CORRELATIONS OF THE DIVISOR FUNCTION

LILIAN MATTHIESEN

ABSTRACT. Let $\tau(n) = \sum_d 1_{d|n}$ denote the divisor function. Based on Erdős's fundamental work on sums of multiplicative functions evaluated over polynomials, we construct a pseudorandom majorant for a slightly smoothed version of τ. By means of the nilpotent Hardy-Littlewood method we give an asymptotic for the following correlation

$$\mathbb{E}_{n\in[-N,N]^d\cap K} \prod_{i=1}^{t} \tau(\psi_i(n)) \ ,$$

where $\Psi = (\psi_1, \ldots, \psi_t)$ is a non-degenerated system of affine-linear forms no two of which are affinely related, and where K is a convex body.

CONTENTS

1. INTRODUCTION

Questions concerning the distribution of the values of elementary arithmetic functions play a central role in analytic number theory. We mention two classes of such questions for multiplicative functions, both of which are related to the results of this paper.

The first class concerns asymptotics for sums

$$\sum_{M\leqslant n\leqslant N+M} f(|P(n)|)$$

of multiplicative functions evaluated over polynomials, a direction which has been substantially influenced by Erdős's work on the sum $\sum \tau(P(n))$, see [3]. We shall employ some ideas introduced in that paper. For newer work on this type of question, see for instance [13] and the references therein.

A second class considers linear correlations of multiplicative functions. Let $f : [N] \to \mathbb{R}$ be multiplicative and let $\psi_1, \ldots, \psi_t : \mathbb{Z}^d \to \mathbb{Z}$ be affine-linear forms, then these problems ask for an asymptotic to the correlation

$$\sum_{n \in K \cap \mathbb{Z}^d} f(\psi_1(n)) \dots f(\psi_t(n)), \tag{1.1}$$

where K is a convex body. Questions of this second type include the generalised Hardy-Littlewood conjecture, which predicts, based on a probabilistic model for the prime numbers, an asymptotic for (1.1) when $f = \Lambda$ is given by the von Mangoldt function. Note that the frequency of arithmetic progressions of a fixed length t in the set of primes can be expressed as a special case

$$\mathbb{E}\big(\Lambda(n_1)\Lambda(n_1+n_2)\dots\Lambda(n_1+(t-1)n_2) \mid n_1+(t-1)n_2 \leqslant N\big)$$

of the $f = \Lambda$ instance of (1.1). The generalised Hardy-Littlewood conjecture has been resolved in the series of papers [6, 7, 8, 9] for those cases where no two forms ψ_i and ψ_j are affinely related. (Thus the prime k-tuples conjecture, which concerns the asymptotic behaviour of

$$\mathbb{E}_{n \leqslant N}\Lambda(n+h_1)\Lambda(n+h_2)\dots\Lambda(n+h_k)$$

for any k-tuple of integers $h_1, \dots, h_k$, remains unsettled.)

The general approach that was used in the aforementioned partial resolution of the Hardy-Littlewood conjecture is described as the 'nilpotent Hardy-Littlewood method' in [7]. It extends the classical Hardy-Littlewood method in the sense that it provides a tool to address a similar type of question as the classical method, but covers a wider range: it can be employed to resolve questions of the above second kind, provided the function f involved shows a certain amount of random-like behaviour. Furthermore, it resembles the classical method in that this approach too requires a (suitably adapted) major and a minor arc analysis (Section 9), cf. [7, §4] for a discussion of this analogy. A very central role in this method is assigned to *pseudorandom majorant* functions. We shall explain the reason for this and its role at the start of Section 4. In the case of the divisor function, the construction of the majorant constitutes the principal task that needs to be accomplished in order to apply the method and thus in order to obtain an asymptotic for (1.1) with $f = \tau$.

For an application of the nilpotent Hardy-Littlewood method the function f is required to have asymptotic density, that is, to satisfy

$$\mathbb{E}_{n \leqslant N} f(n) = \delta + o(1)$$

for some absolute constant $\delta \geqslant 0$. For this reason, we shall work not with the divisor function itself, but with the normalised divisor function $\tilde{\tau} : [N] \to \mathbb{R}_{\geqslant 0}$ which is defined by

$$\tilde{\tau}(n) := \frac{1}{\log N} \sum_{d \mid n} 1$$

and has asymptotic density $\delta = 1$.

A *pseudorandom majorant* for f is a function $\nu : [N] \to \mathbb{R}_{\geqslant 0}$ such that $|f(n)| \leqslant C\nu(n)$ pointwise (for some absolute constant C), and which resembles a random measure in the following sense. The total mass of ν is approximately 1, that is $\mathbb{E}_{n \leqslant N}\nu(n) = 1+o(1)$, and two further conditions modelling independence are satisfied. These are the linear forms and correlation conditions from [6]. The linear forms condition requires asymptotics of the form

$$\sum_{n\in K\cap\mathbb{Z}^d} \nu(\psi_1(n))\dots\nu(\psi_t(n)) = 1 + o(1).$$

Note that this is (1.1) for the majorant ν instead of f. Thus, to enable us to check this condition, the pseudorandom majorant ν has to be of a form that allows a good understanding of its value distribution. In particular, assuming that one failed to establish (1.1) for f directly and hence resorted to other methods of approach, the majorant has to be sufficiently easier to understand than the function f.

In the course of the above cited work on (1.1) for the von Mangoldt function, the problem of finding an asymptotic for (1.1) was also addressed for $f = \mu$, the Möbius function. A key feature of both functions μ and Λ is that they show some regularity in their growth. μ is bounded by 1 pointwise, whereas Λ grows not faster than log. This regularity simplifies the task of constructing a function that is simple enough that one can check the linear forms condition, and which simultaneously satisfies the majorant and the density condition.

The divisor function $\tau(n) = \sum_{d|n} 1$, on the other hand, is known for its irregularities in distribution. The moments $\mathbb{E}_{n\leqslant N}\tau(n)^p \sim (\log N)^{2^p-1}$ grow rapidly in p. While τ has an 'approximate' normal order, that is for every $\varepsilon > 0$ all but $o(N)$ positive integers $n \leqslant N$ satisfy $(\log N)^{(1-\varepsilon)\log 2} < \tau(n) < (\log N)^{(1+\varepsilon)\log 2}$, a theorem of Birch [1] implies that it does not have a normal order in the sense of Hardy and Ramanujan. Instead there is a gap between the 'approximate' normal order $(\log N)^{\log 2}$ and the average order $\log N$, which results from few exceptionally large values of τ. In particular, $\tau(n)$, and similarly $\tilde{\tau}(n)$, can be as large as $\exp(c\log n/\log\log n)$, see [10, §18.1 and §22.13].

We shall show that, nonetheless, there is a pseudorandom majorant $\nu : [N] \to \mathbb{R}_{\geqslant 0}$ for (a W-tricked version of) $\tilde{\tau}$, and that the same basic method that was employed to deal with $f = \mu$ and $f = \Lambda$ can also be employed in this case: The existence of this majorant in combination with the recent complete resolution of the Inverse Conjectures for the Gowers norms [9] allows us to deduce an asymptotic for $\sum_{n\in K\cap\mathbb{Z}^d} \tilde{\tau}(\psi_1(n))\dots\tilde{\tau}(\psi_t(n))$ under the already mentioned condition that no two forms ψ_i and ψ_j are affinely related.

Notation and statement of the main result. We recall some notation from [6] in order to state the result precisely.

Definition 1.1 (Affine-linear forms). *Let $d, t \geqslant 1$ be integers. An* affine-linear form *on $\mathbb{Z}^d$ is a function $\psi : \mathbb{Z}^d \to \mathbb{Z}$ which is the sum of $\psi = \dot{\psi} + \psi(0)$ of a linear form $\dot{\psi} : \mathbb{Z}^d \to \mathbb{Z}$ and a constant $\psi(0) \in \mathbb{Z}$. A* system of affine-linear forms *on $\mathbb{Z}^d$ is a collection $\Psi = (\psi_1, \dots, \psi_t)$ of affine-linear forms on $\mathbb{Z}^d$ that is required to satisfy the following non-degeneracy condition: no affine-linear form is constant and no two forms are rational multiples of each other.*

To formulate quantitative results, the following notion was introduced in [6] to classify the necessary bounds on the coefficients of the forms.

Definition 1.2 (Size at scale N, [6]). *If $N > 0$, define the* size $\|\Psi\|_N$ *of Ψ relative to the scale N by*

$$\|\Psi\|_N = \sum_{i=1}^{t}\sum_{j=1}^{d} |\dot{\psi}_i(e_j)| + \sum_{i=1}^{t} \frac{|\psi_i(0)|}{N} ,$$

where $e_1, \dots, e_d$ is the standard basis of $\mathbb{Z}^d$.

As we will show, the asymptotic behaviour of

$$\sum_{n\in K\cap\mathbb{Z}^d} \tilde{\tau}(\psi_1(n))\ldots\tilde{\tau}(\psi_t(n)) = \frac{1}{(\log N)^t}\sum_{d_1,\ldots d_t}\sum_{n\in K\cap\mathbb{Z}^d} 1_{d_1|\psi_1(n)}\ldots 1_{d_t|\psi_t(n)} \tag{1.2}$$

is determined by the local behaviour of the affine-linear system modulo small primes. To make this precise, we proceed to define local factors at primes.

For a given system $(\psi_1,\ldots,\psi_t)$ of affine-linear forms, positive integers $d_1,\ldots d_t$ and their least common multiple $m := \mathrm{lcm}(d_1,\ldots d_t)$ define *local divisor densities* by

$$\alpha(d_1,\ldots,d_t) := \mathbb{E}_{n\in(\mathbb{Z}/m\mathbb{Z})^d}\prod_{i\in[t]} 1_{d_i|\psi_i(n)} \ .$$

The Chinese remainder theorem implies that α is multiplicative. Thus, we restrict attention to what happens at prime powers $d_i = p^{a_i}$ for a fixed prime p. If the forms ψ_i were independent, one would expect $\alpha(p^{a_1},\ldots,p^{a_t}) = p^{-a_1}\ldots p^{-a_t}$. The prime powers of p would then contribute to (1.2) a factor of

$$\sum_{a_1,\ldots,a_t} p^{-a_1}\ldots p^{-a_t} = (1-p^{-1})^{-t} \ .$$

We therefore introduce for each prime p a *local factor*

$$\beta_p := (1-p^{-1})^t\sum_{a_1,\ldots,a_t\in\mathbb{N}}\alpha(p^{a_1},\ldots,p^{a_t})$$

which measures the irregularities of the divisor densities of the given system Ψ of affine-linear forms. As will be checked in the next section, the local factors satisfy the estimate $\beta_p = 1 + O_{t,d,L}(p^{-2})$. Thus, in particular, their product $\prod_p \beta_p$ converges.

Our main result is the following local-global principle.

Main Theorem. *Let N, d, t, L be positive integers and let $\Psi = (\psi_1,\ldots,\psi_t) : \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine-linear forms with coefficients bounded by $\|\Psi\|_N \leqslant L$ and for which any $\dot{\psi}_i$, $\dot{\psi}_j$, $i \neq j$, are linearly independent. Then*

$$\sum_{n\in[-N,N]^d} 1_K(n)\prod_{i=1}^t\tilde{\tau}(\psi_i(n)) = \mathrm{vol}(K)\prod_p\beta_p + o_{t,d,L}(N^d)$$

for any convex body $K \subseteq [-N,N]^d$ such that $\Psi(K) \subset [-N,N]^t$.

The corresponding asymptotic for the divisor function is an immediate consequence:

Corollary 1.3 (Correlations of the divisor function)**.** *With the assumptions of the Main Theorem, the divisor function τ satisfies*

$$\sum_{n\in[-N,N]^d} 1_K(n)\prod_{i=1}^t\tau(\psi_i(n)) = (\log N)^t\,\mathrm{vol}(K)\prod_p\beta_p + o_{t,d,L}(N^d\log^t N) \ .$$

The condition that no two forms $\dot{\psi}_i$ and $\dot{\psi}_j$ are linearly dependent, which the main theorem places upon the affine-linear system Ψ, is equivalent to saying that the affine-linear system Ψ has *finite complexity*, a notion introduced in [6]. The *infinite complexity* case includes problems of just one free parameter, like the one of estimating

$$\mathbb{E}_{n\leqslant N}\tau(n+a_1)\ldots\tau(n+a_k) \ . \tag{1.3}$$

These remain untouched, as they cannot be addressed by the nilpotent Hardy-Littlewood method. To place the task of estimating (1.3) into context, we mention that Ingham [11] proves the asymptotic

$$\sum_{n=1}^{N} \tau(n)\tau(n+a) = \frac{6}{\pi^2}\sigma_{-1}(a)\; N\log^2 N + O(N\log N)\ ,$$

where $\sigma_{-1}(a) = \sum_{d|a} d^{-1}$. No asymptotics are known when $k \geqslant 3$; c.f. [2, Thm 2] for a recent result into the direction of gaining asymptotics in the $k = 3$ case.

A subsequent paper [12] considers the problem of the type (1.1) for other arithmetic functions such as $r(n)$, the number of representations of n as a sum of two squares. This has some natural arithmetic consequences concerning the number of simultaneous integer zeros of pairs of certain diagonal quadratic forms, which are, in the 8-variables case, out of reach of the classical Hardy-Littlewood method as it currently stands.

2. Local divisor densities

This section contains some lemmas involving local divisor densities that are repeatedly used in analysing singular products. We also provide an estimate for β_p.

Let $\Psi = (\psi_1, \ldots, \psi_t) : \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine-linear forms which satisfies $\|\Psi\|_N \leqslant L$, let $K \subset \mathbb{R}^d$ be a convex body, and let $d_1, \ldots, d_t$ be integers. Divisibility events of the form

$$\sum_{n\in\mathbb{Z}^d\cap K}\prod_{i\in[t]} 1_{d_i|\psi_i(n)}$$

will naturally occur quite frequently in this paper. As in [6], the main tool to deal with these divisibility events is a simple volume packing lemma.

Lemma 2.1 (Volume packing argument)**.** *Let $K \subseteq [-L, L]^d$ be a convex body and Ψ a system of affine-linear forms. Then*

$$\sum_{n\in\mathbb{Z}^d\cap K}\prod_{i\in[t]} 1_{d_i|\psi_i(n)} = \mathrm{vol}(K)\alpha(d_1, \ldots, d_t) + O(L^{d-1}\,\mathrm{lcm}(d_1, \ldots, d_t))\ .$$

Proof. Let $\delta = \mathrm{lcm}(d_1, \ldots, d_t)$ and cover K by translates $\delta\mathbb{Z}^d + [0, \delta)^d$ of the box $[0, \delta)^d$. Each box contains $\delta^d\alpha(d_1, \ldots, d_t)$ points n such that $\prod_{i\in[t]} 1_{d_i|\psi_i(n)} = 1$. Any box that does not lie completely inside K is contained in the $2\bar{d}$-neighbourhood of the boundary of K, which has according to the Steiner theory of parallel surfaces a volume of order $O_d(\delta L^{d-1})$. Putting things together yields the result. □

We proceed to analyse the multiplicative function $\alpha = \alpha_\Psi$ more closely. If p is large compared to t, d, L, then

$$\alpha(p^{a_1}, \ldots, p^{a_t}) = p^{-a_j}\ , \tag{2.1}$$

when a_j is the only non-zero exponent. A prime p is called *exceptional* (with respect to Ψ) when there are forms ψ_i, ψ_j in the system that are affinely related modulo p. If $a_i, a_j > 0$, then considering the number of solutions $n \in (\mathbb{Z}/p^{\max(a_i,a_j)}\mathbb{Z})^d$ to $\psi_i(n) \equiv 0 \pmod{p^{a_i}}$, $\psi_j(n) \equiv 0 \pmod{p^{a_j}}$ yields $\alpha(p^{a_1}, \ldots, p^{a_t}) \leqslant p^{-a_i-a_j}$ if ψ_i and ψ_j are not affinely related. Thus, if p is not an exceptional prime, one has, with $a_{\max} := \max_i a_i$,

$$\alpha(p^{a_1}, \ldots, p^{a_t}) \leqslant p^{-a_{\max}-1}\ , \tag{2.2}$$

if there are at least two non-zero exponents.

Lemma 2.2 (Contribution from dependent divisibility events)**.** *Let Ψ be as above and let p be an unexceptional prime. Then*

$$\sum_{\substack{a_1,\dots,a_t\geqslant 0\\ \text{at least two } a_i\neq 0}} \alpha(p^{a_1},\dots,p^{a_t}) \ll_{t,d,L} \frac{1}{p^2} \ .$$

Proof. The number of t-tuples $(a_1,\dots,a_t)$ of non-negative integers with $\max_i a_i = j$ is at most tj^{t-1}. This together with the bound (2.2) yields

$$\sum_{\substack{a_1,\dots,a_t\geqslant 0\\ \text{at least two } a_i\neq 0}} \alpha(p^{a_1},\dots,p^{a_t}) \ll_{L,t,d} \sum_{j\geqslant 1} \frac{j^t}{p^{j+1}} = \sum_{k\geqslant 2} \frac{1}{p^k}\left(\frac{(2k-1)^t}{p^{k-2}} + \frac{(2k)^t}{p^{k-1}}\right) \ .$$

There is p_0 such that whenever $p > p_0$ then all the brackets in the last sum are less than 1, except the bracket for $k=2$. Thus, for $p > p_0$

$$\sum_{\substack{a_1,\dots,a_t\geqslant 0\\ \text{at least two } a_i\neq 0}} \alpha(p^{a_1},\dots,p^{a_t}) \ll_{L,t,d} \frac{1}{p^2} + \sum_{k\geqslant 3}\frac{1}{p^3} \ll_{L,t,d} \frac{1}{p^2} \ .$$

□

The following lemma immediately implies the convergence of $\prod_p \beta_p$ whenever Ψ contains no two forms ψ_i and ψ_j that are affinely dependent, and thus every exceptional prime is bounded by $O_{t,d,L}(1)$.

Lemma 2.3. *Let Ψ be as above and let p be an unexceptional prime, then*

$$\beta_p = 1 + O_{t,d,L}(p^{-2}) \ . \tag{2.3}$$

Proof. By Lemma 2.2 and the bound (2.1)

$$\begin{aligned}\beta_p &= (1-p^{-1})^t \sum_{a_1,\dots,a_t\in\mathbb{N}} \alpha(p^{a_1},\dots,p^{a_t})\\ &= \left(1-\frac{t}{p}+O_{t,d,L}(p^{-2})\right)\left(1+\frac{t}{p}+O_{t,d,L}(p^{-2})\right)\\ &= 1+O_{t,d,L}(p^{-2})\end{aligned}$$

which proves the result. □

3. Some arithmetical lemmas and a reduction

In this section we record for later reference some early lemmas from [3], adapted to our purposes, and deduce a reduction of the Main Theorem.

Lemma 3.1 (Second moment bound for the divisor function τ)**.** *Let $\Psi = (\psi_1,\dots,\psi_t) : \mathbb{Z}^m \to \mathbb{Z}^t$ be a system of affine-linear form with coefficients bounded by $\|\Psi\|_{[N]} \leqslant L$. Then*

$$\mathbb{E}_{n\in\mathbb{Z}^m\cap K} \prod_{i\in[t]} \tau^2(\psi_i(n)) \ll_{t,m,L} (\log N)^3 \ .$$

Proof. The standard approach to obtain a second moment estimate for the divisor function carries over: Note that $\mathbb{E}_{n\in\mathbb{Z}^m\cap K}1_{d|\psi_i(n)} \ll_{m,L} \frac{1}{d}$, and thus

$$\begin{aligned}
\mathbb{E}_{n\in\mathbb{Z}^m\cap K}\prod_{i\in[t]}\tau^2(\psi_i(n)) &= \sum_{\substack{d_1,\dots,d_t\in\mathbb{N}\\ d_1',\dots d_t'\in\mathbb{N}}} \mathbb{E}_{n\in\mathbb{Z}^m\cap K}\prod_{i\in[t]} 1_{[d_i,d_i']|\psi_i(n)} \\
&= \sum_{\substack{(d_i),(d_i'),(d_i'')\in\mathbb{N}^t\\ \gcd(d_i,d_i',d_i'')=1, i\in[t]}} \mathbb{E}_{n\in\mathbb{Z}^m\cap K}\prod_{i\in[t]} 1_{d_id_i'd_i''|\psi_i(n)} \\
&\leqslant \sum_{\substack{(d_i),(d_i'),(d_i'')\in\mathbb{N}^t\\ \gcd(d_i,d_i',d_i'')=1, i\in[t]}} \mathbb{E}_{n\in\mathbb{Z}^m\cap K}\sum_{i\in[t]} 1_{d_id_i'd_i''|\psi_i(n)} \\
&\ll_{m,L} \sum_{d,d',d''\leqslant N}\frac{t}{dd'd''} \ll_{t,m,L} (\log N)^3 \ .
\end{aligned}$$

□

Lemma 3.2 ("rough" numbers are rare, [3])**.** *Suppose $\Psi = (\psi_1,\dots,\psi_t) : \mathbb{Z}^d \to \mathbb{Z}^t$ is affine-linear and $\|\Psi\|_N \leqslant L$. Let $C_1 > 1$ be a parameter and let S_1 be the set of $m\in\mathbb{Z}$ which are divisible by a large proper prime power $p^a > \log^{C_1} N$, $a \geqslant 2$. Then the density of $n \in \mathbb{Z}^d\cap K$ such that $\psi_i(n)\in S_1$ for at least one $i\in[t]$ is bounded by*

$$\sum_{i\in[t]}\mathbb{E}_{n\in\mathbb{Z}^d\cap K}1_{\psi_i(n)\in S_1} \ll_{L,d,t} \log^{-C_1/2} N \ .$$

Proof. Similarly as above, this is a straightforward adaption of the one-dimensional estimate. Note that $\mathbb{E}_{n\in\mathbb{Z}^d\cap K}1_{p^a|\psi_i(n)} \ll_{L,d} p^{-a}$ for all primes p. Let $a(p)$ be the smallest exponent $a\geqslant 2$ for which $p^a > \log^{C_1} N$. We then have

$$\begin{aligned}
\sum_{i\in[t]}\mathbb{E}_{n\in\mathbb{Z}^d\cap K}1_{\psi_i(n)\in S_1} &\leqslant \sum_p \mathbb{E}_{n\in\mathbb{Z}^d\cap K}\sum_{i\in[t]}1_{p^{a(p)}|\psi_i(n)} \\
&\ll_{L,d} \sum_{p\leqslant \log^{C_1/2} N} t\log^{-C_1} N + \sum_{p>\log^{C_1/2} N} tp^{-2} \\
&\ll_{L,d,t} \log^{-C_1/2} N \ .
\end{aligned}$$

□

Lemma 3.3 ("smooth" numbers are rare, [3])**.** *Let Ψ be as in the previous lemma, let $\gamma < 1$ be a parameter and let S_2 be the set of smooth $m\in\mathbb{Z}$, that is, m for which*

$$\prod_{\substack{p^a\|m\\ p\leqslant N^{1/(\log\log N)^3}}} p^a \geqslant N^{\gamma/\log\log N} \ . \tag{3.1}$$

Then the density of $n\in\mathbb{Z}^d\cap K$ for which $\psi_i(n)\in S_2$ for at least one $i\in[t]$ is bounded by

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K}1_{n\in S_2} \ll_{L,d,t,\gamma,C_1} \log^{-C_1/2} N \ .$$

Proof. Suppose that $\psi_i(n)\in S_2$ but does not belong to the set S_1 from the previous lemma at the same time. Then each prime power in the product (3.1) for $m = \psi_i(n)$ is in particular $\ll_{C_1} N^{1/(\log\log N)^3}$. Since $\psi_i(n) > N^{\gamma/\log\log n}$, we then have

$$\tau(\psi_i(n)) \geqslant 2^{\omega(\psi_i(n))} \gg_{C_1} 2^{(\log\log N)^3\gamma/\log\log N} \gg_{\gamma,C_2} (\log N)^{C_2}$$

for any positive constant C_2. By Lemma 3.1 and Cauchy-Schwarz this can for each value of i only happen on a set of density $< (\log N)^{3-2C_2}$. The result follows with $C_2 \geqslant 3/2 + C_1/4$. □

The next lemma shows that S_1 and S_2 are exceptional sets for the divisor function.

Lemma 3.4 (Contribution from the exceptional sets S_1 and S_2)**.** *Let Ψ be as before and let $C_3 \geqslant 1$ be a parameter. For sufficiently large C_1, we have*

$$\sum_{i\in[t]} \mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{i\in[t]} \tilde{\tau}(\psi_i(n)) 1_{\psi_i(n)\in S_1\cup S_2} \ll_{t,d,L} (\log N)^{-C_3} \ .$$

Proof. This follows by the Cauchy-Schwarz inequality from lemmata 3.1, 3.2 and 3.3 provided C_1 is chosen large enough. □

The previous lemma reduces the task of proving the Main Theorem as follows.

Proposition 3.5. *Let $\bar{\tau} : \mathbb{Z} \to \mathbb{R}$ be any function that agrees with $\tilde{\tau}$ on the complement of $S_1 \cup S_2$ and satisfies $0 \leqslant \bar{\tau}(n) \leqslant \tilde{\tau}(n)$ for $n \in S_1 \cup S_2$. Then the Main Theorem, that is, $\sum_{n\in\mathbb{Z}^d\cap K} \prod_{i\in[t]} \tilde{\tau}(\psi_i(n)) = \mathrm{vol}(K)\prod_p \beta_p + o_{L,t,d}(N^d)$, holds if and only if under the same conditions*

$$\sum_{n\in\mathbb{Z}^d\cap K} \prod_{i\in[t]} \bar{\tau}(\psi_i(n)) = \mathrm{vol}(K)\prod_p \beta_p + o_{L,t,d}(N^d) \ .$$

4. A majorant for the normalised divisor function

Suppose that $A \subseteq [N]$ has cardinality $|A| = \delta N$. Loosely speaking, if $0 < \delta < 1$ is fixed, we refer to such sets A, for N arbitrarily large, as *dense*. In this case, a sufficient condition for A to contain approximately the expected number of finite complexity structures is that A is sufficiently Gowers-uniform. This is to say, the uniformity norm

$$\|1_A - \delta\|_{U^s[N]} := \left(\mathbb{E}_{x\in[N]}\mathbb{E}_{h\in[N]^s} \prod_{\omega\in\{0,1\}^s} (1_A - \delta)(x + \omega\cdot h)\right)^{1/2^s}$$

is small for some s that is determined by the structure one is counting. For instance, the number of 4-term arithmetic progressions in a set A of size $|A| = \delta N$ satisfies

$$\mathbb{E}_{n+3d\leqslant N} 1_A(n)1_A(n+d)1_A(n+2d)1_A(n+3d) \sim \delta^4 \ ,$$

if $\|1_A - \delta\|_{U^3}$ is small. These results remain to be true when one replaces 1_A by a function $f : \mathbb{N} \to \mathbb{C}$ that is bounded independent of N and that has asymptotic density $\mathbb{E}_{n\leqslant N} f(n) = \delta + o(1)$.

If f fails to satisfy these properties, that is, if it is either sparse or unbounded, then a *transference principle* is required. Such a principle was established by Green and Tao in [5, 6] and is based on the observation that a sparse set that is relatively dense in a random-like set behaves in the same way as a dense set.

The first step is to replace the function f by a model $\tilde{f}$ that has asymptotic density. Examples are the replacement of the characteristic function of primes by the von Mangoldt function or the replacement of τ by $\tilde{\tau}$ in our case.

An application of the transference principle requires a *majorant* function $\nu : [N] \to \mathbb{C}$ with $|\tilde{f}(n)| \leqslant C\nu(n)$ for all n which satisfies the linear forms and correlation conditions (c.f. Section 6) of [6, §6], two conditions which are designed to model a random measure. This majorant replaces the "random-like set" from the observation. The relative density condition from the observation is also present in the generalised case. Indeed, part of the definition of ν is that $\mathbb{E}_{n \leqslant N}\nu(n) = 1 + o(1)$, and we further replaced the original function f by a dense model $\tilde{f}$. Thus we have

$$\delta\ \mathbb{E}_{n \leqslant N}\nu(n) = \delta(1 + o(1)) \leqslant (1 + o(1))\mathbb{E}_{n \leqslant N}\tilde{f}(n) \leqslant C(1 + o(1))\mathbb{E}_{n \leqslant N}\nu(n)$$

and hence $\tilde{f}$ can be regarded as being 'dense' in ν.

The Koopman–von Neumann theorem [6, Prop.10.3], or [5, Prop.8.1], then provides a result corresponding to the above observation: Any function f with asymptotic density $\mathbb{E}f$ that is dominated by a pseudorandom measure $|f(n)| \leqslant \nu(n)$ may be decomposed as a sum $f = f_1 + f_2$ where f_1 is bounded and $f_2 - \mathbb{E}f_2$ has small uniformity norms. Thus, $f - \mathbb{E}f$ has small uniformity norms if and only if the bounded function $f_1 - \mathbb{E}f_1$ has, and one can apply the results from the dense setting to f_1. That is, we have 'transferred' the problem to the dense setting, provided there is a way to deal with the error $f_2 - \mathbb{E}f_2$. Such a way is provided by [6, Cor. 11.6].

In the case $f = \Lambda$, Green and Tao [5] construct, building upon work of Goldston and Yildirim, the required pseudorandom majorant by modifying the majorant the proof of Selberg's sieve is based on. The key property of such a majorant resulting from a Selberg sieve is that it has the form of a *truncated* divisor sum

$$\nu(n) := \sum_{d|n, d \leqslant N^\gamma} a_d$$

for certain coefficients a_d and where $\gamma > 0$ is a fixed constant that may be chosen as small as necessary. Its importance lies in the fact that summing only over small divisors ensures that the divisibility events that occur when checking the linear forms condition are almost independent, and thus allows us to deduce *asymptotics* as required for the linear forms condition.

Our aim in this section is to show that a majorant of similar structure can be constructed in the case of the divisor function $\tilde{\tau}$. A first attempt, given the above discussion, might be to take

$$\nu(n) = \tilde{\tau}_\gamma(n) := \frac{1}{\gamma \log N} \sum_{d|n : d \leqslant N^\gamma} 1\ .$$

Unfortunately, however, a result of Tenenbaum [15, Cor.3] asserts that if $\gamma < 1/2$ then for every λ the majorant condition $\tilde{\tau}(n) \leqslant \lambda\tilde{\tau}_\gamma(n)$ fails to hold on a positive proportion of $n \in [N]$. A modification of this idea is therefore required. It turns out that the proportion of such 'bad' n can be bounded by $\lambda^{-c \log\log\lambda}$ for some $c = c(\gamma) > 0$. Denoting by $X(\lambda)$ the set of bad n for λ, then the bound on $|X(\lambda)|$ allows us to sum $\sum_{i \geqslant 1} \lambda_i 1_{X(\lambda_i)}(n)\tilde{\tau}_\gamma(n)$ for suitable sequences (λ_i).

The idea behind this is due to Erdős [3]: Let $N^\gamma < n \leqslant N$. Considering the distribution of prime factors of such a number, one expects that $\tilde{\tau}(n)$ is essentially controlled by the number of small divisors $\tilde{\tau}_\gamma(n)$. But when is this actually the case? A sufficient condition may be obtained as follows. Write $n = p_1^{a_1} \dots p_t^{a_t}$, where the primes are ordered by increasing size, and let $p_1^{a_1} \dots p_{j+1}^{a_{j+1}}$ be the first initial partial product that exceeds N^γ. Then we are guaranteed control of $\tilde{\tau}(n)$ by $\tilde{\tau}_\gamma(n)$ provided p_{j+1} is large, since n has at most $\frac{\log N}{\log p_{j+1}}$ prime factors $> p_{j+1}$. The quality of control depends on the size of p_{j+1}. Suppose n is a 'bad' integer for which the control is of $\tilde{\tau}(n)$ by $\tilde{\tau}_\gamma(n)$ is not good enough, thus, p_{j+1} is quite small. The smaller p_{j+1} is, the worse is the control, but, also, the denser gets the distribution of prime factors of the large initial product of n. Excluding the sparse set of numbers that have a large proper prime power divisor, one expects to find some structure in the 'dense' set of prime factors $< p_{j+1}$. A pigeonhole argument shows that there is some short interval that contains quite a large number of those prime factors $< p_{j+1}$; a very sparse event.

The prime divisor structure of 'bad' integers n that this proof strategy provides will be important later on, because it allows us to *explicitly* describe the exceptional set for the inequality $\tilde{\tau}(n) \leqslant \lambda\tilde{\tau}_\gamma(n)$ at level λ.

The following lemma is a reformulation of Erdős's observations from [3].

Lemma 4.1 (Erdős)**.** *Let $n \leqslant N$ and suppose that $\tilde{\tau}(n) \geqslant 2^s\tilde{\tau}_\gamma(n)$ for some $s > 2/\gamma$. Then one of the following three alternatives holds:*

(i) *n is excessively "rough" in the sense that it is divisible by some prime power p^a, $a \geqslant 2$, with $p^a > \log^{C_1} N$;*

(ii) *n is excessively "smooth" in the sense that if $n = \prod_p p^a$ then*

$$\prod_{p \leqslant N^{1/(\log\log N)^3}} p^a \geqslant N^\gamma;$$

(iii) *n has a "cluster" of prime factors in the sense that there is an i, $\log_2 s - 2 \leqslant i \ll \log\log\log N$ such that n has at least $\gamma s(i + 3 - \log_2 s)/100$ prime factors in the superdyadic range $I_i := [N^{1/2^{i+1}}, N^{1/2^i}]$ and is not divisible by two primes in this range.*

Proof. The alternatives (i) and (ii) correspond to the sets S_1 and S_2 from Section 3 and thus can be regarded as *exceptional*. Suppose that n is unexceptional, that is (i) and (ii) are not satisfied, and that the prime factorisation of n is given by

$$n = p_1^{a_1} \dots p_k^{a_k},$$

where $p_1 < \cdots < p_k$. Let j be the index for which

$$p_1^{a_1} \dots p_j^{a_j} \leqslant N^\gamma < p_1^{a_1} \dots p_{j+1}^{a_{j+1}}, \tag{4.1}$$

and write

$$n' := p_1^{a_1} \dots p_j^{a_j} \ .$$

We claim that $n' \geqslant N^{\gamma/2}$. Indeed, if this is not the case, then $p_{j+1}^{a_{j+1}} \geqslant N^{\gamma/2}$. Since (i) does not hold we have $a_{j+1} = 1$. Thus, since $p_{j+1} \dots p_k | n$, we have $k - j \leqslant 2/\gamma$. Furthermore, using the fact that (i) does not hold once more, we have $a_{j+1} = \cdots = a_k = 1$ and so in this case

$$\tilde{\tau}(n) = 2^{k-j}\tilde{\tau}(n') \leqslant 2^{2/\gamma}\tilde{\tau}(n') \leqslant 2^{2/\gamma}\tilde{\tau}_\gamma(n) < 2^s\tilde{\tau}_\gamma(n) \ ,$$

contrary to assumption. Let $r \geqslant 1$ be the unique integer such that

$$N^{\gamma/(r+1)} < p_j \leqslant N^{\gamma/r} \ .$$

Then

$$a_{j+1} + \cdots + a_k \leqslant \frac{\log N}{\log p_j} \leqslant \frac{r+1}{\gamma} \ ,$$

which means that

$$\tilde\tau(n) = (a_{j+1}+1)\dots(a_k+1)\tilde\tau(n') \leqslant 2^{a_{j+1}+\dots+a_k}\tilde\tau(n') \leqslant 2^{(r+1)/\gamma}\tilde\tau_\gamma(n) \leqslant 2^{2r/\gamma}\tilde\tau_\gamma(n)$$

and thus, recalling the assumption $2^s\tilde\tau_\gamma(n) \leqslant \tilde\tau(n)$, we have

$$r \geqslant s\gamma/2.$$

All prime factors of n' are therefore bounded by $N^{2/s}$.

Since we are not in the exceptional case (ii), the small prime factors have a negligible contribution

$$\prod_{p \leqslant N^{1/(\log\log N)^3}} p^a \leqslant N^{\gamma/\log\log N} \ . \tag{4.2}$$

Consider the smallest collection of superdyadic intervals $I_i = [N^{1/2^{i+1}}, N^{1/2^i}]$ which cover $(N^{1/(\log\log N)^3}, N^{2/s}]$; hence, these i satisfy $\log_2 s - 2 \leqslant i < 6\log\log\log N$. In view of (4.2), the bound $p_j \leqslant N^{2/s}$ and the fact that $n' \geqslant N^{\gamma/2}$, we obtain

$$\prod_i \prod_{\substack{p\in I_i \\ p^a \| n}} p^a \geqslant N^{\gamma/2 - \gamma/\log\log N} > N^{\gamma/4} \ .$$

Since n is unexceptional (and, specifically, (i) does not hold), all of the a's appearing here are equal to one. Thus if the lemma were false, we would have

$$N^{\gamma/4} \leqslant \prod_{i \geqslant \log_2 s - 2} N^{\gamma s(i+3-\log_2 s)/(100\cdot 2^i)} = \exp\left(\log N \sum_{j\geqslant 1} \gamma \frac{j}{2^j}\frac{s}{2^{\log_2 s - 2}}\frac{1}{100}\right) < N^{\gamma/4},$$

a contradiction[1]. □

It is possible to bound the number of $n \leqslant N$ satisfying condition (iii) for some value of i just using their specific structure. Setting $m_0 := \lceil \gamma s(i+3-\log_2 s)/100 \rceil$, write $X(i,s)$ for the set of $n \leqslant N$ divisible by at least $m_0(i,s)$ primes in $[N^{1/2^{i+1}}, N^{1/2^i}]$. Thus

$$N^{-1}|X(i,s)| \leqslant \frac{1}{m_0!}\left(\sum_{p\in I_i}\frac{1}{p}\right)^{m_0} = \frac{1}{m_0!}(\log 2 + o(1))^{m_0} \ .$$

The crude bound $m! \geqslant \left(\frac{m}{e}\right)^m$ yields the estimate

$$N^{-1}|X(i,s)| \leqslant \begin{cases} (c/\gamma s)^{\gamma s} & \text{if } s/4 \leqslant 2^i \leqslant s^2 \\ (c/\gamma s)^{\gamma s i} & \text{if } 2^i > s^2 \ , \end{cases} \tag{4.3}$$

and hence

$$N^{-1}\sum_{i \geqslant \log_2 s - 2} |X(i,s)| \leqslant (c/\gamma s)^{\gamma s}\log_2 s \ . \tag{4.4}$$

In particular, given the paucity of integers n satisfying (i) and (ii) as guaranteed by Lemma 3.4, this together with Lemma 4.1 shows that the density of $n \leqslant N$ for which $\tilde\tau(n) > 2^s\tilde\tau_\gamma(n)$ is bounded by $2^{-c_\gamma s\log s}$. The fast decay of these densities makes the following definition reasonable.

---

[1]The somewhat arbitrary factor of 100 could have been replaced by any other positive number that was large enough to induce this contradiction.

Proposition 4.2 (Majorant for the divisor function)**.** *Fix $\gamma > 0$. Write $U(i,s)$ for the set of all $n \leqslant N$ divisible by exactly $m_0(i,s) := \lceil \gamma s(i+3-\log_2 s)/100 \rceil$ distinct primes in the interval $[N^{1/2^{i+1}}, N^{1/2^i}]$, and not by the square of any such prime. Define $\nu : [N] \to \mathbb{R}_+$ by*

$$C\nu(n) := 2^{2/\gamma}\tilde{\tau}_\gamma(n) + \sum_{s>2/\gamma}^{(\log\log N)^3} \sum_{i=\log_2 s-2}^{6\log\log\log N} \sum_{u\in U(i,s)} 2^s 1_{u|n}\tilde{\tau}_\gamma(n) + 1_{n\in S_1\cup S_2}\tilde{\tau}(n)\ ,$$

where $S_1 \cup S_2$ is the set of all $n \leqslant N$ satisfying either (i) or (ii) of Lemma 4.1, Then there is a value of C (depending on γ) such that $\mathbb{E}_{n\leqslant N}\nu(n) = 1 + o(1)$. For all $n \leqslant N$ we have $\tilde{\tau}(n) \leqslant C\nu(n)$.

Remarks. (1) Since γ will be as small as necessary in every later application, we may as well choose it to be the reciprocal of an integer. This has the advantage that, setting $U(i, 2/\gamma) := \{1\}$ for $i = \log_2 s - 2$ and $U(i, 2/\gamma) := \emptyset$ otherwise, we can write

$$C\nu(n) = \sum_{s=2/\gamma}^{(\log\log N)^3} \sum_{i=\log_2 s-2}^{6\log\log\log N} \sum_{u\in U(i,s)} 2^s 1_{u|n}\tilde{\tau}_\gamma(n) + 1_{n\in S_1\cup S_2}\tilde{\tau}(n)\ .$$

(2) While ν can be shown to be pseudorandom, a further reduction in the next section will allow us to save some work by dropping the exceptional term $1_{n\in S_1\cup S_2}\tilde{\tau}(n)$.
(3) Finally, note that the divisors $u \in U(i,s)$ are truncated divisors themselves, that is, they satisfy $u \leqslant N^\gamma$. Indeed, suppose $i + 3 - \log_2 s = j(> 1)$, and hence $s/2^i = 8/2^j$, then

$$u \leqslant N^{m_0(i,s)/2^i} \leqslant N^{2\gamma sj/(100\cdot 2^i)} \leqslant N^{2\gamma 8j/(100\cdot 2^j)} < N^\gamma\ .$$

Proof. The fact that $\tilde{\tau}(n) \leqslant C\nu(n)$ is an immediate consequence of Lemma 4.1. To show the existence of C, we have to check that the expectation of ν on the integers $\leqslant N$ is bounded independent of N. Note that

$$\mathbb{E}_{n\leqslant N} \sum_{u\in U(i,s)} 1_{u|n}\tilde{\tau}_\gamma(n) \leqslant \frac{1}{m_0!}\left(\sum_{p\in I_i}\frac{1}{p}\right)^{m_0} \frac{1}{\gamma\log N}\sum_{m\leqslant N^\gamma}\frac{1}{m} \leqslant \frac{1}{m_0!}(\log 2 + o(1))^{m_0}\ .$$

This allows us to make use of a bound of type (4.4). In detail,

$$\begin{aligned}
&\mathbb{E}_{n\leqslant N}\sum_{s\geqslant 2/\gamma}\sum_{i\geqslant\log_2 s-2}\sum_{u\in U(i,s)} 1_{u|n}\tilde{\tau}_\gamma(n)2^s \\
&\leqslant \sum_{s\geqslant 2/\gamma}\sum_{i\geqslant\log_2 s-2}\frac{1}{m_0!}(\log 2 + o(1))^{m_0}\ 2^s \\
&\leqslant \sum_{s\geqslant 2/\gamma}\sum_{j\geqslant 1}\left(\frac{100\cdot e\cdot(\log 2 + o(1))}{\gamma sj}\right)^{\gamma sj/100} 2^s \\
&\leqslant \sum_{s\geqslant 2/\gamma}\frac{2^s}{s^{s\gamma/100}}\left(\sum_{j\geqslant 1}\left(\frac{100\cdot e\cdot(\log 2 + o(1))}{\gamma j}\right)^{\gamma j/100}\right)^s
\end{aligned}$$

which converges. We note for later reference that the above expression still converges when the factor 2^s is replaced by a^s with any positive constant a. □

5. W-trick

The nilpotent Hardy-Littlewood method employs the uniformity of a function to deduce an asymptotic for finite complexity correlations. However, the divisor function $\tilde{\tau}$ is not equidistributed in residue classes to small moduli and thus in particular not Gowers-uniform. To remove this obstruction we shall use a so-called W-trick and decompose $\tilde{\tau}$ into a sum of functions which do not detect a difference between these residue classes. This decomposition of $\tilde{\tau}$ can be viewed as a factorisation as product of a uniform function and an almost periodic function.

It is natural to consider the restricted divisor function that does not count divisors with small prime factors at all:

Definition 5.1 (W-tricked divisor function)**.** *Set $w(N) := \frac{1}{2}\log\log N$ and $W := \prod_{p<w(N)} p$. We define W-tricked versions of $\tilde{\tau}$ and $\tilde{\tau}_\gamma$ by*

$$\tilde{\tau}'(n) := \frac{W}{\phi(W)}(\log N)^{-1} \sum_{(d,W)=1} 1_{d|n} \ ,$$

and

$$\tilde{\tau}_\gamma'(n) := \frac{W}{\phi(W)}(\gamma \log N)^{-1} \sum_{\substack{d\leqslant N^\gamma \\ (d,W)=1}} 1_{d|n} \ ,$$

where ϕ denotes Euler's totient function.

Thus $\tilde{\tau}$ decomposes as a product

$$\tilde{\tau}(n) = \tilde{\tau}'(n)\left(\frac{\phi(W)}{W} \sum_{\substack{w\in\mathbb{N}\ : \\ p|w \ \Rightarrow \ p<w(N)}} 1_{w|n}\right) ,$$

where the first factor is expected to be uniform and the second factor is almost periodic. We may, in fact, replace the second factor by a periodic function: Setting

$$\overline{W} := \prod_{p\leqslant w(N)} p^{\left\lfloor \frac{\gamma}{\log\log N} \log_p N \right\rfloor} \leqslant N^{\pi(w(N))\gamma/\log\log N} \leqslant N^{o(1)\gamma} \ ,$$

define the following explicit function $\bar{\tau} : \mathbb{Z} \to \mathbb{R}$ by

$$\bar{\tau}(n) := \tilde{\tau}'(n)\Big(\frac{\phi(W)}{W}\sum_{w|\overline{W}} 1_{w|n}\Big) \ .$$

Since any integer n that is divisible by some $w \nmid \overline{W}$, completely composed of primes $< w(N)$, belongs to the exceptional set S_2 (c.f. Lemma 3.3), $\bar{\tau}$ satisfies the conditions of Proposition 3.5. This will allow us to deduce the Main Theorem from the following Proposition, to be established in Section 8.

Proposition 5.2. *Let $M = N/\overline{W}$ and suppose $K' \subset [M]^d$ is a convex body. Then for any choice of $b_1, \dots, b_t \in [\overline{W}]$*

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K'} \prod_{i\in[t]} \tilde{\tau}'(\overline{W}\tilde{\psi}_i(n) + b_i) = 1 + o_{d,t,L}(M^d/\operatorname{vol}(K'))$$

holds for all finite complexity systems $\tilde{\Psi}$ of affine-linear forms satisfying $\|\Psi\|_{[M]} \leqslant L$.

Proof of the Main Theorem from Proposition 5.2. Assume K and Ψ satisfy the conditions of the Main Theorem. Fix some $a \in [\overline{W}]^d$ and let $\tilde{\psi}_{i,a} : \mathbb{Z}^d \to \mathbb{Z}$ be the linear function for which

$$\psi_i(\overline{W}n + a) = \overline{W}\tilde{\psi}_{i,a}(n) + b_i(a)$$

where $b_i(a) \in [\overline{W}]$. Note that ψ_i and $\tilde{\psi}_{i,a}$ only differ in the constant term. Since $\overline{W} < N^{o(1)\gamma}$, this implies that whenever $\|\Psi\|_{[N]} = O(1)$ then $\|\tilde{\Psi}\|_{[M]} = O(1)$ too.

Define $K'_a \subset [N/\overline{W}]^d$ to be the convex body $\{x \in [0, N/\overline{W}]^d : \overline{W}x + a \in K\}$.

By Proposition 5.2,

$$\begin{aligned}\sum_{n\in\mathbb{Z}^d\cap K}\prod_{i\in[t]}\bar{\tau}(\psi_i(n)) &= \sum_{a\in[\overline{W}]^d}\sum_{\substack{\overline{W}n+a\in\\ \mathbb{Z}^d\cap K}}\prod_{i\in[t]}\left(\frac{\phi(W)}{W}\sum_{w|\overline{W}}1_{w|\psi_i(a)}\right)\tilde{\tau}'(\overline{W}\tilde{\psi}_{i,a}(n)+b_i(a))\\ &= \sum_{a\in[\overline{W}]^d}\prod_{i\in[t]}\left(\frac{\phi(W)}{W}\sum_{w|\overline{W}}1_{w|\psi_i(a)}\right)\frac{\operatorname{vol}K}{\overline{W}^d}(1+o_{t,d,L}(1))\\ &= \frac{\phi^t(W)}{W^t}\prod_{p\leqslant w(N)}\sum_{\substack{e_1,\dots,e_t\in\mathbb{N}\\ p^{e_i}|\overline{W}, i\in[t]}}\alpha(p^{e_1},\dots,p^{e_t})\operatorname{vol}(K)(1+o_{t,d,L}(1))\ .\end{aligned}$$

Since smooth numbers are rare by Lemma 3.3, we have

$$\sum_{\substack{e_1,\dots,e_t\in\mathbb{N}\\ p^{e_i}|\overline{W}, i\in[t]}}\alpha(p^{e_1},\dots,p^{e_t}) = \sum_{e_1,\dots,e_t\in\mathbb{N}}\alpha(p^{e_1},\dots,p^{e_t}) + o(1) = \beta_p(1-p^{-1})^{-t} + o(1)\ ,$$

and since $\prod_{p\leqslant w(N)}(1-p^{-1})^{-1} = \frac{W}{\phi(W)}$, the above implies

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{i\in[t]}\bar{\tau}(\psi_i(n)) = (1+o(1))\prod_{p\leqslant w(N)}\beta_p\ .$$

The local factors' bound (2.3), that is $\beta_p = 1 + O_{t,d,L}(p^{-2})$, and Proposition 3.5 yield the Main Theorem:

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{i\in[t]}\tilde{\tau}(\psi_i(n)) = \prod_p \beta_p + o(1)\ .$$

□

W-TRICKED MAJORANT. In order to prove Proposition 5.2, we require for any given choice of $b = (b_1, \dots, b_t) \in [\overline{W}]^t$ a majorant that simultaneously majorises all of the functions $n \mapsto \tilde{\tau}'(\overline{W}n + b_i)$ for $i = 1, \dots, t$. Define

$$C'\nu'(n) := \sum_{s=2/\gamma}^{(\log\log N)^3}\ \sum_{i=\log_2 s-2}^{6\log\log\log N}\ \sum_{u\in U(i,s)} 2^s 1_{u|n}\tilde{\tau}'_\gamma(n)\ , \qquad (5.1)$$

where C' is such that $\mathbb{E}_{n\leqslant N}\nu'(n) = 1 + o(1)$. This and the definition of $\tilde{\tau}'$ imply $\mathbb{E}_{n\leqslant M}\nu'(\overline{W}n + a) = 1 + o(1)$ for all $a \in [\overline{W}]$.

Thus, a majorant of the required form is given by a constant multiple of

$$\nu'_{\overline{W},b} : \mathbb{Z}^t \to \mathbb{Z}, \qquad \nu'_{\overline{W},b} := \mathbb{E}_{i\in[t]}\nu'(\overline{W}n + b_i)\ .$$

Note furthermore that $\nu'_{\overline{W},b}$ still satisfies the condition $\mathbb{E}_{m\leqslant M}\nu'_{\overline{W},b}(m) = 1 + o(1)$.

6. The linear forms condition

The aim of the following two sections is to show that the following slight modification of the majorant $\nu'_{\overline{W},(b_i)} = \mathbb{E}_{i\in[t]}\nu'(\overline{W}n + b_i)$ is indeed pseudorandom. Let M' be a prime satisfying $M < M' \leqslant O_{t,d,L}(M)$ and define $\nu^*_{\overline{W},(b_i)} : [M'] \to \mathbb{R}^+$ by

$$\nu^*_{\overline{W},(b_i)}(n) = \begin{cases} \frac{1}{2}(1 + \nu'_{\overline{W},(b_i)}(n)) & \text{if } n \leqslant M \\ 1 & \text{if } M < n \leqslant M' \ . \end{cases}$$

As is seen in [6, App.D], $\nu^*_{w,(b_i)}$ is D-pseudorandom if it satisfies the following two propositions, which are technical reductions of the linear forms and correlation conditions from [6].

Proposition 6.1 (D-Linear forms estimate)**.** *Let $1 \leqslant d, t \leqslant D$ and let $(i_1, \dots, i_t) \in [t]^t$ be an arbitrary collection of indices. For any finite complexity system $\Psi : \mathbb{Z}^d \to \mathbb{Z}^t$ with bounded coefficients $\|\Psi\|_N \leqslant D$ and every convex body $K \subseteq [N]^d$ such that $\Psi(K) \subseteq [N/\overline{W}]^d$ the asymptotic*

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} \nu(\overline{W}\psi_j(n) + b_{i_j}) = 1 + O_D(N^{d-1+O_D(\gamma)}/\operatorname{vol}(K)) + o_D(1)$$

holds, provided γ was small enough.

Proposition 6.2 (Correlation estimate)**.** *For every $1 < m_0 \leqslant D$ there exists a function $\sigma_{m_0} : \mathbb{Z}_{M'} \to \mathbb{R}^+$ with bounded moments $\mathbb{E}_{n\in\mathbb{Z}_{M'}}\sigma^q_{m_0}(n) \ll_{m,q} 1$ such that for every interval $I \subset \mathbb{Z}_{M'}$, every $1 \leqslant m \leqslant m_0$ and every m-tuple $(i_1, \dots, i_m) \in [t]^m$, and every choice of (not necessarily distinct) $h_1, \dots, h_m \in \mathbb{Z}_{M'}$ we have*

$$\mathbb{E}_{n\in I} \prod_{j\in[m]} \nu'(\overline{W}(n + h_j) + b_{i_j}) \leqslant \sum_{1\leqslant i<j\leqslant m} \sigma_{m_0}(h_i - h_j) \ ,$$

provided γ was small enough.

The correlation estimate will be deferred to the next section, the verification of the linear forms condition is an immediate consequence of the following proposition.

Proposition 6.3. *Let $\Psi : \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine-linear forms, such that any exceptional prime, that is, any prime p for which there are ψ_i and ψ_j that are affinely related modulo p, satisfies $p \leqslant w(N)$. Then*

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} \nu'(\psi_j(n)) = 1 + O_D(N^{d-1+O_D(\gamma)}/\operatorname{vol}(K)) + o_D(1)$$

for every convex $K \subseteq [N]^d$ such that $\Psi(K) \subseteq [N]^t$.

Proof of Proposition 6.1. The system Ψ of affine-linear forms that appears in the linear forms condition has the property that no two forms ψ_i, ψ_j are affinely related and it further obeys the coefficient bound $\|\Psi\|_N \leqslant L_0$. Thus every exceptional prime p of Ψ satisfies $p = O_{d,t,L_0}(1)$. We have to show that

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} \nu'(\phi_j(n)) = 1 + O_D(N^{d-1+O_D(\gamma)}/\operatorname{vol}(K)) + o_D(1)$$

with

$$\phi_j(n) = \overline{W}\psi_j(n) + b_{i_j} \ .$$

If $p > w(N)$ is a prime, then ϕ_i and ϕ_j are affinely related modulo p if and only if ψ_i and ψ_j are affinely related modulo p, which proves the result in view of Proposition 6.3. □

Proof of Proposition 6.3. The strategy of the proof is to show that all occurring dependent divisibility events $\prod_{j\in[t]} 1_{a_i|\psi_i(n)}$ where the a_i are not pairwise coprime have a negligible contribution. Removing those, the densities of the remaining events will depend on the respective choice of $a_1, \ldots, a_t$ but are, up to a small error, independent of the ψ_i.

Recalling the definition (5.1) of ν', our task is to show that

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} \left(\sum_{s=2/\gamma}^{(\log\log N)^3} \sum_{i=\log_2 s-2}^{6\log\log\log N} \sum_{u_j\in U(i,s)} 1_{u_j|\psi_j(n)} \tilde{\tau}'_\gamma(\psi_j(n)) \right) = C'^t + O_D\left(\frac{N^{d-1+O_D(\gamma)}}{\mathrm{vol}(K)}\right) + o_D(1).$$

An arbitrary cross term that appears when multiplying out is of the form

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} \sum_{u_j\in U(i_j,s_j)} 2^{s_j} \tilde{\tau}'_\gamma(\psi_j(n)) 1_{u_j|(\psi_j(n))} \ . \qquad (6.1)$$

The sets $U(i,s)$ were defined in the statement of Proposition 4.2. We will make use of two of their properties, namely that any prime divisor p of $u \in U(i,s)$ satisfies $p \gg N^{1/(\log\log N)^3}$ and that $u \leqslant N^\gamma$ for $u \in U(i,s)$.

The removal of dependent divisibility events will be carried out in a sequence of steps. The first is the following claim.

Claim 1. *The cross term* (6.1) *equals*

$$\sum_{u_1,\ldots,u_t}{}' \mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} 2^{s_j} \tilde{\tau}'_\gamma(\psi_j(n)) 1_{u_j|\psi_j(n)} + O_D(N^{-(\log\log N)^{-2}}) \ , \qquad (6.2)$$

where the notation $\sum'_{u_1,\ldots,u_t}$ *indicates that the summation is extended only over pairwise coprime choices of* $u_1, \ldots, u_t$*, where* $u_j \in U(i_j, s_j)$ *for each* j.

Remark. Since the sums over s_j and i_j only have $O_D((\log\log N)^4)$ terms, the total contribution of the error term is bounded by $O_D(N^{1/(\log\log N)^{2-\varepsilon}}) = o_D(1)$.

Proof. All we have to do is to bound the contribution of non-coprime choices of $u_1, \ldots, u_t$ to (6.1). Whenever $(u_i, u_j) > 1$, there is some $p > N^{1/(\log\log N)^3}$ such that $p^2 | \prod_{i\in[t]} \psi_i(n)$. By the properties of the function α, in particular by (2.2), we have

$$\sum_{N^{(\log\log N)^{-3}}<p<N^\gamma} \mathbb{E}_{n\in\mathbb{Z}^d\cap K} 1_{p^2|\prod_i \psi_i(n)} \ll_t \sum_{N^{(\log\log N)^{-3}}<p<N^\gamma} p^{-2} = O_t\left(N^{-(\log\log N)^{-3}}\right) \ .$$

By Cauchy-Schwarz, we have

$$\sum_{\substack{u_1,\dots,u_t\\(u_i,u_j)>1}} \mathbb{E}_{n\in\mathbb{Z}^d\cap K} \prod_{j\in[t]} 2^{s_j}\tilde{\tau}'_\gamma(\psi_j(n))1_{u_j|\psi_j(n)} \leqslant$$

$$\Bigg(\mathbb{E}_{n\in\mathbb{Z}^d\cap K} \sum_{\substack{N^{(\log\log N)^{-3}}\\<p<N^\gamma}} 1_{p^2|\prod_i \psi_i(n)}\Bigg)^{\frac12}\Bigg(\mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]} 2^{2s_j}\tilde{\tau}'^2_\gamma(\psi_j(n))\Big(\sum_{\substack{u_j\in\\U(i_j,s_j)}} 1_{u_j|\psi_j(n)}\Big)^2\Bigg)^{\frac12}$$

Note that the second factor may be bounded by

$$\Big(\mathbb{E}_{n\in\mathbb{Z}^d\cap K}\ 2^{2D(\log\log N)^3}\prod_{j\in[t]}\tau^4(\psi_j(n))\Big)^{1/2} \ll_D 2^{D(\log\log N)^3}(\log N)^{O_D(1)}$$

where the 4th moment estimate of τ_γ may be obtained in a similar manner as the second moment in Section 3. This proves the claim since $(2^{(\log\log N)^3}\log N)^{O_D(1)} \ll N^{1/(\log\log N)}$. ☐

We proceed to analyse (6.2). In particular, we show the following.

Claim 2. *The main term of* (6.2) *satisfies*

$$\begin{aligned}&\sum_{u_1,\dots,u_t}{}' \mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]}2^{s_j}\tilde{\tau}'_\gamma(\psi_j(n))1_{u_j|\psi_j(n)}\\ &= (1+o_D(1))\Big(\frac{W}{\phi(W)\gamma\log N}\Big)^t \sum_{u_1,\dots,u_t}{}' \sum_{v_1|u_1,\dots,v_t|u_t}\ \sum_{\substack{d_1,\dots,d_t\\ d_j\leqslant N^\gamma/v_j\\(d_j,u_jW)=1\\j=1,\dots,t}}{}' \prod_{j\in[t]}\frac{2^{s_j}}{u_j}\frac{1}{d_j} \qquad (6.3)\\ &\quad + O(N^{d-1+O_t(\gamma)}/\operatorname{vol}(K))\ .\end{aligned}$$

Remark. Similar as with the previous claim, the fact that the sums over s_j and i_j only have $O_D((\log\log N)^4)$ terms implies that the overall contribution of the error terms from here is still $O(N^{d-1+O_t(\gamma)}/\operatorname{vol}(K))$.

Proof. Inserting the definition of $\tilde{\tau}'_\gamma$, multiplying by its normalisation, and applying the volume packing lemma, Lemma 2.1, we have

$$\begin{aligned}&\Big(\frac{W}{\phi(W)\gamma\log N}\Big)^{-t}\sum_{u_1,\dots,u_t}{}' \mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]}2^{s_j}\tilde{\tau}'_\gamma(\psi_j(n))1_{u_j|\psi_j(n)}\\ &= \sum_{u_1,\dots,u_t}{}' \mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]}2^{s_j}\sum_{v_j|u_j}\sum_{\substack{d_j\leqslant N^\gamma/v_j\\(d_j,u_jW)=1}}1_{d_ju_j|\psi_j(n)}\\ &= \sum_{u_1,\dots,u_t}\prod_{j\in[t]}2^{s_j}\sum_{v_1|u_1,\dots,v_t|u_t}\ \sum_{\substack{d_1,\dots,d_t\\d_i\leqslant N^\gamma/u_i\\(d_i,u_iW)=1\\i=1,\dots t}}\left\{\frac{\alpha(d_1,\dots,d_t)}{u_1\dots u_t}+O\Big(\frac{N^{d-1}}{\operatorname{vol}(K)}\operatorname{lcm}(u_1d_1,\dots,u_td_t)\Big)\right\}.\end{aligned}$$

The error term is of order $O(N^{d-1+O_t(\gamma)}/\operatorname{vol}(K))$. Since $2^{s_j}\leqslant 2^{(\log\log N)^3}$, since $W/(\phi(W)\gamma\log N)\ll 1$ and since the sums over the u_j and v_j have altogether $N^{O_t(\gamma)}$ terms, the total contribution of the error term is also given by $O(N^{d-1+O_t(\gamma)}/\operatorname{vol}(K))$.

Concerning the main term, Lemma 2.2 allows us to also pass to only summing over pairwise coprime choices of $d_1, \dots d_t$: Fix a choice of $u = (u_1, \dots u_t)$ and $v = (v_1, \dots, v_t)$ and let $\mathcal{D}_{u,v}$ be the set of all tuples $(d_1, \dots, d_t)$ satisfying $(d_i, u_i W) = 1$ and $d_i \leqslant N^\gamma / v_i$ for $i = 1, \dots t$. With this notation, the sum over $d_1, \dots, d_t$ in the main theorem, for fixed u, v, satisfies

$$\begin{aligned}
\sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \alpha(d_1,\dots,d_t) &\leqslant \sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \alpha(d_1,\dots,d_t) \\
&\leqslant \sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \alpha(d_1,\dots,d_t) \prod_{\substack{p\nmid d_1\dots d_t \\ p>w(N)}} \Bigg(1 + \sum_{\substack{a_1,\dots,a_t \\ \text{at least two } a_i\neq 0}} \alpha(p^{a_1},\dots,p^{a_t})\Bigg) \\
&\leqslant \sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \alpha(d_1,\dots,d_t) \prod_{p>w(N)} (1 + O_D(p^{-2})) \\
&\leqslant (1 + O_D(1/w(N))) \sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \alpha(d_1,\dots,d_t) \\
&= (1 + O_D(1/w(N))) \sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \frac{1}{d_1\dots d_t} \ ,
\end{aligned}$$

which implies the claim. □

The last remaining step will be to show that, picking up only another $(1 + o_{t,d,L}(1))$ factor, we can move the product over j in front in the term (6.3).

Claim 3. *Summing all terms* (6.3), *we have*

$$\begin{aligned}
&\sum_{s_1,\dots,s_t}\sum_{i_1,\dots,i_t}\left(\frac{W}{\phi(W)\gamma\log N}\right)^t \sideset{}{'}\sum_{u_1,\dots,u_t}\sum_{v_1|u_1,\dots,v_t|u_t}\sideset{}{'}\sum_{\substack{(d_1,\dots,d_t)\\ \in\mathcal{D}_{u,v}}}\prod_{j\in[t]}\frac{2^{s_j}}{u_j}\frac{1}{d_j} \\
&= (1+o_D(1))\prod_{j\in[t]}\sum_{s_j=2/\gamma}^{(\log\log N)^3}\sum_{i_j=\log_2 s_j-2}^{6\log\log\log N}\sum_{u_j\in U(i_j,s_j)}\frac{2^{s_j}}{u_j}\frac{W}{\phi(W)\gamma\log N}\sum_{\substack{d_j\leqslant N^\gamma/v_j\\ (d_j,u_jW)=1}}\frac{1}{d_j} \qquad (6.4)\\
&\quad + O_t(N^{-1/(\log\log N)^3}) \ .
\end{aligned}$$

Proof. The new expression (6.4) includes additional terms containing non-coprime tuples $u_1, \dots, u_t$ or $d_1, \dots, d_t$. To see that these terms only contribute an additional $(1 + o_D(1))$ factor, first consider the d_j's: Note that

$$\prod_{j\in[t]} \frac{1}{d_j} \leqslant \alpha(d_1, \dots, d_t) \ .$$

Thus, an application of Lemma 2.2, similar to the one for the previous claim, yields

$$\sideset{}{'}\sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \prod_{j\in[t]} \frac{2^{s_j}}{u_j}\frac{1}{d_j} = (1 + o_{t,d,L}(1)) \sum_{(d_1,\dots,d_t)\in\mathcal{D}_{u,v}} \prod_{j\in[t]} \frac{2^{s_j}}{u_j}\frac{1}{d_j} \ .$$

It remains to show that we can also drop the coprimality condition on the u_j's. The contribution to (6.4) from non-coprime choices $u_1, \dots, u_t$ can be bounded as follows.

Suppose $(u_{j'}, u_{j''}) > 1$. Then in particular $(u_{j'}, u_{j''}) > N^{1/(\log\log N)^3}$, since any prime factor of a u_j is greater than $N^{1/(\log\log N)^3}$ by definition. Thus

$$\prod_{j\in[t]} \frac{2^{s_j}}{u_j} \leqslant \frac{1}{N^{1/(\log\log N)^3}} \prod_{j\in[t]:j\neq j',j''} \frac{2^{s_j}}{u_j}\left(\frac{2^{2s_{j'}}}{u_{j'}} + \frac{2^{2s_{j''}}}{u_{j''}}\right) .$$

Since

$$\frac{W}{\phi(W)\gamma\log N} \sum_{\substack{d_j\leqslant N^\gamma/v_j\\(d_j,u_jW)=1}} \frac{1}{d_j} \ll 1$$

the contribution to (6.4) from bad $(u_i)_{i\in[t]}$ is at most

$$\binom{t}{2}\frac{1}{N^{1/(\log\log N)^3}} \prod_{j\in[t-1]} \sum_{s_j\geqslant 2/\gamma} \sum_{\substack{i_j\geqslant\\ \log_2 s_j-2}} \sum_{u_j\in U(i_j,s_j)} \frac{2^{2s_j}}{u_j} \ll_t \frac{1}{N^{1/(\log\log N)^3}} ,$$

where the convergence of the three nested sums follows from the proof of Proposition 4.2. This proves the claim. □

To summarise, we have shown that

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]}\nu'(\psi_j(n)) = \mathbb{E}_{n\in\mathbb{Z}^d\cap K}\prod_{j\in[t]}\left(\sum_{s=2/\gamma}^{(\log\log N)^3}\sum_{i=\log_2 s-2}^{6\log\log\log N}\sum_{u_j\in U(i,s)} 1_{u_j|\psi_j(n)}\tilde{\tau}'_\gamma(\psi_j(n))\right)$$

$$= (1+o_D(1))\prod_{j\in[t]}\frac{W}{\phi(W)\gamma\log N}\left(\sum_{s=2/\gamma}^{(\log\log N)^3}\sum_{i=\log_2 s-2}^{6\log\log\log N}\sum_{u\in U(i,s)}\frac{2^s}{u}\sum_{v|u}\sum_{\substack{d\leqslant N^\gamma/v\\(d,uW)=1}}\frac{1}{d}\right)$$

$$+O\left(\frac{N^{d-1+O_D(\gamma)}}{\mathrm{vol}(K)}\right) + o_D(1) .$$

Regarding the last equation in the special and already known case $\mathbb{E}_{n\leqslant N}\nu'(n) = 1+o(1)$ of the linear forms condition implies that each of the factors on the right hand side, which is independent of Ψ, equals $C'(1+o(1))$. This completes the proof of Proposition 6.3.

7. The correlation condition

This section provides a proof of Proposition 6.2.

Due to the similar structure of our majorant to that of the majorant used in [5, 6], the function σ_m can be chosen in the same manner as in [5, 6].

Proposition 7.1 (Green-Tao [5]). *Let $\Delta : \mathbb{Z}\to\mathbb{Z}$ be the polynomial defined by $\Delta(n) = \prod_{1\leqslant j<j'\leqslant m}(\overline{W}n + b_{i_j} - b_{i'_j})$, define $\sigma_m : \mathbb{Z}_{M'}\to\mathbb{R}^+$ to be*

$$\sigma_m(n) := \exp\left(\sum_{p>w(N),\ p|\Delta(n)} O_m(p^{-1/2})\right) .$$

for $n>0$ and suppose $\sigma_m(0) = o(M')$. Then $\mathbb{E}_{n\in\mathbb{Z}_{M'}}\sigma_m^q(n) \ll_{m,q} 1$.

Proof of Proposition 6.2. The proof proceeds in two cases. The first case considers the situation where $h_i = h_j$ for two distinct indices i, j. We aim to use the fact that on the right hand side of the inequality

$$\mathbb{E}_{n\in\mathbb{Z}_{M'}} \prod_{i\in[m]} \nu'_{w,(b_i)}(n+h_i) \leqslant \sum_{1\leqslant i<j\leqslant m} \sigma_{m_0}(h_i-h_j)$$

$\sigma_{m_0}(0)$ occurs while Proposition 7.1 allows us to choose $\sigma_{m_0}(0)$ quite large. Indeed, Hölder's inequality yields

$$\mathbb{E}_{n\in\mathbb{Z}_{M'}} \prod_{i\in[m]} \nu'_{\overline{W},(b_i)}(n+h_i) \leqslant \prod_{j\in[m]} \left(\mathbb{E}_{n\in\mathbb{Z}_{M'}} \nu'(\overline{W}n+b_{i_j})^m\right)^{1/m} .$$

Any cross term of

$$\mathbb{E}_{n\in\mathbb{Z}_{M'}} \nu'(\overline{W}n+b_{i_j})^m$$

is of the form

$$\begin{aligned} &C'^{-m}\mathbb{E}_{n\leqslant M'} \left(\tilde{\tau}'_\gamma(\overline{W}n+b_{i_j})\right)^m \prod_{j\in[m]} \sum_{u_j\in U(i_j,s_j)} 2^{s_j} 1_{u_j|(\overline{W}n+b_{i_j})} \\ &\leqslant C'^{-m}\mathbb{E}_{n\leqslant M'} \left(\tilde{\tau}'_\gamma(\overline{W}n+b_{i_j})\right)^m \sum_{j\in[m]} \sum_{u_j\in U(i_j,s_j)} 2^{ms_j} 1_{u_j|(\overline{W}n+b_{i_j})} .\end{aligned}$$

Since $\tau(n) \ll_\varepsilon n^\varepsilon$, we may continue this estimate by

$$\begin{aligned} &\ll_\varepsilon C'^{-m} \exp(\varepsilon m \log N) \mathbb{E}_{n\leqslant M'} \sum_{j\in[m]} \sum_{u_j\in U(i_j,s_j)} 2^{ms_j} 1_{u_j|(\overline{W}n+b_{i_j})} \\ &\leqslant C'^{-m} \exp(\varepsilon m \log N) \sum_{j\in[m]} \sum_{u_j\in U(i_j,s_j)} \frac{2^{ms_j}}{u_j} .\end{aligned}$$

Note that the proof of Proposition 4.2 implies that

$$\sum_{s>2/\gamma} \sum_{i_j\geqslant \log_2 s-2} \sum_{u_j\in U(i_j,s)} \frac{2^{ms_j}}{u_j}$$

converges. Thus, summing over all cross terms yields

$$\mathbb{E}_{n\in\mathbb{Z}_{M'}} \nu'(\overline{W}n+b_{i_j})^m \ll_{m,\varepsilon} \exp(\varepsilon m \log N) .$$

If $\varepsilon > 0$ is small enough, setting

$$\sigma_{m_0}(0) := O_{m_0,\varepsilon}\Big(\exp(\varepsilon m_0 \log N)\Big) = o(N/\log N) = o(M')$$

we can ensure that

$$\mathbb{E}_{n\in\mathbb{Z}_{M'}} \prod_{i\in[m]} \nu'_{w,(b_i)}(n+h_i) \leqslant \sum_{1\leqslant i<j\leqslant m} \sigma_{m_0}(h_i-h_j)$$

when $h_i = h_j$ for some $i \neq j$.

Next, we consider the case where $h_i \neq h_j$ whenever $i \neq j$. Our approach to estimate

$$\mathbb{E}_{n\leqslant N} \prod_{j\in[m]} \nu'\left(\overline{W}(n+h_j)+b_{i_j}\right)$$

is the same as the one used to check the linear forms condition and we therefore proceed to analyse the local divisor densities: Since the forms $\psi_j(n) = \overline{W}(n+h_j)+b_{i_j}$ are affinely related, all we can say in general is

$$\alpha(p^{a_1}, \dots, p^{a_m}) = O(p^{-\max_i a_i}) \ .$$

If, however, more than one exponent a_i is non-zero, then we have

$$\alpha(p^{a_1}, \dots, p^{a_m}) > 0$$

only if $p \mid \left(\overline{W}(h_j - h_{j'}) + b_{i_j} - b_{i_{j'}}\right)$ for some $j, j' \in [m]$.

Claim. *We have the following estimate*

$$\mathbb{E}_{n\leqslant N} \prod_{j\in[m]} \nu'\left(\overline{W}(n+h_j)+b_{i_j}\right) \ll \prod_{p|\Delta} \sum_{a_1,\dots,a_m} \alpha(p^{a_1}, \dots, p^{a_m}) \tag{7.1}$$

where $\Delta := \prod_{j\neq j'}(\overline{W}(h_j - h_{j'}) + b_{i_j} - b_{i_{j'}})$.

Before we prove the claim, we complete the verification of the correlation estimate. In order to apply the bound on α, note that there are at most jm^{j-1} tuples $(a_1, \dots a_t)$ satisfying $\max_i a_i = j$. Since $p > w(N)$ on the right hand side of (7.1), we have for p large enough that

$$jm^{j-1} < p^{j/2}/2 \ .$$

Thus, for such p

$$\sum_{a_1,\dots,a_m} \alpha(p^{a_1}, \dots, p^{a_m}) \leqslant 1 + \frac{1}{2}p^{-1/2} \sum_{j\geqslant 0} p^{-j/2} \leqslant 1 + p^{-1/2}$$

and therefore

$$\prod_{p|\Delta} \sum_{a_1,\dots,a_m} \alpha(p^{a_1}, \dots, p^{a_m}) \ll \prod_{\substack{p>w(N)\\ p|\Delta}} \left(1 + p^{-1/2}\right) \ .$$

Let $\Delta(n) := \prod_{j\neq j'} \left(\overline{W}n + b_{i_j} - b_{i_{j'}}\right)$ and set

$$\sigma_{m_0}(n) := \exp\left(\sum_{p>w(N),\ p|\Delta(n)} O_{m_0}(p^{-1/2}) \right) \ .$$

for $n > 0$. Since $1 + x \leqslant \exp x$, we have

$$\mathbb{E}_{n\leqslant N} \prod_{j\in[m]} \nu'\left(\overline{W}(n+h_j)+b_{i_j}\right) \ll_m \sum_{1\leqslant j<j'\leqslant m} \sigma_{m_0}\left(h_j - h_{j'}\right) \ .$$

In view of the above Proposition 7.1, this completes the verification of the correlation condition.

Proof of Claim. Consider, similarly to the previous section, an arbitrary cross term

$$C'^{-t}\mathbb{E}_{n\leqslant M'} \prod_{j\in[m]} \sum_{u_j\in U(i_j,s_j)} 2^{s_j} \tilde{\tau}'_\gamma(\overline{W}(n+h_j)+b_{i_j}) 1_{u_j|((\overline{W}(n+h_j)+b_{i_j}))}$$

of $\mathbb{E}_{n\leqslant M'}\prod_{j\in[m]}\nu'\left(\overline{W}(n+h_j)+b_{i_j}\right)$. Dropping the normalising factor $\left(\frac{W}{C'\phi(W)\gamma\log N}\right)^t$ for the moment, the cross term may be rewritten as

$$\mathbb{E}_{n\leqslant M'}\sum_{u_1,\dots,u_m}\sum_{v_1|u_1,\dots,v_m|u_m}\sum_{\substack{d_1,\dots,d_m\\ d_i\leqslant N^\gamma/u_i\\ (d_j,u_jW)=1\\ i=1,\dots,m}}\prod_{j\in[m]}2^{s_j}1_{u_jd_j|(\overline{W}(n+h_j)+b_{i_j})}$$

$$\leqslant \mathbb{E}_{n\leqslant M'}\sum_{u_1,\dots,u_m}\left(\sum_{k=1}^m 2^{ms_k}\tau(u_k)1_{u_k|(\overline{W}(n+h_k)+b_{i_k})}\right)\prod_{j\in[m]}\sum_{\substack{d_j\leqslant N^\gamma\\ (d_j,u_jW)=1}}1_{d_j|(\overline{W}(n+h_j)+b_{i_j})}\ .$$

Since $(d_j,u_jW)=1$, this may, employing the volume packing lemma, be bounded by

$$\ll \sum_{u_1,\dots,u_m}\left(\sum_{k=1}^m 2^{ms_k}\frac{\tau(u_k)}{u_k}\right)\sum_{\substack{d_1,\dots,d_m\leqslant N^\gamma\\ (d_j,u_jW)=1}}\alpha(d_1,\dots,d_m)$$

$$\ll \sum_{u_1,\dots,u_m}\left(\sum_{k=1}^m 2^{ms_k}\frac{\tau(u_k)}{u_k}\right)\prod_{p_1>w(N)}(1+p_1^{-1})^t\prod_{p_2|\Delta}\sum_{a_1,\dots,a_m}\alpha(p_2^{a_1},\dots,p_2^{a_m})\ .$$

Summing over all cross terms and noting that

$$\left(\frac{W}{C'\phi(W)\gamma\log N}\right)^t\prod_{p_1>w(N)}(1+p_1^{-1})^t\ll 1$$

and that (c.f. the proof of Proposition 4.2)

$$\begin{aligned}\sum_{s>2/\gamma}\sum_{i\geqslant\log_2 s-2}\sum_{u\in U(i,s)}2^{ms}\frac{\tau(u)}{u}&=\sum_{s>2/\gamma}\sum_{i\geqslant\log_2 s-2}\sum_{u\in U(i,s)}2^{ms}2^{m_0(i,s)}\frac{1}{u}\\ &\leqslant\sum_{s>2/\gamma}\sum_{i\geqslant\log_2 s-2}2^{ms}\frac{1}{m_0(i,s)!}(2\log 2+o(1))^{m_0(i,s)}\\ &\leqslant\sum_{s>2/\gamma}\frac{2^{ms}}{s^{s\gamma/100}}\sum_{j\geqslant1}\left(\frac{100\cdot e\cdot(2\log 2+o(1))}{\gamma j}\right)^{\gamma sj/100}\\ &\ll 1\end{aligned}$$

completes the proof of the claim.

8. Application of the transference principle

The aim of this section is to deduce the main theorem from a generalised von Neumann theorem and to prove some reductions on the remaining task of checking that the conditions of the generalised von Neumann theorem are satisfied.

The transference principle [6, Prop. 10.3] allows, as was discussed in Section 5, to transfer results that hold for bounded Gowers-uniform functions to Gowers-uniform functions that are dominated by a pseudorandom measure. It was developed in [5, §8] in view of an application to the (unbounded) von Mangoldt function, and was proved by an iteration argument. New and simplified approaches to the transference principle were more recently found by Gowers [4] and Reingold-Tulsiani-Trevisan-Vadhan [14].

The generalised von Neumann theorem asserts that, if f is suitably Gowers-uniform and dominated by a pseudorandom measure, then composing f with linear forms ψ_i that are sufficiently independent yields functions $f \circ \psi_i$ that behave like independent variables in the sense that $\mathbb{E}_n \prod_{i\in[t]} f(\psi_i(n))$ is close to $(\mathbb{E}_n f(n))^t$, which is the expected value, had the $f \circ \psi_i$ genuinely been independent.

Proposition 8.1 (Green-Tao [6], generalised von Neumann theorem)**.** *Let t, d, L be positive integer parameters. Then there are constants C_1 and D, depending on t, d and L, such that the following is true. Let C, $C_1 \leqslant C \leqslant O_{t,d,L}(1)$ be arbitrary and suppose that $N' \in [CN, 2CN]$ is a prime. Let $\nu : \mathbb{Z}_{N'} \to \mathbb{R}^+$ be a D-pseudorandom measure, and suppose that $f_1, \dots, f_t : [N] \to \mathbb{R}$ are functions with $|f_i(x)| \leqslant \nu(x)$ for all $i \in [t]$ and $x \in [N]$. Suppose that $\Psi = (\psi_1, \dots, \psi_t)$ is a finite complexity system of affine-linear forms with $\|\Psi\|_N \leqslant L$. Let $K \subset [-N, N]^d$ be a convex body such that $\Psi(K) \subset [N]^t$. Suppose also that*

$$\min_{1\leqslant j\leqslant t} \|f_j\|_{U^{t-1}[N]} \leqslant \delta \tag{8.1}$$

for some $\delta > 0$. Then we have

$$\sum_{n\in K} \prod_{i\in[t]} f_i(\psi_i(n)) = o_\delta(N^d) + \kappa(\delta)N^d \ .$$

Establishing the Gowers-uniformity condition (8.1) itself is a task that is conceptually equivalent to that of finding an asymptotic for $\sum_{n\in K} \prod_{i\in[t]} f(\psi_i(n))$ directly, and should therefore not be any easier. The specific system of affine-linear forms that appears in the definition of the uniformity norms, however, allows an alternative characterisation of Gowers-uniform functions.

A CHARACTERISATION OF GOWERS-UNIFORM FUNCTIONS. Whether or not a function f is Gowers-uniform is characterised by the non-existence or existence of a polynomial nilsequence[2] that correlates with f. On the one hand, correlation with a nilsequence obstructs uniformity:

Proposition 8.2 (Green-Tao [6], Cor. 11.6)**.** *Let $s \geqslant 1$ be an integer and let $\delta \in (0, 1)$ be real. Let $G/\Gamma = (G/\Gamma, d_{G/\Gamma})$ be an s-step nilmanifold with some fixed smooth metric $d_{G/\Gamma}$, and let $(F(g(n)\Gamma))_{n\in\mathbb{N}}$ be a bounded s-step nilsequence with Lipschitz constant at most L. Let $f : [N] \to \mathbb{R}$ be a function that is bounded in the L_1-norm, that is, assume $\|f\|_{L_1} = \mathbb{E}_{n\in[N]}|f(n)| \leqslant 1$. If furthermore*

$$\mathbb{E}_{n\in[N]} f(n)F(g(n)\Gamma) \geqslant \delta$$

then we have

$$\|f\|_{U^{s+1}[N]} \gg_{s,\delta,L,G/\Gamma} 1 \ .$$

An inverse result to this statement has been known as Inverse Conjecture for the Gowers norms (GI(s) conjectures) for some time and has recently been resolved, see [9]. The inverse conjectures are stated for bounded functions. With our application to the normalised divisor function in mind, we only recall the transferred statement, c.f. [6, Prop. 10.1], here.

[2]For definitions of nilmanifolds and nilsequences, see, for instance, [8].

Proposition 8.3 (Green-Tao [6], Relative inverse theorem for the Gowers norms)**.** *For any $0 < \delta \leqslant 1$ and any $C \geqslant 20$, there exists a finite collection $\mathcal{M}_{s,\delta,C}$ of s-step nilmanifolds G/Γ, each equipped with a metric $d_{G/\Gamma}$, such that the following holds. Given any $N \geqslant 1$, suppose that $N' \in [CN, 2CN]$ is prime, that $\nu : [N'] \to \mathbb{R}^+$ is an $(s+2)2^{s+1}$-pseudorandom measure, suppose that $f : [N] \to \mathbb{R}$ is any arithmetic function with $|f(n)| \leqslant \nu(n)$ for all $n \in [n]$ and such that*

$$\|f\|_{U^{s+1}[N]} \geqslant \delta \ .$$

Then there is a nilmanifold $G/\Gamma \in M_{s,\delta,C}$ in the collection and a 1-bounded s-step nilsequence $(F(g(n)\Gamma))_{n\in\mathbb{N}}$ on it that has Lipschitz constant $O_{s,\delta,C}(1)$, such that we have the correlation estimate

$$|\mathbb{E}_{n\in[N]} f(n)F(g(n)\Gamma)| \gg_{s,\delta,C} 1 \ .$$

This inverse theorem now reduces the required uniformity-norm estimate (8.1) to the potentially easier task of proving that the centralised version of f does not correlate with polynomial nilsequences.

Reduction of the main theorem to a non-correlation estimate. The task of proving the main theorem had been reduced to the proof of the following proposition in Section 5.

Proposition 5.2. *Let $M = N/\overline{W}$ and suppose $K' \subset [M]^d$ is a convex body. Then for any choice $b_1, \dots, b_t \in [\overline{W}]$*

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K'} \prod_{i\in[t]} \tilde{\tau}'(\overline{W}\psi_i(n) + b_i) = 1 + o_{d,t,L}(M^d/\operatorname{vol}(K'))$$

holds for all finite complexity systems Ψ of affine-linear forms satisfying $\|\Psi\|_M \leqslant L$.

Define for $b \in [\overline{W}]$ the function $\tilde{\tau}'_{\overline{W},b} : \mathbb{Z} \to \mathbb{R}$, $\tilde{\tau}'_{\overline{W},b}(n) := \tilde{\tau}'(\overline{W}n + b)$. Rewriting

$$\mathbb{E}_{n\in\mathbb{Z}^d\cap K'} \prod_{i\in[t]} \tilde{\tau}'_{\overline{W},b_i}(\psi_i(n)) - 1 = \mathbb{E}_{n\in\mathbb{Z}^d\cap K'} \prod_{i\in[t]} \Big(\Big(\tilde{\tau}'_{\overline{W},b_i}(\psi_i(n)) - 1\Big) + 1\Big) - 1$$

and multiplying out the product, the constant term cancels out, while all other terms are of a form the generalised von Neumann theorem can be applied to, provided we can show that

$$\|\tilde{\tau}'_{\overline{W},b_i} - 1\|_{U^{t-1}} = o(1)$$

for all $i \in [t]$. By the inverse theorem, it thus suffices to establish the non-correlation estimates

$$|\mathbb{E}_{n\in[N]}(\tilde{\tau}'_{\overline{W},b}(n) - \mathbb{E}\tau'_{w,b})F(g(n)\Gamma)| = o(1)$$

for all $(t-2)$-step nilsequences $F(g(n)\Gamma)$ as in Proposition 8.3 and all $b \in [\overline{W}]$.

9. Non-correlation of the W-tricked divisor function with nilsequences

The aim of this section is to provide the remaining non-correlation estimate which will complete the proof of the main theorem. For all concepts and notation in connection with nilmanifolds and nilsequences that remain undefined in this section we refer to [7] and its companion paper [8].

Let $k \geqslant 1$ be an arbitrary integer, let $F : G/\Gamma \to \mathbb{C}$ be a Lipschitz function on the $(k-1)$-step nilmanifold G/Γ, and let $g : \mathbb{Z} \to G$ be a polynomial nilsequence adapted to some given filtration $G_\bullet$ of G.

Since the W-tricked divisor function does not count divisors with prime factors $p < w(N)$, let $\varpi(n) := \prod_{p^a \| n, p \leqslant w(N)} p^a$ denote for any integer n the largest divisor composed of primes $< w(N)$.

Let $b \in [\overline{W}]$ and note that $\varpi(\overline{W}n + b) = \varpi(b)$. Setting

$$\begin{aligned}\mu_{\overline{W},b} &:= \frac{W}{\phi(W)\log N} \sum_{\substack{d \leqslant (N/\varpi(b))^{1/2} \\ (W,d)=1}} 2(d^{-1} - d/N) \\ &= \frac{W}{\phi(W)\log N} \sum_{\substack{d \leqslant N/\varpi(b) \\ (W,d)=1}} d^{-1} + o(1) = (1+o(1))\mathbb{E}_{n \leqslant M} \tilde{\tau}'_{\overline{W},b}(n) ,\end{aligned}$$

the application of the inverse theorem for the Gowers norm requires the estimation

$$\begin{aligned}&\mathbb{E}_{n \leqslant M} (\tilde{\tau}'_{\overline{W},b}(n) - \mu_{\overline{W},b}) F(g(n)\Gamma) \\ &= 2\mathbb{E}_{n \leqslant M} \sum_{d \leqslant (N/\varpi(b))^{1/2}} (1_{d | \overline{W}n+b} 1_{\overline{W}n+b > d^2} - d^{-1}(1 - d^2/N)) F(g(n)\Gamma) \\ &= o_{F,G/\Gamma}(1) .\end{aligned}$$

To achieve this, we shall employ the strategy and various lemmata from [7]. Some parts of the argument will be generalised to meet our requirements.

The basic strategy is as follows. When trying to establish a non-correlation estimate, it is desirable to have good control on the nilsequence involved. This is for instance the case when the nilsequence is totally equidistributed, that is equidistributed in every sufficiently dense subprogression of the range it is defined on. While a nilsequence in general does not have this property, the factorisation theorem from [8] states that any nilsequence $g : [N] \to G$ may be written as a product $g(n) = \varepsilon(n)g'(n)\gamma(n)$ where $\varepsilon : \mathbb{Z} \to G$ is smooth, $g' : \mathbb{Z} \to G'$ is totally equidistributed in a rational subgroup $G' \leqslant G$, and $\gamma : \mathbb{Z} \to G$ is periodic.

The aim then is to show that, by passing to a collection of subsequences defined on subprogressions of $[N]$, the correlation estimate involving g can be reduced to correlation estimates involving totally equidistributed sequences arising from g'.

One further reduction is possible: Any periodic function of short period can be regarded as a nilsequence. Establishing non-correlation in the special case of periodic sequences is likely to be much easier than the general case. If we pass from $\mathbb{E}_{n \leqslant N} f(n)F(g(n)\Gamma)$ to considering the collection $\mathbb{E}_{n \leqslant (N-i)/d} f(dn+i)F(g(dn+i)\Gamma)$ for $0 \leqslant i < d$, where each sequence $g(dn+i)\Gamma$ takes values in some subnilmanifold G_i/Γ_i of G/Γ, then a non-correlation estimate with periodic sequences allows us to assume that the mean values $\int_{G_i/\Gamma_i} F(x)\,dx$ vanish. Indeed, we may subtract off the periodic correlation

$$\mathbb{E}_{n \leqslant N} f(n) \left(\sum_{i=0}^{d-1} 1_{n \equiv i(d)} \int_{G_i/\Gamma_i} F(x) \right) = o(1) ,$$

that is, we may subtract off the relevant mean values.

This sketch shows the rough strategy from §2 of [7] for reducing a non-correlation estimate to the case where the nilsequence is equidistributed and furthermore the involved Lipschitz function F has zero mean.

The following is [7, Thm. 1.1] adapted to our case.

Theorem 9.1. *Let G/Γ be a nilmanifold of some dimension $m \geqslant 1$, let $G_\bullet$ be a filtration of G of some degree $d \geqslant 1$, and let $g \in poly(\mathbb{Z}, G_\bullet)$ be a polynomial sequence. Suppose that G/Γ has a Q-rational Mal'cev basis $\mathcal{X}$ for some $Q \geqslant 2$, defining a metric $d_{\mathcal{X}}$ on G/Γ. Suppose that $F : G/\Gamma \to [-1,1]$ is a Lipschitz function. Recall that $M = N/\overline{W}$ and that the normalising factor of $\tilde{\tau}'_{\overline{W},b}$ depends on N. We have*

$$|\mathbb{E}_{n\in[M]}\tilde{\tau}'_{\overline{W},b}(n)F(g(n))\Gamma| \ll_{m,d,\gamma,A} Q^{O_{m,d,\gamma,A}(1)}(1+\|F\|)(\log\log\log N)^{-A}$$

for any $A > 0$ and $N \geqslant 2$.

Sketch proof: Since $\mathbb{E}_{n\in[M]}|\tilde{\tau}'_{\overline{W},b}(n)| = O(1)$, the theorem trivially holds unless $Q \ll (\log\log\log N)^{O_{A,m,d}(1)} \ll w(N)$, allowing us to assume $Q^B \ll_B w(N)$, for some $B > 1$ to be chosen later.

Proceeding as in §2 of [7], one may reduce to analysing the case where $(g(n)\Gamma)$ is totally Q'^{-B}-equidistributed for some $Q', Q \leqslant Q' \ll Q^{O_{B,m,d}(1)}$ and $\int F = 0$. The major arc estimate that is required to carry through this reduction is the following. For any progression $P \subseteq [M]$ of common difference $1 \leqslant q < w(N)$ we have

$$\begin{aligned}
&\mathbb{E}_{n\in N} 1_P(n)(\tilde{\tau}'_{\overline{W},b}(n) - \mu_{\overline{W},b}) \\
&= \frac{2W}{\phi(W)\log N}\ \mathbb{E}_{n\in N}\ 1_P(n) \sum_{\substack{d\leqslant (N/\varpi(b))^{1/2} \\ (d,W)=1}} \Big(1_{d|\overline{W}n+b}1_{\overline{W}n+b>d^2} - d^{-1}(1-\frac{d^2}{N})\Big) \\
&\ll N^{-1/2}\ .
\end{aligned}$$

Note that this bound critically depends on the assumption $1 \leqslant q < w(N)$.

The case where $(g(n)\Gamma)$ is totally Q'^{-B}-equidistributed and $\int F = 0$ is a consequence of the next proposition (cf. also [7, Proposition 2.1]), applied with $\delta = Q'^{-B}$, provided is provided B was chosen large enough. ☐

Proposition 9.2 ($\tilde{\tau}'_{\overline{W},b}$ is orthogonal to equidistributed nilsequences)**.** *Suppose that G/Γ has a Q-rational Mal'cev basis $\mathcal{X}$ adapted to $G_\bullet$. Suppose $(g(n)\Gamma)_{n\in\mathbb{Z}}$ is a totally δ-equidistributed sequence. Then for any Lipschitz function $F : G/\Gamma \to [-1,1]$ with $\int_{G/\Gamma} F = 0$ and for any progression $P \subset [M]$ of length at least M/Q, we have*

$$|\mathbb{E}_{n\in[M]}(\tilde{\tau}'_{\overline{W},b}(n)1_P(n)F(g(n)\Gamma)| \ll \delta^c Q^{O(1)}\|F\| \log\log\log N\ .$$

For the proof of this proposition we employ tools from the analysis of *Type I* sums in the proof of [7, Proposition 2.1]. In particular, we will use the following generalisation of that analysis, which is significant for our purposes. The proof, however, follows the aforementioned *Type I* sums analysis in large parts, so that we keep, where we do not cite, close to the presentation in [7].

Lemma 9.3. *Suppose that $(g(n)\Gamma)_{n\in\mathbb{Z}}$ is a totally δ-equidistributed sequence, suppose that $\delta > N^{-\sigma}$ for some positive constant σ depending only on m, d, and γ, and suppose also that $\|F\| = 1$ and that $Q \leqslant \delta^{-c_1}$ for some parameter $c_1 \in (0,1)$. Let $P \subseteq [M]$ be* *a progression of length at least M/Q. For any $1 \leqslant K \leqslant N^{1/2}$ there are only $o(\delta^{O(1)})K$ values of k satisfying $k \in (K, 2K]$ and*

$$\left|k^{-1}\mathbb{E}_{N/\overline{W}<n<2N/\overline{W}}\ 1_{k|\overline{W}n+b}1_P(\overline{W}n+b)F(g(n)\Gamma)\right| \gg \delta^{O(1)}\ .$$

Proof. Suppose for contradiction that there is some K, $1 \leqslant K \leqslant N^{1/2}$, such that the following inequality holds for $\gg \delta^{O(c_1)}K$ values of $k \in (K, 2K]$

$$\begin{aligned}&\left|\frac{1}{k}\mathbb{E}_{N/\overline{W}<n<2N/\overline{W}}\ 1_{k|\overline{W}n+b}1_P(\overline{W}n+b)F(g(n)\Gamma)\right|\\ &= \left|\mathbb{E}_{N/\overline{W}k<m<2N/\overline{W}k}\ 1_P(\overline{W}(km+u_k)+b)F(g(km+u_k)\Gamma)\right|\\ &\gg \delta^{O(c_1)}\ ,\end{aligned}$$

where u_k is the smallest integer for which $k|\overline{W}u_k + b$. u_k exists for all k for which the inequality holds. To remove the indicator function of P, let $\ell \leqslant Q$ denote the common difference of P and split the range of m into progressions of common difference ℓ. Pigeonholing shows that there is some residue $b' \pmod{\ell}$ such that we still find $\gg \delta^{O(c_1)}K$ values of $k \in (K, 2K]$ that satisfy

$$\left|\sum_{m'\in I_k} F(g(k(\ell m'+b')+u_k))\right| \gg \delta^{O(c_1)}\frac{N}{\overline{W}k\ell}\ , \tag{9.1}$$

where $I_k \subseteq [N/\overline{W}2k\ell-1, N/\overline{W}k\ell]$ is an interval. This lower bound means that for those k that satisfy (9.1), the sequence $\tilde{g}_k : \mathbb{Z} \to G$ defined by $\tilde{g}_k(n) := g(k(\ell n + b') + u_k)$, fails to be $\delta^{O(c_1)}$-equidistributed in G/Γ on the range $N_k = [N/\overline{W}2k\ell - 1, N/\overline{W}k\ell]$.

By [8, Thm 2.9] there is a non-trivial horizontal character $\psi_k : G \to \mathbb{R}/\mathbb{Z}$ of modulus $|\psi_k| \ll \delta^{O(c_1)}$ such that

$$\|\psi_k \circ \tilde{g}_k\|_{C^\infty[N_k]} \ll \delta^{-O(c_1)}.$$

For notational simplicity we remove the dependence on b and ℓ. This step is not strictly necessary for the proof. Let $g_k : \mathbb{Z} \to G$ be defined by $g_k(n) := g(kn + u_k)$. Then [7, Lemma 8.4] asserts that there is some integer q_k, $1 < q_k \ll \delta^{-O(c_i)}$ such that

$$\|q_k\psi_k \circ g_k\|_{C^\infty[N_k]} \ll \delta^{-O(c_1)}\ .$$

Pigeonholing over the possible choices of horizontal character $q_k\psi_k$, there is some non-trivial ψ of magnitude $|\psi| \ll \delta^{O(c_1)}$ among them such that

$$\|\psi \circ g_k\|_{C^\infty[N_k]} \ll \delta^{-O(c_1)}$$

for $\gg \delta^{O(c_1)}K$ values of $k \in (K, 2K]$. Let

$$\psi \circ g(n) = \beta_d n^d + \cdots + \beta_0$$

be the projection of the polynomial sequence to $\mathbb{R}/\mathbb{Z}$ by the character ψ. Then

$$\psi \circ g_k(n) = \beta_d k^d n^d + (\text{lower order terms in } n)\ .$$

We now consider just the highest coefficients $\beta_d k^d$. As in [7, p.9], one shows that β_d is close to a rational with small denominator, more precisely, that there is some $\tilde{q}, 1 \leqslant \tilde{q} \ll \delta^{-O(c_1)}$

$$\|\tilde{q}\beta_d\|_{\mathbb{R}/\mathbb{Z}} \ll \delta^{-O(O(c_1))}(N/\overline{W})^{-d}\ . \tag{9.2}$$

Behind this is the following: since $\psi \circ g_k$ has small smoothness norm, the coefficients, in particular $\beta_d k^d$, are close to rationals with small denominator. Waring's theorem tells that one can express many integers as a sum of few dth powers. This allows us to show that $\beta_d n$ is strongly recurrent in $\mathbb{R}/\mathbb{Z}$, and hence β_d is close to a rational with small denominator.

The bound (9.2) means that $\beta_d n^d$ is varying very slowly on progressions of common difference $\tilde{q}$. By pigeonholing, one of these progressions, say $\{n \equiv q' \mod \tilde{q}\}$, contains the numbers u_k for at least $\gg \delta^{O(c_1)}K$ of our selection of values $k \in (K, 2K]$ that are also satisfying (9.1).

For each such k, consider the full expansion of $\psi \circ g_k$:

$$\begin{aligned}\psi \circ g_k(n) &= \sum_{j=1}^{d} \beta_j (kn + u_k)^j \\ &= \beta_d k^d n^d + \left(\beta_{d-1} + \binom{d}{1} u_k \beta_d\right) k^{d-1} n^{d-1} \\ &\quad + \left(\beta_{d-2} + \binom{d-1}{1} u_k \beta_{d-1} + \binom{d}{2} u_k^2 \beta_d\right) k^{d-2} n^{d-2} + \dots .\end{aligned}$$

Since $u_k \equiv q' \pmod{\tilde{q}}$, there are integers $a_{d-1}, \dots, a_0$ such that

$$\left\| \binom{d}{j} u_k^j \beta_d - \frac{a_j}{\tilde{q}} \right\|_{\mathbb{R}/\mathbb{Z}} \ll \delta^{-O(c_1)} (N/\overline{W})^{-d}$$

We aim to use this information to remove the appearance of the u_k, which are varying with k in a way we have no control on, from the coefficient of n^{d-1}, hoping to then run a similar argument as before to show that β_{d-1} is close to being rational.

Writing

$$\psi \circ g_k(n) = \sum_{j=1}^{d} \tilde{\beta}_{j,k} k^j n^j ,$$

the assertion

$$\|q \tilde{\beta}_{j,k} k^j\|_{\mathbb{R}/\mathbb{Z}} \ll (N/\overline{W}K)^{-j} \|\psi \circ g_k\|_{C^\infty[N_k]} \ll (N/\overline{W}K)^{-j} \delta^{-O(c_1)}$$

holds if and only if

$$\left\| q\left(\tilde{\beta}_{j,k} - \binom{d}{j} u_k^j \beta_d + \frac{a_j}{\tilde{q}}\right) k^j \right\|_{\mathbb{R}/\mathbb{Z}} \ll (N/\overline{W}K)^{-j} \|\psi \circ g_k\|_{C^\infty[N_k]} \ll (N/\overline{W}K)^{-j} \delta^{-O(c_1)} .$$

Thus, we can remove all occurrence of β_d in the $\tilde{\beta}_j$ for $j < d$. For $j = d-1$ this also removes all occurrences of u_k since

$$\tilde{\beta}_{d-1,k} = \beta_{d-1} + \binom{d}{1} u_k \beta_d ,$$

We proceed inductively: We know that there is $q = O(1)$ such that for $\gg \delta^{O(c_1)}K$ values of k from our selection of $k \in (K, 2K]$ the following holds

$$\|q k^{d-1} (\beta_{d-1} + \frac{a_{d-1}}{\tilde{q}})\|_{\mathbb{R}/\mathbb{Z}} \ll (N/\overline{W}K)^{-d+1} \delta^{-O(c_1)} .$$

As before, one deduces via Waring's theorem that $\beta_{d-1} + \frac{a_{d-1}}{\tilde{q}}$, and hence β_{d-1}, is close to a rational with small denominator, say $\tilde{\tilde{q}}$. Pass to a subprogression of common difference

$\tilde{\tilde{q}}$ such that for many of our k the number u_k belongs to that subprogression, note that we can remove the appearance of β_{d-1} in all $\tilde{\beta}_j$ for $j < d-1$, and the appearance of u_k in $\tilde{\beta}_{d-2}$. Show that β_{d-2} is close to a rational with small denominator and repeat.

Finally, we see that there is $\bar{q}, 1 \leqslant \bar{q} \ll \delta^{-O(c_1)}$ such that

$$\|\bar{q}\beta_j\|_{\mathbb{R}/\mathbb{Z}} \ll \delta^{-O(c_1)} N^{-j} \ .$$

This means that $\|\bar{q}\psi \circ g(n)\|_{\mathbb{R}/\mathbb{Z}}$ is small on a reasonably long interval: exactly as in [7], we have for fixed small $\varepsilon > 0$ ($\varepsilon = 1/10$)

$$\|\bar{q}\psi \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \ll n\delta^{-O(c_1)} N^{-1} \leqslant \varepsilon$$

for all $n \leqslant N' = \delta^{Cc_1} N$ provided C is large enough.

By taking $\tilde{F}$ to be a composition of $\bar{q}\psi$ with a smooth cut-off of the interval $[-\varepsilon, \varepsilon]$ one constructs a Lipschitz function $\tilde{F} : G/\Gamma \to [-1, 1]$ for which

$$|\mathbb{E}_{n\in[N']}\tilde{F}(g(n)\Gamma)| \geqslant 1 > \delta\|\tilde{F}\|_{\mathrm{Lip}} \ .$$

If c_1 is chosen small enough, this contradicts the assumption that g was δ-equidistributed and, hence, proves the Lemma. □

Proof of Proposition 9.2. Exactly as in the proof of [7, Proposition 2.1], one shows that the result is trivially true in all cases that are not covered by the assumptions of Lemma 9.3.

Since $(g(n)\Gamma)_{n\in\mathbb{Z}}$ is totally δ-equidistributed and since $\int_{G/\Gamma} F = 0$, it suffices to show that

$$|\mathbb{E}_{n\in[M]}\tilde{\tau}'_{\overline{W},b}(n) 1_P(n) F(g(n)\Gamma)| \ll \delta^{O(1)} \log\log\log N \ .$$

This, however, follows from Lemma 9.3 via dyadic summation:

$$\begin{aligned}
&\mathbb{E}_{n\in[M]}\tilde{\tau}'_{\overline{W},b}(n) 1_P(n) F(g(n)\Gamma) \\
&= \frac{W}{\phi(W)\log N} \sum_{\substack{d<(N/\varpi(b))^{1/2} \\ (d,W)=1}} d^{-1}\ \mathbb{E}_{n\in[N/d]} 1_{d|\overline{W}n+b} 1_P(n) F(g(n)\Gamma) \\
&= \frac{W}{\phi(W)\log N} \sum_{\substack{j\leqslant \\ \frac{1}{2}\log_2(N/\varpi(b))}} \sum_{\substack{d\sim 2^j \\ (d,W)=1}} \sum_{\ell : d^2 \leqslant 2^\ell \leqslant N} \frac{2^{\ell-1-\log_2 N}}{d} \mathbb{E}_{n\in[2^{\ell-1},2^\ell]} 1_{d|\overline{W}n+b} 1_P(n) F(g(n)\Gamma) \\
&\ll \frac{W}{\phi(W)\log N}\left(\sum_{\substack{d<N^{1/2} \\ (d,W)=1}} d^{-1}\delta^{O(1)} + \log_2 N o(\delta^{O(1)}) \right) \\
&\ll \delta^{O(1)} \log w(N) \\
&\ll \delta^{O(1)} \log\log\log N \ .
\end{aligned}$$

□

Acknowledgements. I should like to thank my PhD supervisor Ben Green for suggesting the original problem that led to the results of this paper, Ben Green and Tom Sanders for many valuable discussions and encouragement, and Tim Browning for drawing my attention to the paper [13].

References

[1] B. J. Birch, *Multiplicative functions with non-decreasing normal order,* J. London Math. Soc. **42** (1967), no. 1, 149–151.

[2] T. D. Browning *The divisor problem for binary cubic forms,* arXiv:1006.3476v1, preprint.

[3] P. Erdős, *On the sum* $\sum_{k=1}^{x} d(f(k))$, J. London Math. Soc. **27** (1952), no. 1, 7–15.

[4] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem,* arXiv:0811.3103v1, preprint.

[5] B. J. Green and T. C. Tao, *The primes contain arbitrarily long arithmetic progressions,* Annals of Math. **167** (2008), No. 2, 481–547.

[6] ———, *Linear equations in primes,* Annals of Math., **171** (2010), No. 3 , 1753–1850.

[7] ———, *The Möbius function is strongly orthogonal to nilsequences,* arXiv:0807.1736v3, preprint.

[8] ———, *The quantitative behaviour of polynomial orbits on nilmanifolds,* Annals of Math., to appear. Preprint available at arXiv:0709.3562v5.

[9] B. J. Green, T. C. Tao and T. Ziegler, *An inverse theorem for the Gowers* $U^k[N]$ *norm,* in preparation.

[10] G. H. Hardy, E. M. Wright *An Introduction to the Theory of Numbers,* Oxford University Press, Fifth Edition, 1979.

[11] A. E. Ingham. *Some asymptotic formulae in the theory of numbers,* J. London Math. Soc., **2** (1927), no. 3, 202–208.

[12] L. Matthiesen, *Linear correlations amongst numbers represented by positive definite binary quadratic forms,* in preparation.

[13] M. Nair, G. Tenenbaum, *Short sums of certain arithmetic functions,* Acta Math. **180** (1998), no. 1, 119–144.

[14] O. Reingold, L. Trevisan, M. Tulsiani and S. Vadhan, *Dense Subsets of Pseudorandom Sets,* In: Proceedings of the 49th IEEE Symposium on Foundations of Computer Science (2008), 76–85.

[15] G. Tenenbaum, *Lois de répartition des diviseurs. IV,* Ann. Inst. Fourier (Grenoble) **29** (1979), no. 3, 1–15.

Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
E-mail address: L.Matthiesen@dpmms.cam.ac.uk