(cf. the proof of Lemma 3.4). Observe that for $i, j \in A$ we may assume $|h_i - h_j| \leqslant \epsilon_k N$, since the expectation vanishes otherwise. By Proposition 9.6 and Lemma 9.9, we therefore have

$$\mathbb{E}\Big(\prod_{i \in A} g(x + h_i) \Bigm| x \in \mathbb{Z}_N \Big) \leqslant \sum_{1 \leqslant i < j \leqslant m} \tau(h_i - h_j) + o_m(1).$$

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

Proof of Proposition 9.1. This is immediate from Lemma 9.4, Lemma 9.7, Proposition 9.8, Proposition 9.10 and the definition of k-pseudorandom measure, which is Definition 3.3. □

10. Correlation estimates for Λ_R

To conclude the proof of Theorem 1.1 it remains to verify Propositions 9.5 and 9.6. That will be achieved in this section, assuming an estimate (Lemma 10.4) for a certain class of contour integrals involving the ζ-function. The proof of that estimate is given in the preprint [17], and will be repeated in the Appendix for sake of completeness. The techniques of this section are also rather close to those in [17]. We are greatly indebted to Dan Goldston for sharing this preprint with us.

The linear forms condition for Λ_R. We begin by proving Proposition 9.5. Recall that for each $1 \leqslant i \leqslant m$ we have a linear form $\psi_i(\mathbf{x}) = \sum_{j=1}^t L_{ij} x_j + b_i$ in t variables $x_1, \dots, x_t$. The coefficients L_{ij} satisfy $|L_{ij}| \leqslant \sqrt{w(N)}/2$, where $w(N)$ is the function, tending to infinity with N, which we used to set up the W-trick. We assume that none of the t-tuples $(L_{ij})_{j=1}^t$ are zero or are rational multiples of any other. Define $\theta_i := W\psi_i + 1$.

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

$$\mathbb{E}\big(\Lambda_R(\theta_1(\mathbf{x}))^2 \dots \Lambda_R(\theta_m(\mathbf{x}))^2 \bigm| \mathbf{x} \in B\big) = (1 + o_{m,t}(1)) \left(\frac{W \log R}{\phi(W)} \right)^m.$$

The first step is to eliminate the role of the box B. We can use Definition 9.2 to expand the left-hand side as

$$\mathbb{E}\Big(\prod_{i=1}^m \sum_{\substack{d_i, d_i' \leqslant R \\ d_i, d_i' | \theta_i(\mathbf{x})}} \mu(d_i)\mu(d_i') \log \frac{R}{d_i} \log \frac{R}{d_i'} \Bigm| \mathbf{x} \in B \Big)$$

which we can rearrange as

$$\sum_{d_1, \dots, d_m, d_1', \dots, d_m' \leqslant R} \Big(\prod_{i=1}^m \mu(d_i)\mu(d_i') \log \frac{R}{d_i} \log \frac{R}{d_i'} \Big) \mathbb{E}\Big(\prod_{i=1}^m 1_{d_i, d_i' | \theta_i(\mathbf{x})} \Bigm| \mathbf{x} \in B \Big). \qquad (10.1)$$

Because of the presence of the Möbius functions we may assume that all the d_i, d_i' are square-free. Write $D := [d_1, \dots, d_m, d_1', \dots, d_m']$ to be the least common multiple of the d_i and d_i', thus $D \leqslant R^{2m}$. Observe that the expression $\prod_{i=1}^m 1_{d_i, d_i' | \theta_i(\mathbf{x})}$ is periodic with period D in each of the components of x, and can thus can be safely defined on $\mathbb{Z}_D^t$. Since B is a product of intervals of length at least R^{10m}, we thus see that

$$\mathbb{E}\Big(\prod_{i=1}^m \mathbf{1}_{d_i,d_i'|\theta_i(\mathbf{x})}\Big|\mathbf{x}\in B\Big) = \mathbb{E}\Big(\prod_{i=1}^m \mathbf{1}_{d_i,d_i'|\theta_i(\mathbf{x})}\Big|\mathbf{x}\in \mathbb{Z}_D^t\Big) + O_{m,t}(R^{-8m}).$$

The contribution of the error term $O_m(R^{-8m})$ to (10.1) can be crudely estimated by $O_{m,t}(R^{-6m}\log^{2m} R)$, which is easily acceptable. Our task is thus to show that

$$\sum_{d_1,\dots,d_m,d_1',\dots,d_m'\leqslant R}\Big(\prod_{i=1}^m \mu(d_i)\mu(d_i')\log\frac{R}{d_i}\log\frac{R}{d_i'}\Big)\mathbb{E}\Big(\prod_{i=1}^m \mathbf{1}_{d_i,d_i'|\theta_i(\mathbf{x})}\Big|\mathbf{x}\in \mathbb{Z}_D^t\Big)$$
$$= (1+o_{m,t}(1))\left(\frac{W\log R}{\phi(W)}\right)^m. \quad (10.2)$$

To prove (10.2), we shall perform a number of standard manipulations (as in [17]) to rewrite the left-hand side as a contour integral of an Euler product, which in turn can be rewritten in terms of the Riemann ζ-function and some other simple factors. We begin by using the Chinese remainder theorem (and the square-free nature of d_i, d_i') to rewrite

$$\mathbb{E}\Big(\prod_{i=1}^m \mathbf{1}_{d_i,d_i'|\theta_i(\mathbf{x})}\Big|\mathbf{x}\in \mathbb{Z}_D^t\Big) = \prod_{p|D}\mathbb{E}\Big(\prod_{i:p|d_id_i'} \mathbf{1}_{\theta_i(\mathbf{x})\equiv 0 (\mathrm{mod}\, p)}\Big|\mathbf{x}\in \mathbb{Z}_p^t\Big).$$

Note that the restriction that p divides D can be dropped since the multiplicand is 1 otherwise. In particular, if we write $X_{d_1,\dots,d_m}(p) := \{1\leqslant i\leqslant m : p|d_i\}$ and

$$\omega_X(p) := \mathbb{E}\Big(\prod_{i\in X} \mathbf{1}_{\theta_i(\mathbf{x})\equiv 0 (\mathrm{mod}\, p)}\Big|\mathbf{x}\in \mathbb{Z}_p^t\Big) \quad (10.3)$$

for each subset $X\subseteq\{1,\dots,m\}$, then we have

$$\mathbb{E}\Big(\prod_{i=1}^m \mathbf{1}_{d_i,d_i'|\theta_i(\mathbf{x})}\Big|\mathbf{x}\in \mathbb{Z}_D^t\Big) = \prod_p \omega_{X_{d_1,\dots,d_m}(p)\cup X_{d_1',\dots,d_m'}(p)}(p).$$

We can thus write the left-hand side of (10.2) as

$$\sum_{d_1,\dots,d_m,d_1',\dots,d_m'\in\mathbb{Z}^+}\Big(\prod_{i=1}^m \mu(d_i)\mu(d_i')(\log\frac{R}{d_i})_+(\log\frac{R}{d_i'})_+\Big)\prod_p \omega_{X_{d_1,\dots,d_m}(p)\cup X_{d_1',\dots,d_m'}(p)}(p).$$

To proceed further, we need to express the logarithms in terms of multiplicative functions of the d_i, d_i'. To this end, we introduce the vertical line contour Γ_1 parameterised by

$$\Gamma_1(t) := \frac{1}{\log R} + it; \quad -\infty < t < +\infty \quad (10.4)$$

and observe the contour integration identity

$$\frac{1}{2\pi i}\int_{\Gamma_1}\frac{x^z}{z^2}\,dz = (\log x)_+$$

valid for any real $x > 0$. The choice of $\frac{1}{\log R}$ for the real part of Γ_1 is not currently relevant, but will be convenient later when we estimate the contour integrals that emerge (in particular, R^z is bounded on Γ_1, while $1/z^2$ is not too large). Using this identity, we can rewrite the left-hand side of (10.2) as

$$(2\pi i)^{-2m} \int_{\Gamma_1} \cdots \int_{\Gamma_1} F(z,z') \prod_{j=1}^{m} \frac{R^{z_j+z'_j}}{z_j^2 z_j'^2}\, dz_j dz'_j \tag{10.5}$$

where there are $2m$ contour integrations in the variables $z_1, \dots, z_m, z'_1, \dots, z'_m$ on Γ_1, $z := (z_1, \dots, z_m)$ and $z' := (z'_1, \dots, z'_m)$, and

$$F(z,z') := \sum_{d_1,\dots,d_m,d'_1,\dots,d'_m \in \mathbb{Z}^+} \left(\prod_{j=1}^{m} \frac{\mu(d_j)\mu(d'_j)}{d_j^{z_j} d_j'^{z'_j}} \right) \prod_p \omega_{X_{d_1,\dots,d_m}(p) \cup X_{d'_1,\dots,d'_m}(p)}(p). \tag{10.6}$$

We have changed the indices from i to j to avoid conflict with the square root of -1. Observe that the summand in (10.6) is a multiplicative function of $D = [d_1, \dots, d_m, d'_1, \dots, d'_m]$ and thus we have (formally, at least) the Euler product representation $F(z,z') = \prod_p E_p(z,z')$, where

$$E_p(z,z') := \sum_{X,X' \subseteq \{1,\dots,m\}} \frac{(-1)^{|X|+|X'|} \omega_{X \cup X'}(p)}{p^{\sum_{j \in X} z_j + \sum_{j \in X'} z'_j}}. \tag{10.7}$$

From (10.3) we have $\omega_\emptyset(p) = 1$ and $\omega_X(p) \leqslant 1$, and so $E_p(z,z') = 1 + O_\sigma(1/p^\sigma)$ when $\Re(z_j), \Re(z'_j) > \sigma$ (we obtain more precise estimates below). Thus this Euler product is absolutely convergent to $F(z,z')$ in the domain $\{\Re(z_j), \Re(z'_j) > 1\}$ at least.

To proceed further we need to exploit the hypothesis that the linear parts of $\psi_1, \dots, \psi_m$ are non-zero and not rational multiples of each other. This shall be done via the following elementary estimates on $\omega_X(p)$.

Lemma 10.1 (Local factor estimate). *If $p \leqslant w(N)$, then $\omega_X(p) = 0$ for all non-empty X; in particular, $E_p = 1$ when $p \leqslant w(N)$. If instead $p > w(N)$, then $\omega_X(p) = p^{-1}$ when $|X| = 1$ and $\omega_X(p) \leqslant p^{-2}$ when $|X| \geqslant 2$.*

Proof. The first statement is clear, since the maps $\theta_j : \mathbb{Z}_p^t \to \mathbb{Z}_p$ are identically 1 when $p \leqslant w(N)$. The second statement (when $p > w(N)$ and $|X| = 1$) is similar since in this case θ_j uniformly covers $\mathbb{Z}_p$. Now suppose $p > w(N)$ and $|X| = 2$. We claim that none of the s pure linear forms $W(\psi_i - b_i)$ is a multiple of any other $(\mathrm{mod}\, p)$. Indeed, if this were so then we should have $L_{ij} L_{i'j}^{-1} \equiv \lambda (\mathrm{mod}\, p)$ for some λ, and for all $j = 1, \dots, t$. 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 $a/q \not\equiv a'/q' (\mathrm{mod}\, p)$ unless $a = a'$, $q = q'$. It follows that the two pure linear forms $\psi_i - b_i$ and $\psi_{i'} - b_{i'}$ are rational multiples of one another, contrary to assumption. Thus the set of $\mathrm{x} \in (\mathbb{Z}/p\mathbb{Z})^t$ for which $\theta_i(\mathrm{x}) \equiv 0 (\mathrm{mod}\, p)$ for all $i \in X$ is contained in the intersection of two skew affine subspaces of $(\mathbb{Z}/p\mathbb{Z})^t$, and as such has cardinality at most p^{t-2}. □

This lemma implies, comparing with (10.7), that

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

$$+ \mathbf{1}_{p>w(N)} \sum_{\substack{X,X' \subseteq \{1,\ldots,m\} \\ |X \cup X'| \geqslant 2}} \frac{O(1/p^2)}{p^{\sum_{j \in X} z_j + \sum_{j \in X'} z_j'}}, \tag{10.8}$$

where the $O(1/p^2)$ numerator does not depend on z, z'. To take advantage of this expansion, we factorise $E_p = E_p^{(1)} E_p^{(2)} E_p^{(3)}$, where

$$E_p^{(1)}(z,z') := \frac{E_p(z,z')}{\prod_{j=1}^m (1 - \mathbf{1}_{p>w(N)} p^{-1-z_j})(1 - \mathbf{1}_{p>w(N)} p^{-1-z_j'})(1 - \mathbf{1}_{p>w(N)} p^{-1-z_j-z_j'})^{-1}}$$

$$E_p^{(2)}(z,z') := \prod_{j=1}^m (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j})^{-1} (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j'})^{-1} (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j-z_j'})$$

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

Writing $G_j := \prod_p E_p^{(j)}$ for $j = 1,2,3$, one thus has $F = G_1 G_2 G_3$ (at least for $\Re(z_j), \Re(z_j')$ sufficiently large). If we introduce the Riemann ζ-function $\zeta(s) := \prod_p (1 - \frac{1}{p^s})^{-1}$ then we have

$$G_3(z,z') = \prod_{j=1}^m \frac{\zeta(1 + z_j + z_j')}{\zeta(1+z_j)\zeta(1+z_j')} \tag{10.9}$$

so in particular G_3 can be continued meromorphically to all of $\mathbb{C}^{2m}$. As for the other two factors, we have the following estimates which allow us to continue these factors a little bit to the left of the imaginary axes.

Definition 10.2. For any $\sigma > 0$, let $\mathcal{D}_\sigma^m \subseteq \mathbb{C}^{2m}$ denote the domain

$$\mathcal{D}_\sigma^m := \{z_j, z_j' : -\sigma < \Re(z_j), \Re(z_j') < 100, j = 1, \ldots, m\}.$$

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

$$\|G\|_{C^k(\mathcal{D}_\sigma^m)} := \sup_{a_1,\ldots,a_m,a_1',\ldots,a_m'} \Big\| \big(\frac{\partial}{\partial z_1}\big)^{a_1} \cdots \big(\frac{\partial}{\partial z_m}\big)^{a_m} \big(\frac{\partial}{\partial z_1'}\big)^{a_1'} \cdots \big(\frac{\partial}{\partial z_m'}\big)^{a_m'} G \Big\|_{L^\infty(\mathcal{D}_\sigma^m)}$$

where $a_1, \ldots, a_m, a_1', \ldots, a_m'$ range over all non-negative integers with total sum at most k.

Lemma 10.3. *The Euler products $\prod_p E_p^{(j)}$ for $j = 1,2$ are absolutely convergent in the domain $\mathcal{D}_{1/6m}^m$. In particular, G_1, G_2 can be continued analytically to this domain. Furthermore, we have the estimates*

$$\|G_1\|_{C^m(\mathcal{D}_{1/6m}^m)} \leqslant O_m(1)$$
$$\|G_2\|_{C^m(\mathcal{D}_{1/6m}^m)} \leqslant O_{m,w(N)}(1)$$
$$G_1(0,0) = 1 + o_m(1)$$
$$G_2(0,0) = (W/\phi(W))^m.$$

Remark. The choice $\sigma = 1/6m$ is of course not best possible, but in fact any small positive quantity depending on m would suffice for our argument here. The dependence of $O_{m,w(N)}(1)$ on $w(N)$ is not important, but one can easily obtain (for instance) growth bounds of the form $w(N)^{O_m(w(N))}$.

Proof. First consider $j = 1$. From (10.8) and Taylor expansion we have the crude bound $E_p^{(1)}(z, z') = 1 + O_m(p^{-2+4/6m})$ in $\mathcal{D}^m_{1/6m}$, which gives the desired convergence and also the $C^m(\mathcal{D}^m_{1/6m})$ bound on G_1; the estmate for $G_1(0,0)$ also follows since the Euler factors $E_p^{(1)}(z, z')$ are identically 1 when $p \leqslant w(N)$. The bound for G_2 are easy since this is just a finite Euler product involving at most $w(N)$ terms; the formula for $G_2(0,0)$ follows from direct calculation since $\frac{\phi(W)}{W} = \prod_{p<w(n)}(1 - \frac{1}{p})$. □

To estimate (10.5), we now invoke the following contour integration lemma.

Lemma 10.4. [17] *Let R be a positive real number. Let $G = G(z, z')$ be an analytic function of $2m$ complex variables on the domain $\mathcal{D}^m_\sigma$ for some $\sigma > 0$, and suppose that*

$$\|G\|_{C^m(\mathcal{D}^m_\sigma)} = \exp(O_{m,\sigma}(\log^{1/3} R)). \tag{10.10}$$

Then

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

$$= G(0, \ldots, 0) \log^m R + \sum_{j=1}^m O_{m,\sigma}(\|G\|_{C^j(\mathcal{D}^m_\sigma)} \log^{m-j} R) + O_{m,\sigma}(e^{-\delta\sqrt{\log R}})$$

for some $\delta = \delta(m) > 0$.

Proof. While this lemma is essentially in [17], we shall give a complete proof in the Appendix for sake of completeness. □

We apply this lemma with $G := G_1 G_2$ and $\sigma := 1/6m$. From Lemma 10.3 and the Leibnitz rule we have the bounds

$$\|G\|_{C^j(\mathcal{D}^m_{1/6m})} \leqslant O_{j,m,w(N)}(1) \text{ for all } 0 \leqslant j \leqslant m,$$

and in particular we obtain (10.10) by choosing $w(N)$ to grow sufficiently slowly in N. Also we have $G(0,0) = (1+o_m(1))(\frac{W}{\phi(W)})^m$ from that lemma. We conclude (again taking $w(N)$ sufficiently slowly growing in N) that the quantity in (10.5) is $(1+o_m(1))(\frac{W \log R}{\phi(W)})^m$, as desired. This concludes the proof of Proposition 9.5. □

Higher order correlations for Λ_R. We now prove Proposition 9.6, using arguments very similar to those used to prove Proposition 9.5. The main differences here are that the number of variables t is just equal to 1, but on the other hand all the linear forms are equal to each other, $\psi_i(x_1) = x_1$. In particular, these linear forms are now rational multiples of each other and so Lemma 10.1 no longer applies. However, the arguments before that Lemma are still valid; thus we can still write the left-hand side of (9.1) as an expression of the form (10.5) plus an acceptable error, where F is again defined by (10.6) and E_p is defined by (10.7); the difference now is that $\omega_X(p)$ is the quantity

$$\omega_X(p) := \mathbb{E}\bigg(\prod_{i \in X} 1_{W(x+h_i)+1 \equiv 0 (\mathrm{mod}\, p)} \,\bigg|\, x \in \mathbb{Z}_p \bigg).$$

Again we have $\omega_\emptyset(p) = 1$ for all p. The analogue of Lemma 10.1 is as follows.

Lemma 10.5. *If $p \leqslant w(N)$, then $\omega_X(p) = 0$ for all non-empty X; in particular, $E_p = 1$ when $p \leqslant w(N)$. If instead $p > w(N)$, then $\omega_X(p) = p^{-1}$ when $|X| = 1$ and $\omega_X(p) \leqslant p^{-1}$ when $|X| \geqslant 2$. Furthermore, if $|X| \geqslant 2$ then $\omega_X(p) = 0$ unless p divides $\Delta := \prod_{1 \leqslant i < j \leqslant s} |h_i - h_j|$.*

Proof. When $p \leqslant w(N)$ then $W(x + h_i) + 1 \equiv 1 (\operatorname{mod} p)$ and the claim follows. When $p > w(N)$ and $|X| \geqslant 1$, $\omega_X(p)$ is equal to $1/p$ when the residue classes $\{h_i (\operatorname{mod} p) : i \in X\}$ are all equal, and zero otherwise, and the claim again follows. □

In light of this lemma, the analogue of (10.8) is now

$$E_p(z, z') = 1 - \mathbf{1}_{p > w(N)} \sum_{j=1}^{m} (p^{-1-z_j} + p^{-1-z_j'} - p^{-1-z_j-z_j'}) + \mathbf{1}_{p > w(N), p | \Delta} \lambda_p(z, z') \quad (10.11)$$

where $\lambda_p(z, z')$ is an expression of the form

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

and the $O(1/p)$ quantities do not depend on z, z'. We can thus factorise

$$E_p = E_p^{(0)} E_p^{(1)} E_p^{(2)} E_p^{(3)},$$

where

$$E_p^{(0)} = 1 + \mathbf{1}_{p > w(N), p | \Delta} \lambda_p(z, z')$$
$$E_p^{(1)} = \frac{E_p}{E_p^{(0)} \prod_{j=1}^{m} (1 - \mathbf{1}_{p > w(N)} p^{-1-z_j})(1 - \mathbf{1}_{p > w(N)} p^{-1-z_j'})(1 - \mathbf{1}_{p > w(N)} p^{-1-z_j-z_j'})^{-1}}$$
$$E_p^{(2)} = \prod_{j=1}^{m} (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j})^{-1} (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j'})^{-1} (1 - \mathbf{1}_{p \leqslant w(N)} p^{-1-z_j-z_j'})$$
$$E_p^{(3)} = \prod_{j=1}^{m} (1 - p^{-1-z_j})(1 - p^{-1-z_j'})(1 - p^{-1-z_j-z_j'})^{-1}.$$

Write $G_j = \prod_p E_p^{(j)}$. Then, as before, $F = G_0 G_1 G_2 G_3$ and G_3 is given by (10.9) as before. As for G_0, G_1, G_2, we have the following analogue of Lemma 10.3.

Lemma 10.6. *Let $0 < \sigma < 1/6m$. Then the Euler products $\prod_p E_p^{(l)}$ for $l = 0, 1, 2$ are absolutely convergent in the domain $\mathcal{D}_\sigma^m$. In particular, G_0, G_1, G_2 can be continued* *analytically to this domain. Furthermore, we have the estimates*

$$\|G_0\|_{C^r(\mathcal{D}_\sigma^m)} \leqslant O_m\left(\frac{\log R}{\log\log R}\right)^r \prod_{p|\Delta}(1+O_m(p^{2m\sigma-1})) \quad \textit{for } 0 \leqslant r \leqslant m \tag{10.12}$$

$$\|G_0\|_{C^m(\mathcal{D}^m_{1/6m})} \leqslant \exp(O_m(\log^{1/3} R)) \tag{10.13}$$

$$\|G_1\|_{C^m(\mathcal{D}^m_{1/6m})} \leqslant O_m(1)$$

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

$$G_0(0,0) = \prod_{p|\Delta}(1+O_m(p^{-1/2})) \tag{10.14}$$

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

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

Proof. The estimates for G_1 and G_2 proceed exactly as in Lemma 10.3 (the additional factors of $\lambda_p(z,z')$ which appear on both the numerator and denominator of $E_p^{(1)}$ cancel to first order, and thus do not present any new difficulties); it is the estimates for G_0 which are the most interesting.

We begin by proving (10.12). Fix l. First observe that $G_0 = \prod_{p|\Delta} E_p^{(0)}$. Now the number of primes dividing Δ is at most $O(\log\Delta/\log\log\Delta)$. Using the crude bound

$$\Delta = \prod_{1\leqslant i<j\leqslant m} |h_i - h_j| \leqslant N^{m^2} \leqslant R^{O_m(1)}, \tag{10.15}$$

we thus see that the number of factors in the Euler product for G_0 is $O_m(\frac{\log R}{\log\log R})$. Upon differentiating r times for any $0 \leqslant r \leqslant m$ using the Leibnitz rule, one gets a sum of $O_m((\log R/\log\log R)^r)$ terms, each of which consists of $O_m(\log R/\log\log R)$ factors, each of which is equal to some derivative of $1+\lambda_p(z,z')$ of order between 0 and r. On $\mathcal{D}_\sigma^m$, each factor is bounded by $1+O_m(p^{2m\sigma-1})$ (in fact, the terms containing a non-zero number of derivatives will be much smaller since the constant term 1 is eliminated). This gives (10.12).

Now we prove (10.13). In light of (10.12), it suffices to show that

$$\prod_{p|\Delta}(1+O_m(p^{2m\sigma-1})) \leqslant \exp(O_m(\log^{1/3} R)).$$

Taking logarithms and using the hypothesis $\sigma < 1/6m$ (and (10.15)), we reduce to showing

$$\sum_{p|\Delta} p^{-2/3} \leqslant O(\log^{1/3}\Delta).$$

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

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

as desired.

The bound (10.14) now follows from the crude estimate $E_p^{(0)}(z,z') = 1 + O_m(p^{-1/2})$. □

We now apply Lemma 10.4 with $\sigma := 1/6m$ and $G := G_0G_1G_2$. Again by the Leibnitz rule we have the bound (10.10), and furthermore

$$\|G\|_{C^r(\mathcal{D}^m_\sigma)} \leqslant O_m(1)O_{m,w(N)}(1)\left(\frac{\log R}{\log\log R}\right)^r \prod_{p|\Delta}\left(1+O_m(p^{-1/2})\right).$$

for all $0 \leqslant r \leqslant m$. From Lemma 10.6 and Lemma 10.4 we can then estimate (10.5) as

$$\leqslant (1+o_m(1))\left(\frac{W}{\phi(W)}\right)^m \log^m R \prod_{p|\Delta}\left(1+O_m(p^{-1/2})\right)$$
$$+ O_{m,w(N)}\left(\frac{\log^m R}{\log\log R}\right)\prod_{p|\Delta}\left(1+O_m(p^{-1/2})\right) + O_m(e^{-\delta\sqrt{\log R}}).$$

The claim (9.1) then follows by choosing $w(N)$ (and hence W) sufficiently slowly growing in N (and hence in R). Proposition 9.6 follows. □

Remark. It should be clear that the above argument not only gives an upper bound for the left-hand side of (9.1), but in fact gives an asymptotic, by working out $G_0(0,0)$ more carefully; this is worked out in detail (in the $W = 1$ case) in [17].

11. Further remarks

In this section we discuss some extensions and refinements of our main result. First of all, notice that our proof actually shows that that there is some constant $\gamma(k)$ such that the number of k-term progressions of primes, all less than N, is at least $(\gamma(k) + o(1))N^2/\log^k N$. This is because the error term in (3.9) does not actually need to be $o(1)$, but merely less than $\frac{1}{2}c(k,\delta) + o(1)$ (for instance). Working backwards through the proof, this eventually reveals that the quantity $w(N)$ does not actually need to be growing in N, but can instead be a fixed number depending only on k (although this number will be very large because our final bounds $o(1)$ decayed to zero extremely slowly). Thus W can be made independent of N, and so the loss incurred by the W-trick when passing from primes to primes equal to 1 mod W is bounded uniformly in N. Nevertheless the bound we obtain on $\gamma(k)$ is extremely poor, in part because of the growth of constants in the best known bounds $c(k,\delta)$ on Szemerédi's theorem in [19], but also because we have not attempted to optimise the decay rate of the $o(1)$ factors and hence will need to take $w(N)$ to be extremely large. In the other direction, standard sieve theory arguments show that the number of k-term progressions of primes all less than N are at most $O_k(N^2/\log^k N)$, and so the lower bounds are only off by a constant depending on k.

As we remarked earlier, our method also extends to prove Theorem 1.2, namely that any subset of the primes with positive relative upper density contains a k-term arithmetic progression. The only significant change[23] to the proof is that one must use the pigeonhole principle to replace the residue class $n \equiv 1 (\mathrm{mod}\, W)$ by a more general residue class $n \equiv b (\mathrm{mod}\, W)$ for some b coprime to W, since the set A in Theorem 1.2 does not need

[23] Also, since we are only assuming positivity of the upper density and not the lower density, we only have good density control for an infinite sequence $N_1, N_2, \ldots \to \infty$ of integers, which may not be prime. However one can easily use Bertrands postulate (for instance) to make the N_j prime, giving up a factor of $O(1)$ at most.

to obey a Dirichlet-type theorem in these residue classes. However it is easy to verify that this does not significantly affect the rest of the argument, and we leave the details to the reader.

Applying Theorem 1.2 to the set of primes $p \equiv 1 (\mathrm{mod}\, 4)$, we obtain the previously unknown fact that there are arbitrarily long progressions consisting of numbers which are the sum of two squares. For this problem, more satisfactory results were known for small k than was the case for the primes. Let S be the set of sums of two squares. It is a simple matter to show that there are infinitely many 4-term arithmetic progressions in S. Indeed, Heath-Brown [26] observed that the numbers $(n-1)^2+(n-8)^2$, $(n-7)^2+(n+4)^2$, $(n+7)^2+(n-4)^2$ and $(n+1)^2+(n+8)^2$ always form such a progression; in fact, he was able to prove much more, in particular finding an asymptotic for the number of 4-term progressions in S, all of whose members are at most N (weighted by $r(n)$, the number of representations of n as the sum of two squares).

It is reasonably clear that our method will produce long arithmetic progressions for many sets of primes for which one can give a lower bound which agrees with some upper bound coming from a sieve, up to a multiplicative constant. Invoking Chen's famous theorem[7] to the effect that there are $\gg N/\log^2 N$ primes $p \leqslant N$ for which $p+2$ is a prime or a product of two primes, it ought to be a simple matter to adapt our arguments to show that there are arbitrarily long arithmetic progressions $p_1, \dots, p_k$ of primes, such that each $p_i + 2$ is either prime or the product of two primes; indeed there should be $N/\log^{2k} N$ such progressions with entries less than N. Whilst we do not plan[24] to write a detailed proof of this fact, we will in [23] give a proof of the case $k = 3$ using harmonic analysis.

The methods in this paper suggest a more general "transference principle", in that if a type of pattern (such as an arithmetic progression) is forced to arise infinitely often within sets of positive density, then it should also be forced to arise infinitely often inside the prime numbers, or more generally inside any subset of a pseudorandom set (such as the "almost primes") of positive relative density. Thus, for instance, one is led to conjecture a Bergelson-Leibman type result (cf. [4]) for primes. That is, one could hope to show that if $F_i : \mathbb{N} \to \mathbb{N}$ are polynomials with $F(0) = 0$, then there are infinitely many configurations $(a + F_1(d), \dots, a + F_k(d))$ in which all k elements are prime. This however seems to require some modification[25] to our current argument, in large part because of the need to truncate the step parameter d to be at most a small power of N. In a similar spirit, the work of Furstenberg and Katznelson [11] on multidimensional analogues of Szemerédi's theorem, combined with this transference principle, now suggests that one should be able to show[26] that the Gaussian primes in $\mathbb{Z}[i]$ contain infinitely many constellations of any prescribed shape, and similarly for

[24]Very briefly, the idea is to replace the function $\Lambda_R(Wn+1)$ in the definition of the pseudorandom measure ν with a variant such as $\Lambda_R(Wn+b)\Lambda_R(Wn+b+2)$ for some $1 \leqslant b < W$ for which $b, b+2$ are both coprime to W; one can use Chen's theorem and the pigeonhole principle to locate a b for which this majorant will capture a large number of Chen primes. We leave the details to the reader.

[25]Note added in press: such a result has been obtained by the second author and T. Ziegler, to appear in *Acta Math.*

[26]Note added in press: such a result has been obtained by the second author, *J. d'Analyse Mathématique* 99 (2006), 109–176.

other number fields. Furthermore, the later work of Furstenberg and Katznelson [12] on density Hales-Jewett theorems suggests that one could also show that for any finite field F, the monic irreducible polynomials in $F[t]$ contain affine subspaces over F of arbitrarily high dimension. Again, these results would require non-trivial modifications to our argument for a number of reasons, not least of which is the fact that the characteristic factors for these more advanced generalizations of Szemerédi's theorem are much less well understood.

Appendix A. Proof of Lemma 10.4

In this appendix we prove Lemma 10.4. This Lemma was essentially proven in [17], but for the sake of self-containedness we provide a complete proof here (following very closely the approach in [17]).

Throughout this section, $R \geqslant 2$, $m \geqslant 1$, and $\sigma > 0$ will be fixed. We shall use $\delta > 0$ to denote various small constants, which may vary from line to line (the previous interpretation of δ as the average value of a function f will now be irrelevant). We begin by recalling the classical zero-free region for the Riemann ζ function.

Lemma A.1 (Zero-free region). *Define the classical zero free region $\mathcal{Z}$ to be the closed region*

$$\mathcal{Z} := \{s \in \mathbb{C} : 10 \geqslant \Re s \geqslant 1 - \frac{\beta}{\log(|\Im s| + 2)}\}$$

for some small $0 < \beta < 1$. Then if β is sufficiently small, ζ is non-zero and meromorphic in $\mathcal{Z}$ with a simple pole at 1 and no other singularities. Furthermore we have the bounds

$$\zeta(s) - \frac{1}{s-1} = O(\log(|\Im s| + 2)); \quad \frac{1}{\zeta(s)} = O(\log(|\Im s| + 2))$$

for all $s \in \mathcal{Z}$.

Proof. See Titchmarsh [41, Chapter 3]. □

Fix β in the above lemma; we may take β to be small enough that $\mathcal{Z}$ is contained in the region where $1 - \sigma < \Re(s) < 101$. We will allow all our constants in the $O()$ notation to depend on β and σ, and omit explicit mention of these dependencies from our subscripts.

In addition to the contour Γ_1 defined in (10.4), we will need the two further contours Γ_0 and Γ_2, defined by

$$\begin{aligned} \Gamma_0(t) &:= -\frac{\beta}{\log(|t|+2)} + it, \qquad -\infty < t < \infty \\ \Gamma_2(t) &:= 1 + it, \quad -\infty < t < \infty. \end{aligned} \tag{A.1}$$

Thus Γ_0 is the left boundary of $\mathcal{Z} - 1$ (which therefore lies to the left of the origin), while Γ_1 and Γ_2 are vertical lines to the right of the origin. The usefulness of Γ_2 for us lies in the simple observation that $\zeta(1 + z + z')$ has no poles when $z \in \mathcal{Z} - 1$ and $z' \in \Gamma_2$, but we will not otherwise attempt to estimate any integrals on Γ_2.

We observe the following elementary integral estimates.

Lemma A.2. *Let A, B be fixed constants with $A > 1$. Then we have the bounds.*

$$\int_{\Gamma_0} \log^B(|z|+2) \left|\frac{R^z dz}{z^A}\right| \leqslant O_{A,B}(e^{-\delta\sqrt{\log R}}); \tag{A.2}$$

$$\int_{\Gamma_1} \log^B(|z|+2) \left|\frac{R^z dz}{z^2}\right| \leqslant O_B(\log R). \tag{A.3}$$

Here $\delta = \delta(A, B, \beta) > 0$ is a constant independent of R.

Proof. We first bound the left-hand side of (A.2). Substitute in the parametrisation (A.1). Since $\Gamma_0'(t) = O(1)$ and $|z| \gg |t| + \beta$ we have, for any $T \geqslant 2$,

$$\begin{aligned}\int_{\Gamma_0} \log^B(|z|+2) \left|\frac{R^z dz}{z^A}\right| &\leqslant O_B\Big(\int_0^\infty R^{-\beta/(\log(|t|+2))} \frac{\log^B(|t|+2)}{(|t|+\beta)^A} dt\Big) \\ &\leqslant O_B\Big(\log^B T \int_0^T R^{-\beta/\log(t+2)} dt + \int_T^\infty \frac{\log^B t}{t^A} dt\Big) \\ &\leqslant O_{A,B}(T \log^B T \exp(-\beta \log R/\log T) + T^{1-A} \log^B T).\end{aligned}$$

Choosing $T = \exp(\sqrt{\beta \log R/2})$ one obtains the claimed bound. The bound (A.3) is much simpler, and can be obtained by noting that R^z is bounded on Γ_1, and substituting in (10.4) splitting the integrand up into the ranges $|t| \leqslant 1/\log R$ and $|t| > 1/\log R$. □

The next lemma is closely related to the case $m = 1$ of Lemma 10.4.

Lemma A.3. *Let $f(z, z')$ be analytic in $\mathcal{D}_\sigma^1$ and suppose that*

$$|f(z, z')| \leqslant \exp(O_m(\log^{1/3} R))$$

uniformly in this domain. Then the integral

$$I := \frac{1}{(2\pi i)^2} \int_{\Gamma_1} \int_{\Gamma_1} f(z, z') \frac{\zeta(1+z+z')}{\zeta(1+z)\zeta(1+z')} \frac{R^{z+z'}}{z^2 z'^2} dz dz'$$

obeys the estimate

$$I = f(0,0) \log R + \frac{\partial f}{\partial z'}(0,0) + \frac{1}{2\pi i} \int_{\Gamma_1} f(z, -z) \frac{dz}{\zeta(1+z)\zeta(1-z) z^4} + O_m(e^{-\delta\sqrt{\log R}})$$

for some $\delta = \delta(\sigma, \beta) > 0$ independent of R.

Proof. We observe from Lemma A.1 that we have enough decay of the integrand in the domain $\mathcal{D}_\sigma^1$ to interchange the order of integration, and to shift contours in either one of the variables z, z' while keeping the other fixed, without any difficulties when $\Im(z), \Im(z') \to \infty$; the only issue is to keep track of when the contour passes through a pole of the integrand. In particular we can shift the z' contour from Γ_1 to Γ_2, since we do not encounter any of the poles of the integrand while doing so. Let us look at the integrand for each fixed $z' \in \Gamma_2$, viewing it as an analytic function of z. We now attempt to shift the z contour of integration to Γ_0. In so doing the contour passes just one pole, a simple one at $z = 0$. The residue there is $\frac{1}{2\pi i} \int_{\Gamma_2} f(0, z') \frac{R^{z'}}{z'^2} dz'$, and so we have $I = I_1 + I_2$, where

$$I_1 := \frac{1}{2\pi i}\int_{\Gamma_2} f(0,z')\frac{R^{z'}}{z'^2}dz'$$

$$I_2 := \frac{1}{(2\pi i)^2}\int_{\Gamma_2}\int_{\Gamma_0} f(z,z')\frac{\zeta(1+z+z')R^{z+z'}}{\zeta(1+z)\zeta(1+z')z^2z'^2}dzdz'.$$

To evaluate I_1, we shift the z' contour of integration to Γ_0. Again there is just one pole, a double one at $z' = 0$. The residue there is $f(0,0)\log R + \frac{\partial f}{\partial z'}(0,0)$, and so

$$\begin{aligned} I_1 &= f(0,0)\log R + \frac{\partial f}{\partial z'}(0,0) + \frac{1}{2\pi i}\int_{\Gamma_0} f(0,z')\frac{R^{z'}}{z'^2}dz' \\ &= f(0,0)\log R + \frac{\partial f}{\partial z'}(0,0) + O_m(e^{-\delta\sqrt{\log R}}), \end{aligned}$$

for some $\delta > 0$, the latter step being a consequence of our bound on f and (A.2) (in the case $B = 0$).

To estimate I_2, we first swap the order of integration and, for each fixed z, view the integrand as an analytic function of z'. We move the z' contour from Γ_2 to Γ_0, this again being allowed since we have sufficient decay in vertical strips as $|\Im z'| \to \infty$. In so doing we pass exactly two simple poles, at $z' = -z$ and $z' = 0$. The residue at the first is exactly

$$\frac{1}{2\pi i}\int_{\Gamma_0} f(z,-z)\frac{dz}{\zeta(1+z)\zeta(1-z)z^4},$$

which is one of the terms appearing in our formula for I.

The residue at $z' = 0$ is

$$\int_{\Gamma_0} f(z,0)\frac{R^z}{z^2}dz,$$

which is $O(e^{-\delta\sqrt{\log R}})$ for some $\delta > 0$ by (A.2). The value of I_2 is the sum of these two quantities and the integral over the new contour Γ_0, which is

$$\int_{\Gamma_0}\int_{\Gamma_0} f(z,z')\frac{\zeta(1+z+z')R^{z+z'}}{\zeta(1+z)\zeta(1+z')z^2z'^2}dzdz'. \tag{A.4}$$

In this integrand we have $|f| = \exp(O_m(\log^{1/3} R))$ and, by Lemma A.1, $1/|\zeta(1+z)| \ll \log(|\Im z|+2)$ and $1/|\zeta(1+z')| \ll \log(|\Im z'|+2)$. Assume that $\beta < 1/10$, as we obviously may. We claim that

$$|\zeta(1+z+z')| \ll (1+|z|+|z'|)^{1/4} \ll (1+|z|)^{1/4}(1+|z'|)^{1/4} \tag{A.5}$$

for all $z, z' \in \Gamma_0$. Once this is proven it follows from (A.2), applied with $A = 7/4$ and $A = 2$, that the integral (A.4) is bounded by $O_m(e^{-\delta\sqrt{\log R}})$ for some $\delta > 0$. Now if $1/2 \leqslant \sigma \leqslant 1$ and $|t| \geqslant 1/100$ we have the convexity bound $|\zeta(\sigma+it)| \ll_\epsilon |t|^{1-\sigma+\epsilon}$ (cf. [41, Chapter V]), and so (A.5) is indeed true provided that $|\Im(z+z')| \geqslant 1/100$. However since $z, z' \in \Gamma_0$ one may see that if $|\Im(z)|, |\Im(z')| \leqslant t$ then $|z+z'| \gg 1/\log(t+2)$. It follows from Lemma A.1 that (A.5) holds when $|\Im(z+z')| \leqslant 1/100$ as well.

Thus we now have estimates for I_1 and I_2 up to errors of $O_m(e^{-\delta\sqrt{\log R}})$. Putting all of this together completes the proof of the lemma. □

Proof of Lemma 10.4. Let $G = G(z, z')$ be an analytic function of $2m$ complex variables on the domain $\mathcal{D}_\sigma^m$ obeying the derivative bounds (10.10). We will allow all our implicit constants in the $O()$ notation to depend on m, β, σ. We are interested in the integral

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

and wish to prove the estimate

$$I(G,m) := G(0,\ldots,0)(\log R)^m + \sum_{j=1}^{m} O(\|G\|_{C^j(\mathcal{D}_\sigma^s)}(\log R)^{m-j}) + O(e^{-\delta\sqrt{\log R}}).$$

The proof is by induction on m. The case $m = 1$ is a swift deduction from Lemma A.3, the only issue being an estimation of the term

$$\frac{1}{2\pi i}\int_{\Gamma_0} G(z_1,-z_1)\frac{dz_1}{\zeta(1+z_1)\zeta(1-z_1)z_1^4}.$$

It is not hard to check (using Lemma A.1) that

$$\int_{\Gamma_0} \left|\frac{dz_1}{\zeta(1+z_1)\zeta(1-z_1)z_1^4}\right| = O(1), \tag{A.6}$$

and so this term is $O(\sup_{z\in\mathcal{D}_\sigma^1}|G(z)|) = O(\|G\|_{C^1(\mathcal{D}_\sigma^1)})$.

Suppose then that we have established the result for $m \geqslant 1$ and wish to deduce it for $m+1$. Applying Lemma A.3 in the variables z_{m+1}, z_{m+1}', we get $I(G, m+1) =$

$$\begin{aligned}
&\frac{\log R}{(2\pi i)^{2m}} \int_{\Gamma_1} \cdots \int_{\Gamma_1} G(z_1,\ldots,z_m,0,z_1',\ldots,z_m',0) \prod_{j=1}^{m} \frac{\zeta(1+z_j+z_j')}{\zeta(1+z_j)\zeta(1+z_j')} \frac{R^{z_j+z_j'}}{z_j^2 z_j'^2} dz_j dz_j' \\
&+ \frac{1}{(2\pi i)^{2m}} \int_{\Gamma_1} \cdots \int_{\Gamma_1} H(z_1,\ldots,z_m,z_1',\ldots,z_m') \prod_{j=1}^{m} \frac{\zeta(1+z_j+z_j')}{\zeta(1+z_j)\zeta(1+z_j')} \frac{R^{z_j+z_j'}}{z_j^2 z_j'^2} dz_j dz_j' \\
&+ O(e^{-\delta\sqrt{\log R}}) \\
=& I(G(z_1,\ldots,z_m,0,z_1',\ldots,z_m',0),m)\log R + I(H,m) + O(e^{-\delta\sqrt{\log R}})
\end{aligned}$$

where $\delta > 0$ and $H : \mathcal{D}_\sigma^m \to \mathbb{C}$ is the function

$$\begin{aligned}
H(z_1,\ldots,z_m,z_1',\ldots,z_m') &:= \frac{\partial G}{\partial z_{m+1}'}(z_1,\ldots,z_m,0,z_1',\ldots,z_m',0) \\
&+ \frac{1}{2\pi i}\int_{\Gamma_0} G(z_1,\ldots,z_m,z_{m+1},z_1',\ldots,z_m',-z_{m+1})\frac{dz_{m+1}}{\zeta(1+z_{m+1})\zeta(1-z_{m+1})z_{m+1}^4}.
\end{aligned}$$

The error term $O(e^{-\delta\sqrt{\log R}})$ which we claim here arises by applying (10.10) and several applications of (A.3).

Now both of the functions $G(z_1,\ldots,z_m,0,z_1',\ldots,z_m',0)$ and $H(z_1,\ldots,z_m,z_1',\ldots,z_m')$ are analytic on $\mathcal{D}_\sigma^m$ and (appealing to (A.6)) we have $\|H\|_{C^j(\mathcal{D}_\sigma^m)} = O_m(\|G\|_{C^{j+1}(\mathcal{D}_\sigma^{m+1})})$

for $0 \leqslant j \leqslant m$. Using the inductive hypothesis, we therefore obtain $I(G, m+1) =$

$$\begin{aligned} & G(0,\ldots,0)(\log R)^{m+1} + \sum_{j=1}^{m} O_m(\|G(\cdot,0,\cdot,0)\|_{C^j(\mathcal{D}^m_\sigma)}(\log R)^{m+1-j}) \\ & + H(0,\ldots,0)(\log R)^m + \sum_{j=1}^{m} O_m(\|H\|_{C^j(\mathcal{D}^m_\sigma)}(\log R)^{m-j}) + O(e^{-\delta\sqrt{\log R}}) \\ = \; & G(0,\ldots,0)(\log R)^{m+1} + \sum_{j=1}^{m} O_m(\|G\|_{C^j(\mathcal{D}^{m+1}_\sigma)}(\log R)^{m+1-j}) \\ & + H(0,\ldots,0)(\log R)^m + \sum_{j=1}^{m} O_m(\|G\|_{C^{j+1}(\mathcal{D}^{m+1}_\sigma)}(\log R)^{m-j}) + O(e^{-\delta\sqrt{\log R}}) \\ = \; & G(0,\ldots,0)(\log R)^{m+1} + \sum_{j=1}^{m+1} O_m(\|G\|_{C^j(\mathcal{D}^{m+1}_\sigma)}(\log R)^{m+1-j}) + O(e^{-\delta\sqrt{\log R}}), \end{aligned}$$

which is what we wanted to prove. □

References

[1] I. Assani, *Pointwise convergence of ergodic averages along cubes*, preprint.

[2] A. Balog, *Linear equations in primes,* Mathematika **39** (1992) 367–378.

[3] ———, *Six primes and an almost prime in four linear equations,* Can. J. Math. **50** (1998), 465–486.

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

[5] J. Bourgain, *A Szemerédi-type theorem for sets of positive density in* $\mathbb{R}^k$, Israel J. Math **54** (1986), no. 3, 307–316.

[6] ———, *On triples in arithmetic progression*, GAFA **9** (1999), 968–984.

[7] 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.

[8] S. Chowla, *There exists an infinity of 3—combinations of primes in A. P.*, Proc. Lahore Philos. Soc. **6**, (1944). no. 2, 15–16.

[9] P. Erdős, P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261–264.

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

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

[12] H. Furstenberg, Y. Katznelson, *A density version of the Hales-Jewett theorem*, J. d'Analyse Math. **57** (1991), 64–119.

[13] H. Furstenberg, Y. Katznelson and D. Ornstein, *The ergodic-theoretical proof of Szemerédi's theorem,* Bull. Amer. Math. Soc. **7** (1982), 527–552.

[14] H. Furstenberg, B. Weiss, *A mean ergodic theorem for* $1/N \sum_{n=1}^{N} f(T^n x) g(T^{n^2} x)$, Convergence in ergodic theory and probability (Columbus OH 1993), 193–227, Ohio State Univ. Math. Res. Inst. Publ., 5. de Gruyter, Berlin, 1996.

[15] D. Goldston and C.Y. Yıldırım *Higher correlations of divisor sums related to primes, I: Triple correlations,* Integers **3** (2003) A5, 66pp.

[16] ———, *Higher correlations of divisor sums related to primes, III: k-correlations,* preprint (available at AIM preprints)

[17] ———, *Small gaps between primes, I,* preprint available at `http://www.arxiv.org/abs/math/0504336`.