The Chen primes contain arbitrarily long arithmetic progressions

by

BINBIN ZHOU (Shanghai)

1. Introduction and main results. Prime is one of the most important notions in number theory. Recently, an astonishing breakthrough was made by B. Green and T. Tao [11] who showed that there are arbitrarily long arithmetic progressions of primes; this result is known as the Green–Tao Theorem.

THEOREM 1.1. *The prime numbers contain infinitely many arithmetic progressions of length k for all k.*

In fact, they say something a little stronger:

THEOREM 1.2 (Green–Tao Theorem). *Let A be any subset of the prime numbers of positive relative upper density, so $\limsup_{N\to\infty} \pi(N)^{-1}|A\cap[1,N]| > 0$, where $\pi(N)$ denotes the number of primes less than or equal to N. Then A contains infinitely many arithmetic progressions of length k for all k.*

If one replaces "primes" in the statement of the Green–Tao Theorem by the set $\mathbb{Z}^+$ of all positive integers, then this is a famous theorem of Szemerédi [19], [5], [9]. The special case $k = 3$ of the Green–Tao Theorem was established by B. Green [10] using methods of Fourier analysis, and the special case $k = 3$ of Szemerédi's theorem is the Roth–Bourgain theorem [18], [1]. Moreover, D. R. Heath-Brown obtained some important results for arithmetic progressions in primes [16].

On the other hand, J. R. Chen [2], [3] proved his famous theorem.

DEFINITION. A prime p is a *Chen prime* if $p + 2$ is either a prime or a product $p_1 p_2$ with $p_1, p_2 \geq p^{1/10}$.

THEOREM 1.3 (Chen Theorem relative to twin primes). *Let $x(1, 2)$ denote the number of Chen primes less than or equal to x. For sufficiently*

large integer x we have

$$x(1,2) \geq \frac{0.67xC_x}{\log^2 x}$$

where

$$C_x := \prod_{\substack{p|x \\ p>2}} \frac{p-1}{p-2} \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right).$$

REMARK. It is important for us to have this extra information. In Iwaniec's unpublished notes [17] one may find a proof of Chen's theorem which leads to the exponent $3/11$ in place of $1/10$. Actually, one can modify Chen's theorem in a simple way to get

LEMMA 1.4. *Let N be a large integer. Then the number of Chen primes in the interval $(N/2, N]$ is at least $c_1 N/\log^2 N$, for some absolute constant c_1, and the smallest prime factor of $p + 2$ is bounded below by $p^{1/10}$.*

Combining the two aspects, recently B. Green and T. Tao [12] proved

THEOREM 1.5 (Green–Tao). *There are infinitely many 3-term arithmetic progressions of Chen primes.*

A little earlier, Tolev [20] used the sieve method and the Hardy–Littlewood circle method to show that there are infinitely many 3-term progressions $p_1 < p_2 < p_3$ of primes such that $p_i + 2$ is a product of at most r_i primes, where (r_1, r_2, r_3) can be taken to be $(5, 5, 8)$ or $(4, 5, 11)$.

Using the method of [11] and [12], we prove the main result of this paper.

THEOREM 1.6. *The Chen prime numbers contain infinitely many arithmetic progressions of length k for all k.*

There is a conjecture due to Erdős and Turán:

CONJECTURE 1.7 (Erdős and Turán [4]). *Suppose that $A = \{a_1 < a_2 < \cdots\}$ is an infinite sequence of integers such that $\sum 1/a_i = \infty$. Then A contains arbitrarily long arithmetic progressions.*

REMARK. It is well known that the sum $\sum_p \frac{1}{p}$ is divergent, and easy to verify that $\sum_{p_c} \frac{1}{p_c}$ (p_c denote the Chen primes) is convergent, hence the Erdős–Turán conjecture would imply Theorem 1.1, but not Theorem 1.6.

Now, let us recall some notation from [11], [12]. For a statement P, we will occasionally write 1_P to denote the indicator of P, thus $1_P = 1$ if P is true and $1_P = 0$ if P is false. If A is a set we use 1_A to denote the function $1_A(x) := 1_{x \in A}$. We will always write $|A| = \sum_x 1_A(x)$ for the cardinality of A.

If $f : A \to \mathbb{R}$ is a function and A is a non-empty set, we write $E(f) := E(f(x) \mid x \in A)$ for the average of f on A, that is, $E(f) := |A|^{-1} \sum_{x \in A} f(x)$.

There is a deep and beautiful theorem of [11] on which this paper is heavily based:

THEOREM 1.8 (Green–Tao, Szemerédi's theorem relative to a pseudo-random measure). *Let $k \geq 3$ and $0 < \delta \leq 1$ be fixed parameters. Suppose that $v : Z_N \to \mathbb{R}^+$ is a k-pseudorandom measure (see Definition 3.3 of [11]). Let $f : Z_N \to \mathbb{R}^+$ be any non-negative function obeying the bound*

$$0 \leq f(x) \leq v(x) \quad \text{for all } x \in Z_N$$

and

$$E(f) \geq \delta.$$

Then

$$E\big(f(x)f(x+r)\cdots f(x+(k-1)r) \mid x, r \in Z_N\big) \geq c(k,\delta) - o_{k,\delta}(1)$$

where $c(k,\delta) > 0$ (the decay rate $o_{k,\delta}(1)$ depends of course on the decay rates in the linear forms and correlation conditions).

Proof. See Theorem 3.5 of [11].

2. A pseudorandom measure majorizes the Chen primes. We are now ready to apply Theorem 1.8 to the specific situation involving arithmetic progressions in the Chen primes. As in [11], we first employ a device called the W-trick, which effectively removes the arithmetic obstructions to pseudorandomness arising from the very small primes. Let $W = W(N)$ be any function tending slowly ($W(N) \ll \log \log N$ will suffice) to infinity with N, so that $1/W(N) = o(1)$, let $W := \prod_{p \leq W(N)} p$ be the product of the primes up to $W(N)$, and $\epsilon_k = 1/2^k(k+4)!$. Define the modified von Mangoldt function (relative to Chen primes) $\widetilde{\Lambda}_C : \mathbb{Z}^+ \to \mathbb{R}^+$ by

$$(2.1) \quad \widetilde{\Lambda}_C(n) := \begin{cases} \dfrac{\phi(W)}{W} \log(Wn + b) & \text{when } Wn + b \text{ is a Chen prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Now let us use Chen's theorem and the pigeonhole principle to capture a large number of Chen primes. Let N be a sufficiently large prime number and let $X_W \subseteq Z_W$ denote the residue classes $b \in Z_W$ such that b and $b + 2$ are coprime to W. Observe that

$$(2.2) \quad |X_W| = W \prod_{3 \leq p \leq W(N)} \left(1 - \frac{2}{p}\right) \leq W \prod_{3 \leq p \leq W(N)} \left(1 - \frac{1}{p}\right)^2$$

$$= 4W \left(\frac{\phi(W)}{W}\right)^2.$$

Note that all but $O(W)$ of Chen primes lie in one of the residue classes

in X_W. From Lemma 1.4, we have

$$(2.3) \qquad \sum_{b \in X_W} |\{\epsilon_k N \le n \le 2\epsilon_k N : Wn + b \text{ is a Chen prime}\}|$$

$$\gg \left[\frac{W\epsilon_k N}{\log^2(W\epsilon_k N)} - O(W) \right] \gg \frac{W\epsilon_k N}{\log^2 N}.$$

Here the implied constant is absolute. Combining (2.2), (2.3) and the pigeonhole principle, we can therefore choose $b \in X_W$ such that

$$(2.4) \qquad |X| \gg \frac{\epsilon_k N}{\log^2 N} \left(\frac{W}{\phi(W)} \right)^2$$

where X is the set

$$X := \{\epsilon_k N \le n \le 2\epsilon_k N : Wn + b \text{ is a Chen prime}\}.$$

Now we can majorize the Chen primes by a pseudorandom measure as follows:

PROPOSITION 2.1. *Let N be a sufficiently large prime. Then there is a k-pseudorandom measure $v : Z_N \to \mathbb{R}^+$ such that $v(n) \ge [k^{-1}2^{-k-5}\widetilde{\Lambda}_C(n)]^2$ for all $\epsilon_k N \le n \le 2\epsilon_k N$.*

Proof of Theorem 1.6 assuming Proposition 2.1. Let N be a large prime. Define $f \in L^1(Z_N)$ by setting $f(n) = [k^{-1}2^{-k-5}\widetilde{\Lambda}_C(n)]^2$ for $\epsilon_k N \le n \le 2\epsilon_k N$ and $f(n) = 0$ otherwise. From the definition of $f(n)$ and (2.4), we observe that

$$E(f) = \frac{k^{-2}2^{-2k-10}}{N} \sum_{\epsilon_k N \le n \le 2\epsilon_k N} \widetilde{\Lambda}_C(n)^2$$

$$\ge \frac{k^{-2}2^{-2k-10}}{N} \times |X| \times \left[\frac{\phi(W)}{W} \log(W\epsilon_k N + b) \right]^2 \ge c'\epsilon_k k^{-2}2^{-2k-10}.$$

Here c' is an absolute positive constant. Now apply Theorem 1.8 and Proposition 2.1 to conclude that

$$E\big(f(x)f(x+r)\cdots f(x+(k-1)r) \mid x, r \in Z_N\big) \ge c(k, c'\epsilon_k k^{-2}2^{-2k-10}) - o(1).$$

Observe that the degenerate case $r = 0$ can only contribute at most $O(N^{-1}\log^{2k} N) = o(1)$ to the left side and can thus be discarded. The rest of the argument goes exactly as in Theorem 1.1 of [11].

REMARK. The value $c(k, c'\epsilon_k k^{-2}2^{-2k-10})$ is very small indeed, especially for large k, and this also occurs in [11], but would not affect our desired result, and B. Green and T. Tao develop a more recent approach in a series of papers [13]–[15] to give the correct asymptotic value for $k = 4$.

Thus to obtain arbitrarily long arithmetic progressions in the Chen primes, it will suffice to prove Proposition 2.1. This will be the purpose

of the remainder of this section. To obtain a majorant for $\widetilde{\Lambda}_C$, we introduce

DEFINITION 2.2 (Goldston–Pintz–Yıldırım truncated divisor sum). Let R be a parameter (in applications it will be a small power of N). Define

$$\Lambda_R(n) = \sum_{\substack{d|n \\ d \leq R}} \mu(d) \log(R/d) = \sum_{d|n} \mu(d) \log(R/d)_+.$$

These truncated divisor sums have been studied in several papers, most notably in the works of Goldston, Pintz and Yıldırım [6]–[8] concerning the problem of finding small gaps between primes. We shall obtain asymptotics for these truncated primes to prove that the measure v defined below is pseudorandom.

DEFINITION 2.3. Let $R = N^{k^{-1}2^{-k-4}}$. We define the function $v: Z_N \to \mathbb{R}^+$ by

(2.5) $\quad v(n)$

$$:= \begin{cases} \left[\dfrac{\phi(W)}{W}\right]^2 \dfrac{\Lambda_R(Wn+b)^2 \Lambda_R(Wn+b+2)^2}{\log^2 R} & \text{when } \epsilon_k N \leq n \leq 2\epsilon_k N, \\ 1 & \text{otherwise,} \end{cases}$$

for all $0 \leq n < N$, where we identify $\{0, \ldots, N-1\}$ with Z_N in the usual manner.

This v will be our majorant for Proposition 2.1. We first verify that it is indeed a majorant.

LEMMA 2.4. *We have $v(n) \geq 0$ for all $n \in Z_N$, and $v(n) \geq f(n) = [k^{-1}2^{-k-5}\widetilde{\Lambda}_C(n)]^2$ for all $\epsilon_k N \leq n \leq 2\epsilon_k N$ (if N is sufficiently large depending on k).*

Proof. The first claim is trivial. From the definition of R, we see that $Wn + b > R$ if N is sufficiently large, and from Lemma 1.4, if p is a Chen prime then $p + 2$ is either a prime or a product $p_1 p_2$ with $p_1, p_2 > p^{1/10}$. Hence if $Wn + b + 2$ is prime, then $Wn + b + 2 > R$; if $Wn + b + 2$ is a product $p_1 p_2$, also $p_1, p_2 > (Wn+b)^{1/10} > N^{k^{-1}2^{-k-4}} = R$ if N is sufficiently large. So we get $\Lambda_R(Wn + b) = \Lambda_R(Wn + b + 2) = \log R$, which means that $v(n) = [\phi(W)/W]^2 \log^2 R \geq [k^{-1}2^{-k-5}\widetilde{\Lambda}_C(n)]^2 = f(n)$ by construction of R and N (assuming $W(N)$ is sufficiently slowly growing in N).

We will prove later that v is actually a measure (in the sense of [11], i.e. $E(v) = 1 + o(1)$). First, we introduce two crucial propositions.

PROPOSITION 2.5 (on linear forms condition). *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 $j = 1, \ldots, t$. *Assume that the t-tuples $(L_{ij})_{j=1}^{t}$ are never identically zero, and that no two t-tuples are rational multiples of each other. Write $\theta_i := W\psi_i + b$ and $\theta_{i+m} := \theta_i + 2$ for all $i = 1, \ldots, m$. Suppose that B is a product $\prod_{i=1}^{t} I_i \subset R^t$ of t intervals I_i, each of length at least R^{10m}. Then (if $W(N)$ is sufficiently slowly growing in N)*

$$E\big(\Lambda_R(\theta_1(X))^2 \cdots \Lambda_R(\theta_m(X))^2 \Lambda_R(\theta_{m+1}(X))^2 \cdots \Lambda_R(\theta_{2m}(X))^2 \mid X \in B\big)$$
$$= (1 + o_{m,t}(1))\left(\frac{W \log R}{\phi(W)}\right)^{2m}.$$

PROPOSITION 2.6 (Correlation condition). *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 $i = 1, \ldots, m$, and let $\triangle$ denote the integer*

$$\triangle := \prod_{1 \leq i < j \leq m} (h_i - h_j)(W(h_i - h_j) + 2)(W(h_i - h_j) - 2).$$

Then (for N sufficiently large depending on m, and assuming $W(N)$ to grow sufficiently slowly in N)

$$(2.6) \quad E\big(\Lambda_R(W(x + h_1) + b)^2 \cdots \Lambda_R(W(x + h_m) + b)^2$$
$$\times \Lambda_R(W(x + h_1) + b + 2)^2 \cdots \Lambda_R(W(x + h_m) + b + 2)^2 \mid x \in B\big)$$
$$\leq (1 + o_m(1))\left(\frac{W \log R}{\phi(W)}\right)^{2m} \prod_{\substack{p|\triangle \\ p > W(N)}} (1 + O_m(p^{-1/2})).$$

Assuming both Propositions 2.5 and 2.6, we can now conclude the proof of Proposition 2.1, but first we show that v is indeed a measure.

LEMMA 2.7. *The function v constructed in Definition 2.3 obeys the estimate $E(v) = 1 + o(1)$.*

Proof. Apply Proposition 2.5 with $m := t := 1$, $\psi_i(x_1) := x_1$ and $B := [\epsilon_k N, 2\epsilon_k N]$ (taking N sufficiently large depending on k, of course). Comparing with Definition 2.3, we thus have

$$E(v(x) \mid x \in [\epsilon_k N, 2\epsilon_k N]) = 1 + o(1).$$

But from the same definition we clearly have

$$E(v(x) \mid x \in Z_N \setminus [\epsilon_k N, 2\epsilon_k N]) = 1.$$

Combining these two results yields the lemma.

Now we verify the linear forms condition (Definition 3.1 of [11]), which is proven in a similar spirit to the above lemma.

PROPOSITION 2.8. *The function v satisfies the $(k \cdot 2^{k-1}, 3k-4, k)$-linear forms condition.*

Proof. Let $\psi_i(X) := \sum_{j=1}^{t} L_{ij}x_j + b_i$ be linear forms of the type which features in Definition of 3.1 of [11]. We wish to show that

$$(2.7) \qquad E\big(v(\psi_1(X)) \cdots v(\psi_m(X)) \mid X \in Z_N^m\big) = 1 + o(1).$$

The argument is almost exactly the same as that in Proposition 9.8 of [11]. The only difference is that whenever $(u_1, \ldots, u_t)$ is nice (see [11]), we can replace each of the $v(\psi_i(X))$ by either $\big[\frac{W \log R}{\phi(W)}\big]^2 \Lambda_R(\theta_i(X))^2 \Lambda_R(\theta_{i+m}(X))^2$ or 1, and when $(u_1, \ldots, u_t)$ is not nice, we can crudely bound v by $1 + \big[\frac{W \log R}{\phi(W)}\big]^2 \Lambda_R(\theta_i(X))^2 \Lambda_R(\theta_{i+m}(X))^2$.

Later, we use Proposition 2.6 to show that v satisfies the correlation condition (Definition 3.2 of [11]). First, we must look at the average size of the "arithmetic" factor $\prod_{p>W(N),\, p|\triangle}(1 + O_m(p^{-1/2}))$ appearing in the next lemma, which is one of the main differences compared with Green–Tao's paper [11].

LEMMA 2.9. *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 for all distinct $h_1, \ldots, h_j \in [\epsilon_k N, 2\epsilon_k N]$ we have*

$$\prod_{\substack{p|\triangle \\ p>W(N)}} (1 + O_m(p^{-1/2})) \leq \prod_{1 \leq i < j \leq m} \tau(h_i - h_j),$$

where $\triangle$ is defined in Proposition 2.6, and $E(\tau^q(n) \mid 0 < |n| \leq N) = O_{m,q}(1)$ for all $0 < q < \infty$.

Proof. We observe that

$$\prod_{\substack{p|\triangle \\ p>W(N)}} (1 + O_m(p^{-1/2}))$$

$$\leq \prod_{1 \leq i < j \leq m} \bigg(\prod_{\substack{p|(h_i-h_j)(W(h_i-h_j)+2)(W(h_i-h_j)-2) \\ p>W(N)}} (1 + p^{-1/2}) \bigg)^{O_m(1)}.$$

By the arithmetic-geometric mean inequality (absorbing all constants into the $O_m(1)$ factor) we can thus take

$$\tau_m(n) := O_m(1) \prod_{\substack{p|n(Wn+2)(Wn-2) \\ p>W(N)}} (1 + p^{-1/2})^{O_m(1)}$$

for all $n \neq 0$. (The value of τ 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

$$E\Big(\prod_{\substack{p|n(Wn+2)(Wn-2)\\p>W(N)}}(1+p^{-1/2})^{O_m(q)}\,\Big|\,0<|n|\le N\Big)=O_{m,q}(1)$$

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

$$E\Big(\prod_{\substack{p|n(Wn+2)(Wn-2)\\p>W(N)}}(1+p^{-1/2})^{O_m(q)}\,\Big|\,0<|n|\le N\Big)$$

$$\le O_{m,q}(1)E\Big(\prod_{\substack{p|n\\p>W(N)}}(1+p^{-1/4})\prod_{\substack{p|Wn+2\\p>W(N)}}(1+p^{-1/4})\prod_{\substack{p|Wn-2\\p>W(N)}}(1+p^{-1/4})\,\Big|$$

$$0<n\le N\Big)$$

$$\le O_{m,q}(1)\Big[E\Big(\prod_{\substack{p|n\\p>W(N)}}(1+p^{-1/8})\,\Big|\,I\Big)\Big]^{1/2}\Big[E\Big(\prod_{\substack{p|Wn+2\\p>W(N)}}(1+p^{-1/8})\,\Big|\,I\Big)\Big]^{1/4}$$

$$\times\Big[E\Big(\prod_{\substack{p|Wn-2\\p>W(N)}}(1+p^{-1/8})\,\Big|\,I\Big)\Big]^{1/4}.$$

The last step uses Hölder's inequality, and I denotes $0<n\le N$. Moreover, $\prod_{p|n}(1+p^{-1/8})\le\sum_{d|n}d^{-1/8}$, hence

$$E\Big(\prod_{\substack{p|n\\p>W(N)}}(1+p^{-1/8})\,\Big|\,0<n\le N\Big)\le E\Big(\prod_{p|n}(1+p^{-1/8})\,\Big|\,0<n\le N\Big)$$

$$\le\frac{1}{N}\sum_{1\le n\le N}\sum_{d|n}d^{-1/8}\le\frac{1}{N}\sum_{d=1}^{N}\frac{N}{d}d^{-1/8}=\sum_{d=1}^{N}d^{-9/8}.$$

It is readily verified that

$$\prod_{\substack{p|Wn+2\\p>W(N)}}(1+p^{-1/8})\le\sum_{\substack{d|Wn+2\\p_d>W(N)}}d^{-1/8}$$

(p_d denotes the smallest prime divisor of d); we recall the definition $W:=\prod_{p\le W(N)}p$, together with $p_d>W(N)$, hence $(d,W)=1$. For every d satisfying $(d,W)=1$, we denote by n_1 the smallest positive integer such that $d\,|\,Wn_1+2$, and for every n_i such that $d\,|\,Wn_i+2$, we have $d\,|\,n_i-n_1$, i.e. when n runs over all $1\le n\le N$, there are at most $1+\lfloor N/d\rfloor\le 1+N/d$ distinct n such that $d\,|\,Wn+2$, and observe that only values with $d\le NW+2\le 2NW$

can contribute. Hence

$$E\Big(\prod_{\substack{p|Wn+2 \\ p>W(N)}} (1+p^{-1/8}) \,\Big|\, 0 < n \le N\Big) \ll \frac{1}{N} \sum_{d \le 2NW} (1+N/d)d^{-1/8}$$

$$\ll N^{-1} \sum_{d \le 2NW} d^{-1/8} + \sum_{d \le 2NW} d^{-9/8} \ll N^{-1}(NW)^{7/8} + O(1).$$

However, $W(N) \ll \log\log N$, we have $W = O(N^{1/7})$, so that

$$N^{-1}(NW)^{7/8} = O(1),$$

hence $O_{m,q}(1)$ as desired.

We are now ready to verify the correlation condition.

PROPOSITION 2.10. *The measure v satisfies the 2^{k-1}-correlation condition.*

Proof. Let us begin by recalling what we wish to prove. For any $1 \le m \le 2^{k-1}$ and $h_1, \ldots, h_m \in Z_N$ we must show a bound

$$(2.8) \quad E(v(x+h_1)v(x+h_2)\cdots v(x+h_m) \mid x \in Z_N) \le \sum_{1 \le i < j \le m} \tau(h_i - h_j)$$

where the weight function $\tau = \tau_m$ is bounded in L^q for all q. The argument follows that for Proposition 9.10 of [11]; the only difference is that we set

$$\tau(0) := \exp(2Cm\log N/\log\log N),$$

and use the crude bound $\|v\|_{L^\infty} \ll \exp(2C\log N/\log\log N)$. Then the claim follows thanks to our choice of $\tau(0)$.

Now, set

$$g(n) := \left[\frac{\phi(W)}{W}\right]^2 \frac{\Lambda_R(Wn+b)^2 \Lambda_R(Wn+b+2)^2}{\log^2 R} 1_{[\epsilon_k N, 2\epsilon_k N]}(n).$$

Now we have enough tools to obtain the result.

Proof of Proposition 2.1. This is immediate from Lemma 2.4, Lemma 2.7, Proposition 2.8, Proposition 2.10 and the definition of k-pseudorandom measures (see Definition 3.3 of [11]).

3. Correlation estimate for Λ_R. To conclude the proof of Theorem 1.6, it remains to verify Propositions 2.5 and 2.6.

We begin by proving Proposition 2.5. Recall that for each $1 \le i \le m$, we have a linear form $\psi_i(X) := \sum_{j=1}^{t} L_{ij}x_j + b_i$ in t variables $x_1, \ldots, x_t$. Define $\theta_i := W\psi_i + b$ and $\theta_{i+m} := W\psi_i + b + 2$ for all $1 \le i \le m$.

Let $B := \prod_{j=1}^{t} I_j$ be a product of intervals I_j, each of length at least R^{10m}. We wish to prove the estimate

$$(3.1) \quad E\big(\Lambda_R(\theta_1(X))^2 \cdots \Lambda_R(\theta_{m+1}(X))^2 \cdots \Lambda_R(\theta_{2m}(X))^2 \mid X \in B\big)$$

$$= (1 + o_{m,t}(1))\Big(\frac{W \log R}{\phi(W)}\Big)^{2m}.$$

The first step is to eliminate the role of the box B. We can use Definition 2.2 to expand the left-hand side as

$$E\Big(\prod_{i=1}^{2m} \sum_{\substack{d_i, d_i' \leq R \\ d_i, d_i' \mid \theta_i(X)}} \mu(d_i)\mu(d_i') \log\frac{R}{d_i} \log\frac{R}{d_i'} \,\Big|\, X \in B\Big),$$

which we can rearrange as

$$\sum_{d_1,\ldots,d_{2m},d_1',\ldots,d_{2m}' \leq R} \Big(\prod_{i=1}^{2m} \mu(d_i)\mu(d_i') \log\frac{R}{d_i} \log\frac{R}{d_i'}\Big) E\Big(\prod_{i=1}^{2m} 1_{d_i,d_i' \mid \theta_i(X)} \,\Big|\, X \in B\Big).$$

The rest of the argument follows Section 10 of [11], the only difference is the contribution of the error term $O_m(R^{-6m})$ corresponding to (10.1) in [11], which can be crudely estimated by $O_{m,t}(R^{-2m} \log^{4m} R)$, which is clearly acceptable; then we just simply need to replace m by $2m$, and consider the behavior of $\omega_X(p)$. We recall the definitions of $\omega_X(p)$ and E_p (or see (10.3) and (10.7) of [11]).

If we write $X_{d_1,\ldots,d_{2m}}(p) := \{1 \leq i \leq 2m : p \mid d_i\}$ then

$$\omega_X(p) := E\Big(\prod_{i \in X} 1_{\theta_i(X) \equiv 0 \,(\mathrm{mod}\, p)} \,\Big|\, X \in Z_p^t\Big)$$

for each subset $X \subseteq \{1, \ldots, 2m\}$, and

$$E_p(z, z') := \sum_{X,X' \subseteq \{1,\ldots,2m\}} \frac{(-1)^{|X|+|X'|}\omega_{X \cup X'}(p)}{p^{\sum_{j \in X} z_j + \sum_{j \in X'} z_j'}}.$$

The main difference from [11] is in the proof of the next lemma, especially Case II.

LEMMA 3.1 (Local factor estimate). *If $p \leq W(N)$, then $\omega_X(p) = 0$ for all non-empty X; in particular, $E_p = 1$ when $p \leq W(N)$. If instead $p > W(N)$, then $\omega_X(p) = p^{-1}$ when $|X| = 1$ and $\omega_X(p) \leq p^{-2}$ when $|X| \geq 2$.*

Proof. The first statement is clear, since the maps $\theta_j : Z_p^t \to Z_p$ are identically b or $b+2$, which are coprime to W when $p \leq W(N)$. The second statement (when $p > W(N)$ and $|X| = 1$) is similar since in this case θ_j uniformly covers Z_p (in the terminology of [11]). Now suppose $p > W(N)$ and $|X| = 2$ ($X = \{i, j\}$), and set $Y_1 := \{1, \ldots, m\}$ and $Y_2 := \{m+1, \ldots, 2m\}$. Our discussion is divided into the following two cases.

CASE I: $\{i,j\} \subseteq Y_1$ or $\{i,j\} \subseteq Y_2$. In this case we argue exactly as in Lemma 10.1 of [11]. We claim that none of the s pure linear forms $W(\psi_i - b_i)$ is a multiple of any other (modulo p). Indeed, otherwise we would have $L_{ij}L_{i'j}^{-1} = \lambda \pmod{p}$ for some λ and all $j = 1, \ldots, t$. But if a/q and a'/q' are two rational numbers in lowest terms, with $|a|, |a'|, q, q' < \sqrt{W(N)}/2$, then clearly $a/q \neq a'/q' \pmod{p}$ unless $a = a'$ and $q = q'$. It follows that the two pure linear forms $\psi_i - b_i$ and $\psi_{i'} - b_{i'}$ are rational multiples of each other, contrary to assumption. Thus the set of $X \in Z_p^t$ for which $\theta_i \equiv 0 \pmod{p}$ for all $i \in X$ is contained in the intersection of two skew affine subspaces of Z_p^t, and as such has cardinality at most p^{t-2}.

CASE II: i, j (assume that $i < j$) belong to different Y_l ($l = 1, 2$). If $j - i = m$ (here $\theta_{i+m} - \theta_i = 2$), then there is no solution $X_0 \in Z_p^t$ such that $\theta_i(X_0) \equiv \theta_j(X_0) \equiv 0 \pmod{p}$ when $p > W(N)$. If $j - i \neq m$, notice that $\theta_i(X) = W(\sum_{l=1}^t L_{il}x_l + b_i) + b$ and $\theta_j(X) = W(\sum_{l=1}^t L_{j-m,l}x_l + b_{j-m}) + b + 2$. Since $p > W(N)$, we can take the difference 2 ($b + 2 - b = 2$) into $b_i - 2W_p^{-1}$ ($i \leq m$, and W_p^{-1} satisfies $W \cdot W_p^{-1} \equiv 1 \pmod{p}$), where b_i is not restricted, and then the argument is analogous to that in Case I.

The third statement is that when $p > W(N)$ and $|X| \geq 3$, we can take any two indices $\{i,j\}$ such that $\{i,j\} \subseteq Y_1$ or $\{i,j\} \subseteq Y_2$, which is just Case I when $p > W(N)$ and $|X| = 2$. The rest of the argument is as in [11], we only need to replace m by $2m$. Thus we can conclude the proof of Proposition 2.5 following Green–Tao's [11] method.

Higher order correlation for Λ_R. We now prove Proposition 2.6; the main difference from [11] is now also that (recall $Y_1 = \{1, \ldots, m\}$ and $Y_2 = \{m+1, \ldots, 2m\}$)

$$\omega_X(p)$$
$$:= E\Big(\prod_{i \in X \cap Y_1} 1_{W(x+h_i)+b \equiv 0 \,(\mathrm{mod}\, p)} \prod_{i \in X \cap Y_2} 1_{W(x+h_{i-m})+b+2 \equiv 0 \,(\mathrm{mod}\, p)} \,\Big|\, x \in Z_p \Big).$$

LEMMA 3.2. *If $p \leq W(N)$ then $\omega_X(p) = 0$ for all non-empty X; in particular, $E_p = 1$ when $p \leq W(N)$. If instead $p > W(N)$, then $\omega_X(p) = p^{-1}$ when $|X| = 1$ and $\omega_X(p) \leq p^{-1}$ when $|X| \geq 2$. Furthermore, if $|X| \geq 2$ then $\omega_X(p) = 0$ unless $p \,|\, \triangle := \prod_{1 \leq i < j \leq s}(h_i - h_j)(W(h_i - h_j) + 2)(W(h_i - h_j) - 2)$, $s \leq m$.*

Proof. If $p \leq W(N)$ then $W(x + h_i) + b \equiv b \pmod{p}$ and the claim follows. When $p > W(N)$ and $|X| = 1$, $\omega_X(p)$ is equal to $1/p$. When $p > W(N)$ and $|X| \geq 2$, $\omega_X(p) \neq 0$; then the residue classes $\{h_i \pmod{p} : i \in X \cap Y_1\}$ are all equal or the residue classes $\{Wh_i, Wh_{i'-m} + 2 \pmod{p} : i \in X \cap Y_1 \text{ or } i' \in X \cap Y_2\}$ are all equal; i.e. when we assume that $i < j$, the first case is that p divides some $h_i - h_j$, and the latter case is that p divides

some $W(h_i - h_j) + 2$ or $W(h_i - h_j) - 2$; it is easy to see that $\omega_X(p)$ is equal to $1/p$ in the first case, and zero otherwise, and the claim again follows.

The remaining argument is almost exactly as in [11], the main difference is the next lemma. First, we recall some notation from Definition 10.2 of [11]. For any $\sigma > 0$, let $D_\sigma^{2m} \subseteq \mathbb{C}^{4m}$ denote the domain

$$D_\sigma^{2m} := \{z_j, z_j' : -\sigma < \operatorname{Re}(z_j), \operatorname{Re}(z_j') < 100, \ j = 1, \ldots, 2m\}.$$

If $G = G(z, z')$ is an analytic function of $4m$ complex variables on D_σ^{2m}, we define the $C^k(D_\sigma^{2m})$ norm of G for any integer $k \geq 0$ as

$$\|G\|_{C^k(D_\sigma^{2m})}$$

$$:= \sup_{a_1 + \cdots + a_{2m}' \leq k} \left\| \left(\frac{\partial}{\partial z_1}\right)^{a_1} \cdots \left(\frac{\partial}{\partial z_{2m}}\right)^{a_{2m}} \left(\frac{\partial}{\partial z_1'}\right)^{a_1'} \cdots \left(\frac{\partial}{\partial z_{2m}'}\right)^{a_{2m}'} G \right\|_{L_\infty(D_\sigma^{2m})}$$

where $a_1, \ldots, a_{2m}'$ range over all non-negative integers with total sum at most k. Moreover, $\lambda_p(z, z')$ is an expression of the form

$$\lambda_p(z, z') = \sum_{\substack{X, X' \subseteq \{1, \ldots, 2m\} \\ |X \cup X'| \geq 2}} \frac{O(1/p)}{p^{\sum_{j \in X} z_j + \sum_{j \in X'} z_j'}}$$

where the $O(1/p)$ quantity does not depend on z, z'. Furthermore,

$$E_p^{(0)} = 1 + 1_{p > W(N),\, p | \triangle} \lambda_p(z, z'),$$

$$E_p^{(1)} = \frac{E_p}{E_p^{(0)} \prod_{j=1}^{2m} (1 - 1_{p > W(N)} p^{-1 - z_j})(1 - 1_{p > W(N)} p^{-1 - z_j'})(1 - 1_{p > W(N)} p^{-1 - z_j - z_j'})^{-1}},$$

$$E_p^{(2)} = \prod_{j=1}^{2m} (1 - 1_{p > W(N)} p^{-1 - z_j})^{-1} (1 - 1_{p > W(N)} p^{-1 - z_j'})^{-1} (1 - 1_{p > W(N)} p^{-1 - z_j - z_j'}),$$

$$E_p^{(3)} = \prod_{j=1}^{2m} (1 - p^{-1 - z_j})(1 - p^{-1 - z_j'})(1 - p^{-1 - z_j - z_j'})^{-1}.$$

Write $G_j := \prod_p E_p^{(j)}$ for $j = 0, 1, 2, 3$.

LEMMA 3.3. *Let $0 < \sigma < 1/12m$. Then the Euler products $\prod_p E_p^{(l)}$ for $l = 0, 1, 2$ are absolutely convergent in the domain D_σ^{2m}. In particular, G_0, G_1, G_2 can be continued analytically to this domain. Furthermore, we have the estimates*

$$(3.2) \quad \|G_0\|_{C^r(D_{1/12m}^{2m})}$$

$$\leq O_m\left(\frac{\log R}{\log \log R}\right)^r \prod_{\substack{p | \triangle \\ p > W(N)}} (1 + O_m(p^{4m\sigma - 1})) \quad \text{for } 0 \leq r \leq 2m,$$

(3.3)
$$\|G_0\|_{C^{2m}(D_{1/12m}^{2m})} \leq \exp(O_m(\log^{1/3} R)),$$
$$\|G_1\|_{C^{2m}(D_{1/12m}^{2m})} \leq O_m(1), \quad \|G_2\|_{C^{2m}(D_{1/12m}^{2m})} \leq O_{m,W(N)}(1),$$

(3.4)
$$G_0(0,0) = \prod_{\substack{p|\triangle \\ p>W(N)}} (1 + O_m(p^{-1/2})),$$

$$G_1(0,0) = 1 + o_m(1), \quad G_2(0,0) = (W/\phi(W))^{2m}.$$

Proof. The difference from Lemma 10.6 of [11] is mainly in three places. The first is

(3.5)
$$\triangle = \prod_{1\leq i<j\leq m} (h_i - h_j)(W(h_i - h_j) + 2)(W(h_i - h_j) - 2)$$
$$\leq N^{3m^2} \leq R^{O_m(1)}.$$

The second is that to prove (3.3), in light of (3.2), it suffices to prove that

$$\prod_{\substack{p|\triangle \\ p>W(N)}} (1 + O_m(p^{4m\sigma-1})) \leq \exp(O_m(\log^{1/3} R)).$$

Taking logarithms and using the hypothesis $\sigma < 1/12m$ (and (3.5)), we are reduced to showing

$$\sum_{\substack{p|\triangle \\ p>W(N)}} p^{-2/3} \leq \sum_{p|\triangle} p^{-2/3} \leq O(\log^{1/3} \triangle).$$

But there are at most $O(\log \triangle / \log\log \triangle)$ primes dividing $\triangle$, hence the left-hand side can be crudely bounded by

$$\sum_{1\leq n\leq O(\log \triangle/\log\log \triangle)} n^{-2/3} = O(\log^{1/3} \triangle)$$

as desired.

The third difference is that the bound (3.4) now follows from the crude estimate $E_p^{(0)}(z,z') = 1 + O_m(p^{-1/2})$, and the condition $p > W(N)$ (main difference from (10.14) of [11]) in (3.4) is crucial for proving Lemma 2.9.

Now, we can conclude the proof of Proposition 2.6 following Green–Tao's [11] argument.

REMARK. We have in fact shown that the number of k-tuples $(p_1, \ldots, p_k)$ which are arithmetic progressions of Chen primes and with each $p_1 \leq N$ is $\gg_k N^2/\log^{2k} N$. It is clear that if one had a lower bound $\pi_2(N) \gg N/\log^2 N$ for the number of twin primes less than N (this, of course, is one of the most open conjectures in prime number theory) then a simple adaptation of our argument would produce infinitely many k-tuples $(p_1, \ldots, p_k)$ of twin primes in arithmetic progressions, and the number of such k-tuples less than N would be $\gg_k N^2/\log^{2k} N$. Hence we conclude

PROPOSITION 3.4. *Assume the Hardy–Littlewood l-tuple conjecture, i.e.
setting $T_l(N) = \{p \in \{1, \ldots, N\} : p, p + d_1, \ldots, p + d_{l-1}$ are all prime\},
assume that $\#T_l(N) \geq cN/\log^l N$ (with c an absolute positive constant).
Then $T_l(N)$ contains infinitely many arithmetic progressions of length k for
all k, and the number of such k-tuples less than N (sufficiently large integer
only depending on k and l) is $\gg_{k,l} N^2/\log^{lk} N$.*

Outline of proof. This is a slight modification of the proof of Theorem 1.6. First, we make

DEFINITION 3.5.

$$(3.6) \qquad \widetilde{\Lambda}_l(n) := \begin{cases} \dfrac{\phi(W)}{W} \log(Wn + b) & \text{if } Wn + b \in T_l(N), \\ 0 & \text{otherwise.} \end{cases}$$

DEFINITION 3.6. Let $R = N^{l^{-1}k^{-1}2^{-k-4}}$. We define the function $v : Z_N \to \mathbb{R}^+$ by

$$(3.7) \quad v(n)$$
$$:= \begin{cases} \left[\dfrac{\phi(W)}{W}\right]^l \dfrac{\Lambda_R(Wn + b)^2 \Lambda_R(Wn + b + d_1)^2 \cdots \Lambda_R(Wn + b + d_{l-1})^2}{\log^l R} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{when } \epsilon_k N \leq n \leq 2\epsilon_k N, \\ 1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise.} \end{cases}$$

Set

$$(3.8) \qquad f(n) = \begin{cases} [l^{-1}k^{-1}2^{-k-5}\widetilde{\Lambda}_l(n)]^l & \text{when } \epsilon_k N \leq n \leq 2\epsilon_k N, \\ 0 & \text{otherwise.} \end{cases}$$

There is also an important revision corresponding to (3.1): Let $B := \prod_{j=1}^t I_j$
be a product of intervals I_j, each of length at least R^{10lm} so that the error
term corresponding to (10.1) of [11] would easily be acceptable. It is easy to
see that $\triangle$ corresponding to Lemma 2.9 would be

$$\triangle := \prod_{1 \leq i < j \leq m} (h_i - h_j)(W(h_i - h_j) + d_1)(W(h_i - h_j) - d_1) \cdots (W(h_i - h_j) - d_{l-1}).$$

The above modifications ensure that no additional difficulties arise in the proof.

References

[1]   J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. 9 (1999), 968–984.
[2]   J. R. Chen, *On the representation of a large even integer as the sum of a prime and a product of at most two primes*, Sci. Sinica 16 (1973), 157–176.
[3]   —, *Further improvement on the constant in the proposition "1+2": On the representation of a large even integer as the sum of a prime and a product of at most two primes II*, ibid. 21 (1978), 477–494 (in Chinese).
[4]   P. Erdős and P. Turán, *On some sequences of integers*, J. London Math. Soc. 11 (1936), 261–264.
[5]   H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math. 31 (1977), 204–256.
[6]   D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math., to appear.
[7]   D. A. Goldston and C. Y. Yıldırım, *Higher correlations of divisor sums related to primes, I: triple correlations*, Integers 3 (2003), A5, 66 pp.
[8]   —, —, *Small gaps between primes, I*, preprint; http://front.math.ucdavis.edu/0504336.
[9]   T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. 11 (2001), 465–588.
[10] B. Green, *Roth's theorem in the primes*, Ann. of Math. 161 (2005), 1609–1636.
[11] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. 167 (2008), 481–547.
[12] —, —, *Restriction theory of the Selberg sieve, with applications*, J. Théor. Nombres Bordeaux 18 (2006), 147–182.
[13] —, —, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc. 51 (2008), 73–153.
[14] —, —, *Quadratic uniformity of the Möbius function*, Ann. Inst. Fourier (Grenoble) 58 (2008), 1863–1935.
[15] —, —, *Linear equations in primes*, Ann. of Math., to appear.
[16] D. R. Heath-Brown, *Three primes and an almost prime in arithmetic progression*, J. London Math. Soc. (2) 23 (1981), 396–414.
[17] H. Iwaniec, *Sieve Methods*, Graduate course, Rutgers, 1996.
[18] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 245–252.
[19] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.
[20] D. I. Tolev, *Arithmetic progressions of prime-almost-prime twins*, ibid. 88 (1999), 67–98.

Department of Mathematics
Shanghai University
Shanghai, 200444
P.R. China
E-mail: dayandanghunan@163.com