the divisor function $d(n)$ we have the crude bound $\|\nu\|_L^\infty \ll \exp(C \log N/\log \log N)$, and the claim follows thanks to our choice of $\tau(0)$.

Suppose then that the $h_i$ are distinct. Write $g(n) := \phi(W) \Lambda^2_{\mathcal{R}}(Wn + 1) \mathbf{1}_{[N+\epsilon_kM,N+2\epsilon_kM]}(n)$. Then by construction of $\nu$ (Definition 2.2), we have

$$\mathbb{E}(\nu(x + h_1) \ldots \nu(x + h_m) \mid x \in \mathbb{Z}_M) \leq \mathbb{E}((1 + g(x + h_1)) \ldots (1 + g(x + h_m)) \mid x \in \mathbb{Z}_M).$$

The right-hand side may be rewritten as

$$\sum_{A \subseteq \{1, \ldots, m\}} \mathbb{E}\left(\prod_{i \in A} g(x + h_i) \mid x \in \mathbb{Z}_M\right)$$

Observe that for $i, j \in A$ we may assume $|h_i - h_j| \leq \epsilon_kM$, since the expectation vanishes otherwise. By Proposition 3.2 and Lemma 3.3 we therefore have

$$\mathbb{E}\left(\prod_{i \in A} g(x + h_i) \mid x \in \mathbb{Z}_M\right) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j) + o_m(1).$$

Summing over all $A$, and adjusting the weights $\tau$ by a bounded factor (depending only on $m$ and hence on $k$), we obtain the result.

4. Proof of the main theorem

In this section we conclude the proof of Theorem 1.1

**Definition 4.1.** Let $\nu : \mathbb{Z}_M \rightarrow \mathbb{R}^+$ be a function. We say that $\nu$ is $k$-pseudorandom measure if it obeys the estimate $\mathbb{E}(\nu) = 1 + o(1)$ and satisfies the $(k \cdot 2^{k-1}, 3k-4, k)$-linear forms condition as well as the $2^{k-1}$-correlation condition.

**Theorem 4.2** (Green-Tao). The function $\nu : \mathbb{Z}_M \rightarrow \mathbb{R}^+$ in Definition 2.2 is a $k$-pseudorandom measure that majorises $f$ in Theorem 1.1

**Proof.** That theorem follows from Lemmas 2.3, 2.6 and Propositions 2.7, 3.4. □

**Remarks.** I have attributed this theorem to Green and Tao, because the above argument is a straightforward generalisation of that of [4, Proposition 9.1].

The proof of Theorem 1.1 is base on the following theorem.
Theorem 4.3 (Green-Tao). Let \( k \geq 3 \) and \( 0 < \delta \leq 1 \) be fixed parameters. Suppose that \( \nu : \mathbb{Z}_M \to \mathbb{R}^+ \) is \( k \)-pseudorandom measure. Let \( f : \mathbb{Z}_M \to \mathbb{R}^+ \) be any non-negative function obeying the bound
\[
0 \leq f(x) \leq \nu(x) \text{ for all } x \in \mathbb{Z}_M
\] (4.1)
and
\[
\mathbb{E}(f) \geq \delta.
\] (4.2)
Then we have
\[
\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) | x, r \in \mathbb{Z}_M) \geq c(k, \delta) - o_{k, \delta}(1)
\] (4.3)
where \( c(k, \delta) > 0 \) stands for a constant depending only on \( k \) and \( \delta \).

That theorem is a great generalization of the following theorem.

Theorem 4.4 (Szemerédi’s theorem). Let \( k \geq 3 \) and \( 0 < \delta \leq 1 \) be fixed parameters. Let \( f : \mathbb{Z}_M \to \mathbb{R}^+ \) be any function which is bounded by a bound independent of \( M \). Suppose that
\[
\mathbb{E}(f) \geq \delta.
\] (4.4)
Then we have
\[
\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) | x, r \in \mathbb{Z}_M) \geq c(k, \delta) - o_{k, \delta}(1)
\] (4.5)
where \( c(k, \delta) > 0 \) is the same constant which appears in Theorem 4.3. (The decay rate \( o_{k, \delta}(1) \), on the other hand, decays significantly faster than that in Theorem 4.3.)

Remarks. The \( k = 3 \) case of Szemerédi’s theorem was established by Roth [5]. The general case as well as the \( k = 4 \) case was proved by Szemerédi [6, 7]. The formulation here is different from the original one, but can be deduce from the original one. The argument was first worked out by Varnavides [9]). A direct proof of Theorem 4.4 can be found in [8].

Proof of Theorem 1.1. By our assumption on \( M \), we see that
\[
\mathbb{E}(f) = \frac{k^{-12-k^{-5}}}{M} \sum_{N+\epsilon_kM \leq n \leq N+2\epsilon_kM} \tilde{\Lambda}(n) = k^{-12-k^{-5}}\epsilon_k(1+o(1)).
\]
We now apply Theorem 4.2 and Theorem 4.3 to conclude that
\[
\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) | x, r \in \mathbb{Z}_N) \geq c(k, k^{-12-k^{-5}}\epsilon_k) - o(1).
\]
Theorem 1.1 follows by setting \( c_k = c(k, k^{-12-k^{-5}}\epsilon_k) \).

REFERENCES

[1] D. Goldston and C.Y. Yıldırım Higher correlations of divisor sums related to primes, I: Triple correlations, Integers 3 (2003) A5, 66pp.
[2] ———, Higher correlations of divisor sums related to primes, III: \( k \)-correlations, preprint (available at AIM preprints)
[3] ———, Small gaps between primes, I, preprint.
[4] B.J. Green and T. Tao, The primes contains arbitrarily long arithmetic progressions, to appear in Ann. Math.
[5] K.F. Roth, On certain sets of integers, J. London Math. Soc. 28 (1953), 245-252.
[6] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104.

[7] ________, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299–345.

[8] T. Tao, *A quantitative ergodic theory proof of Szemerédi’s theorem*, preprint.

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

School of Mathematical Science, Beijing Normal University, Beijing 100875, E-mail address: clliu@bnu.edu.cn