

Algebra & Number Theory

Volume 19
2025
No. 10

Smooth numbers are orthogonal to nilsequences

Lilian Matthiesen and Mengdi Wang



Smooth numbers are orthogonal to nilsequences

Lilian Matthiesen and Mengdi Wang

The aim of this paper is to study distributional properties of integers without large or small prime factors. Define an integer to be $[y', y]$ -smooth if all of its prime factors belong to the interval $[y', y]$. We identify suitable weights $g_{[y', y]}(n)$ for the characteristic function of $[y', y]$ -smooth numbers that allow us to establish strong asymptotic results on their distribution in short arithmetic progressions. Building on these equidistribution properties, we show that (a W -tricked version of) the function $g_{[y', y]}(n) - 1$ is orthogonal to nilsequences. Our results apply in the almost optimal range $(\log N)^K < y \leq N$ of the smoothness parameter y , where $K \geq 2$ is sufficiently large, and to any $y' < \min(\sqrt{y}, (\log N)^c)$.

As a first application, we establish for any $y > N^{1/\sqrt{\log_9 N}}$ asymptotic results on the frequency with which an arbitrary finite complexity system of shifted linear forms $\psi_j(\mathbf{n}) + a_j \in \mathbb{Z}[n_1, \dots, n_s]$, $1 \leq j \leq r$, simultaneously takes $[y', y]$ -smooth values as the n_i vary over integers below N .

1. Introduction	1881
2. Outline of the proof and overview of the paper	1886
3. Smooth numbers	1888
4. Smooth numbers in short intervals	1892
5. Major arc analysis: smooth numbers in arithmetic progressions and a W -trick	1898
6. Weyl sums for numbers without small and large prime factors	1911
7. Strongly recurrent polynomial sequences over smooth numbers	1917
8. Noncorrelation with nilsequences: a reduction and general lemmas	1922
9. Noncorrelation with equidistributed nilsequences	1932
10. The proof of Theorem 1.3	1940
Acknowledgements	1945
References	1945

1. Introduction

Let $y > 0$ be a real number. A positive integer n is called y -smooth if its largest prime factor is at most y . The y -smooth numbers below N form a subset of the integers below N which is, in general, sparse but enjoys good equidistribution properties in arithmetic progressions and short intervals. These distributional properties turn y -smooth numbers into an important technical tool for many arithmetic questions. As an example for one of the striking applications of smooth numbers within analytic number theory, we mention [Vaughan 1989], which introduced smooth numbers in combination with a new iterative method to the

MSC2020: primary 11N37; secondary 11B30, 11D04, 11L15, 11N25.

Keywords: smooth numbers, multiplicative functions, nilsequences, Gowers uniformity norms, higher-order Fourier analysis.

study of bounds in Waring's problem. Wooley [1992] extended these methods and achieved substantial improvements on Waring's problem by working with smooth numbers. We refer to Granville's survey [2008] (see in particular Section 6) for a more comprehensive overview of applications of smooth numbers within number theory.

In the present paper, we prove new results on the equidistribution of smooth numbers in short intervals and arithmetic progressions. Our principal aim is to prove higher uniformity of y -smooth numbers in a sense that will be made precise below and for y ranging over an almost optimal range. In addition, we prove that provided y is not too small, the set of y -smooth numbers is sufficiently well distributed to guarantee the existence of nontrivial solutions to arbitrary finite complexity systems of linear equations.

In order to state our main results precisely, we first introduce a subset of the y -smooth numbers as well as a weighted version of its characteristic function that are both central to the rest of this paper. Given any real numbers $0 < y' \leq y$, we may consider the set of y -smooth numbers that are free from prime factors smaller than y' . We call such numbers $[y', y]$ -smooth and denote their set by

$$S([y', y]) := \{n \in \mathbb{N} : p \mid n \Rightarrow p \in [y', y]\}.$$

Given any $x > 0$, the subset of $[y', y]$ -smooth numbers $\leq x$ and its cardinality are denoted by

$$S(x, [y', y]) := S([y', y]) \cap [1, x] \quad \text{and} \quad \Psi(x, [y', y]) := |S(x, [y', y])|.$$

Our notation extends the following standard notation for y -smooth numbers:

$$S(y) := S([1, y]), \quad S(x, y) := S(x, [1, y]), \quad \text{and} \quad \Psi(x, y) = \Psi(x, [1, y]).$$

With this notation, we define the weighted characteristic function

$$g_{[y', y]}(n) = \frac{n}{\alpha(n, y)\Psi(n, [y', y])} \mathbf{1}_{S([y', y])}(n) \quad (n \in \mathbb{N})$$

of $[y', y]$ -smooth numbers, where $\alpha(n, y)$ denotes the saddle point¹ associated to $S(n, y)$. If $1 \leq A < W$ are coprime integers, we further define a W -tricked version of $g_{[y', y]}$ by setting

$$g_{[y', y]}^{(W, A)}(m) = \frac{\phi(W)}{W} g_{[y', y]}(Wm + A) \quad (m \in \mathbb{N}).$$

In the definitions of the functions $g_{[y', y]}$ and $g_{[y', y]}^{(W, A)}$, the normalisations are chosen such that their average values are roughly 1.

Following these preparations, we are now ready to state the main result of this paper. The notation around nilsequences will be recalled in Section 8. Throughout this paper we write $\log_k x$ to denote the k -fold iterated logarithm of x .

Theorem 1.1 (higher uniformity). *Let N be a large positive parameter and let $K' \geq 1$, $K > 2K'$ and $d \geq 0$ be integers. Let $\frac{1}{2} \log_3 N \leq y' \leq (\log N)^{K'}$ and suppose that $(\log N)^K < y < N^n$ for some sufficiently*

¹We recall the definition in Section 3.

small $\eta \in (0, 1)$ depending on the value of d . Let $(G/\Gamma, G_\bullet)$ be a filtered nilmanifold of complexity Q_0 and degree d . Finally, let $w(N) = \frac{1}{2} \log_3 N$, $W = \prod_{p < w(N)} p$ and define $\delta(N) = \exp(-\sqrt{\log_4 N})$.

If K is sufficiently large depending on the degree d of G_\bullet , then the estimate

$$\left| \frac{W}{N} \sum_{n \leq (N-A)/W} (g_{[y', y]}^{(W, A)}(n) - 1) F(g(n)\Gamma) \right| \ll_d (1 + \|F\|_{\text{Lip}}) \delta(N) Q_0 + \frac{1}{\log w(N)} \quad (1-1)$$

holds uniformly for all $1 \leq A \leq W$ with $\gcd(A, W) = 1$, all polynomial sequences $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ and all 1-bounded Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$.

Remark 1.2. Since $\gcd(A, W) = 1$, we may, when working with $Wn + A \in S([y', y])$, restrict without loss of generality to the case $y' \geq w(N)$. In applications where y is not too small, the contribution from $S(y) \setminus S([w(N), y])$ can often be taken care of separately, leading to a result for $S(y)$ in the end.

Theorem 1.1 constitutes the first of two parts necessary in order to establish asymptotic results on the number of $[y', y]$ -smooth solutions to finite complexity systems of linear forms using the nilpotent circle method. We emphasise that the setting of smooth numbers studied in the present paper is significantly more difficult than those settings considered in previous applications such as, e.g., [Green and Tao 2010], [Matthiesen 2018; 2020] or [Matthiesen 2012], which concern primes, a large class of multiplicative functions, and numbers representable by binary quadratic forms, respectively. The reason for this increase in difficulty lies partly in the sparsity of the set of y -smooth numbers and partly in the unavailability of sieve methods to study this set.

The parameter $w(N)$ is determined by the distribution of y -smooth numbers in arithmetic progressions and chosen in such a way that the subset of y -smooth numbers in any fixed reduced residue class modulo $W(N) = \prod_{p < w(N)} p$ is sufficiently well-distributed in arithmetic progressions. At the end of Section 2, we will discuss in more detail the need for applying such a “ W -trick” when studying y -smooth numbers. In addition to the W -trick, the value y' may be used to further influence how well the resulting set of $[y', y]$ -smooth numbers is distributed in progressions. Increasing the parameter y' beyond the value of $w(N)$ leads to better error terms for larger moduli in the distribution of $[y', y]$ -smooth numbers in progressions. The proof of Theorem 1.1 relies on information on the distribution of $[y', y]$ -smooth numbers in short intervals and in arithmetic progressions, which we establish in Sections 4 and 5.

From a technical perspective, our focus in Theorem 1.1 has been to establish a result in which the lower bound on the range of the smoothness parameter y is as small as possible in terms of N , while the W -trick (i.e., the value of W) is still independent of y . This allows one to combine this result in applications with inductive or recursive arguments in the y -parameter. We remark that when focussing on larger values of y , the function $w(N)$ in the statement can be chosen larger and consequently the bounds in this result improve.

As a first application, we estimate, for large values of y , the frequency with which an arbitrary finite complexity system of shifted linear forms $\psi_j(\mathbf{n}) + a_j \in \mathbb{Z}[n_1, \dots, n_s]$, $1 \leq j \leq r$, simultaneously takes $[y', y]$ -smooth values as the n_i vary over integers below N .

Theorem 1.3 (linear equations in smooth numbers). *Let $r, s \geq 2$ be integers and let $N > 2$ be a parameter. We consider the following setup:*

(i) *Let $\psi_1, \dots, \psi_r : \mathbb{Z}^s \rightarrow \mathbb{Z}$ be linear forms that are pairwise linearly independent over \mathbb{Q} , and let $a_1, \dots, a_r \in [-N, N] \cap \mathbb{Z}$ be integers. Let L denote the maximum of the absolute values of the coefficients of ψ_1, \dots, ψ_r .*

(ii) *Let $\mathfrak{K} \subseteq [-1, 1]^s$ be a fixed convex set of positive s -dimensional volume $\text{vol } \mathfrak{K} > 0$, and suppose that the dilated copy*

$$N\mathfrak{K} = \{Nk \in \mathbb{R}^s : k \in \mathfrak{K}\}$$

satisfies $\psi_j(N\mathfrak{K}) + a_j \subseteq [1, N]$ for all $1 \leq j \leq r$.

(iii) *Let $\frac{1}{2} \log_3 N \leq y' \leq (\log N)^{K'}$ for some fixed $K' \geq 1$. Let $\eta \in (0, 1)$ be sufficiently small in terms of r and s and suppose that $y \leq N^\eta$.*

If $y \leq N^\eta$ is sufficiently large to ensure that $\Psi(N, y) > N/\log_8 N$, then

$$\sum_{\mathbf{n} \in \mathbb{Z}^s \cap N\mathfrak{K}} \prod_{j=1}^r g_{[y', y]}(\psi_j(\mathbf{n}) + a_j) = \text{vol}(\mathfrak{K})N^s \prod_{p < y'} \beta_p + o_{r,s,L}(N^s),$$

as $N \rightarrow \infty$ and where

$$\beta_p = \frac{1}{p^s} \sum_{\mathbf{u} \in (\mathbb{Z}/p\mathbb{Z})^s} \prod_{j=1}^r \frac{p}{p-1} \mathbf{1}_{\psi_j(\mathbf{u}) + a_j \not\equiv 0 \pmod{p}}.$$

We note as an aside that, as shown in [Green and Tao 2010, Section 4], this result implies an asymptotic count of the number of solutions in $[y', y]$ -smooth numbers to systems of linear equations satisfying the non-degeneracy condition stated in [loc. cit., Theorem 1.8]. Theorem 1.3 is the first instance of a result of its kind for smooth numbers that applies in sparse situations where $\Psi(N, y) = o(N)$. We mainly include this result for illustration and remark that we have not tried to optimise the lower bound on $\Psi(N, y)/N$.² Theorem 1.3 follows by combining Theorem 1.1 with a “trivial” majorising function. Once suitable majorising functions are available on the full range of y on which Theorem 1.1 applies, a version of Theorem 1.3 will follow on that range of y , making Theorem 1.3 redundant. For this reason, we chose not to optimise the bounds.

The following unweighted version of Theorem 1.3 is an easy consequence of an asymptotic lower bound on the weight factor that appears in the definition of $g_{[y', y]}(n)$.

Corollary 1.4. *With the notation and under the assumptions of Theorem 1.3 the following holds. If $y \leq N^\eta$ is sufficiently large to ensure that $\Psi(N, y) > N/\log_8 N$ and $\frac{1}{2} \log_3 N \leq y' \leq (\log N)^{K'}$ for some fixed $K' \geq 1$, then the number $\mathcal{N}(N, \mathfrak{K})$ of $\mathbf{n} \in \mathbb{Z}^s \cap N\mathfrak{K}$ for which the given system of linear polynomials $(\psi_j(\mathbf{n}) + a_j)_{1 \leq j \leq r}$ takes simultaneous $[y', y]$ -smooth values satisfies*

$$\mathcal{N}(N, \mathfrak{K}) := \sum_{\mathbf{n} \in \mathbb{Z}^s \cap N\mathfrak{K}} \prod_{j=1}^r \mathbf{1}_{S([y', y])}(\psi_j(\mathbf{n}) + a_j) \gg \text{vol}(\mathfrak{K})N^{s-r} \Psi(N, [y', y])^r \prod_p \beta_p$$

for all sufficiently large N .

²As this application only involves fairly large values in y , the bounds in Theorem 1.3 could be improved by establishing better bounds in Theorem 1.1 for large values of y .

Previous work and discussion. Before turning towards Theorem 1.1, we shall first discuss previous work in the direction of Theorem 1.3, partly because this work illustrates for which range of the smoothness parameter y one can hope to prove such a result.

In the dense case where $y > N^\varepsilon$ for any $\varepsilon > 0$, Theorem 1.3 has been proved by Lachand [2017] with $S([y', y])$ replaced by $S(y)$ and with $W = 1$, $A = 0$. The special case of Lachand's result where $s = 2$ and $a_1 = \dots = a_r = 0$ also follows from work of Balog, Blomer, Dartyge and Tenenbaum [Balog et al. 2012]. The latter paper studies smooth values of binary forms and can be applied to $F(n_1, n_2) = \prod_{i=1}^r \psi_i(n_1, n_2)$. In a similar spirit, Fouvry [2010] investigated smooth values of absolutely irreducible polynomials $F(\mathbf{X}) \in \mathbb{Z}[X_1, \dots, X_n]$ with $n \geq 2$ and $\deg F \geq 3$ and showed that

$$\sum_{\mathbf{n} \in \mathbb{Z}^n \cap [-N, N]} \mathbf{1}_{S(N^{d-\delta})(F(\mathbf{n}))} \gg N^n$$

for all $\delta < \frac{4}{3}$ and all $N > N_0(F)$, which also corresponds to the dense setting.

Concerning the sparse setting, i.e., smaller values of y , it follows from [Lagarias and Soundararajan 2011; 2012] that the lower threshold on y for which our counting function satisfies

$$\mathcal{N}(N, \mathfrak{R}) \rightarrow \infty, \quad \text{as } N \rightarrow \infty,$$

is $y > (\log N)^\kappa$ for some $\kappa \geq 1$. More precisely, Lagarias and Soundararajan studied primitive solutions to the equation $A+B=C$ in smooth numbers. Using our notation they proved in [Lagarias and Soundararajan 2012, Theorem 1.5] that GRH implies

$$\sum_{\substack{0 < n_1 + n_2 \leq N \\ \gcd(n_1, n_2) = 1}} \mathbf{1}_{S(y)}(n_1) \mathbf{1}_{S(y)}(n_2) \mathbf{1}_{S(y)}(n_1 + n_2) \gg \frac{\Psi(N, y)^3}{N} \quad (1-2)$$

for any $y \geq (\log N)^\kappa$ and $\kappa > 8$. On the other hand, they showed in [loc. cit., Theorem 1.1] that the *abc* conjecture implies that the left-hand side above is bounded independent of N if $y = (\log N)^\kappa$ and $\kappa < 1$. In view of this latter result, and with applications in the spirit of Theorem 1.3 in mind, the range of y in which we prove Theorem 1.1 is optimal up to the size of the exponent K . Indeed, the left-hand side above is a special case of the expression $\mathcal{N}(N, \mathfrak{R})$ studied in Corollary 1.4 when ignoring the parameter y' . When combined with a simple majorising function, Theorem 1.3 and Corollary 1.4 follow from Theorem 1.1, the Green–Tao–Ziegler inverse theorem [Green et al. 2012] for the U^k -norms³ and the generalised von Neumann theorem [Green and Tao 2010]. The existence of a suitable majorising function for $g_{[y', y]}$ on a larger range of y would imply an analogue of Theorem 1.3 on the intersection of that range and the range on which Theorem 1.1 holds. Such an analogue can only hold provided that $y > (\log N)^K$ for some sufficiently large K .

³To be precise, we require the quantitative version of the inverse theorem due to Manners [2018] in order to deduce explicit bounds from a very simple majorising function. A majorising function of the correct average order on a larger range of y would allow us to use the original qualitative inverse theorem [Green et al. 2012] instead. Since the first version of our paper appeared on the arXiv, a very substantial improvement to [Manners 2018] has been obtained by Leng, Sah and Sawhney [Leng et al. 2024]. We note that, in view of the bounds we obtain in Theorem 1.1, the range of y to which Theorem 1.3 and Corollary 1.4 apply would, without further optimisation, not significantly improve by exchanging in our proof the application of [Manners 2018] for [Leng et al. 2024].

Unconditional versions of the result (1-2) of [Lagarias and Soundararajan 2012, Theorem 1.5], but with more restrictive ranges for y , have been established by a number of authors. Harper [2016, Corollary 1] significantly improved those ranges and showed, almost matching the conditional range, that if $K > 1$ is sufficiently large then the asymptotic formula

$$\sum_{0 < n_1 + n_2 \leq N} \mathbf{1}_{S(y)}(n_1) \mathbf{1}_{S(y)}(n_2) \mathbf{1}_{S(y)}(n_1 + n_2) = \frac{\Psi(N, y)^3}{2N} \left(1 + O\left(\frac{\log(u+1)}{\log y}\right) \right)$$

holds for all $y \geq (\log N)^K$, where $u = (\log N)/\log y$.

Concerning smooth solutions to Diophantine equations in many variables, there is a vast literature on Waring's problem in smooth numbers. Building on [Harper 2016] mentioned before, Drappeau and Shao [2016] studied Waring's problem in y -smooth numbers for $y \geq (\log N)^K$ and $K > 1$ sufficiently large, which is the currently largest possible range of y . We refer to that paper for references to previous work on this question.

Turning towards Theorem 1.1, the closest previous result is due to Lachand [2017] who proved, starting out from a Möbius inversion formula for $\mathbf{1}_{S(y)}$, that

$$\sum_{n \leq N} \left(\mathbf{1}_{S(y)} - \frac{\Psi(N, y)}{N} \right) F(g(n)\Gamma) = o((1 + \|F\|_{\text{Lip}})\Psi(N, y))$$

for any $\varepsilon > 0$ and uniformly for $x \geq y \geq \exp((\log N)/(\log \log N)^{1-\varepsilon})$. We shall explain at the end of Section 2 why, when restricting to large values of y , neither a W -trick nor the restriction to integers with prime factors in the interval $[y', y]$ is necessary.

As exponential phases $e(\theta n^k)$ for $\theta \in \mathbb{R}$, $k \in \mathbb{N}$, form a special case of the nilsequences $F(g(n)\Gamma)$ in Theorem 1.1, previous work on this question includes all work on exponential sum (and Weyl sum) estimates over smooth numbers. The latter is a well-studied subject. Our proof indeed builds on some of this previous work, in particular on that of [Harper 2016; Drappeau and Shao 2016; Wooley 1995].

2. Outline of the proof and overview of the paper

The proof of Theorem 1.1 occupies most of this paper. After recalling some of the background on smooth numbers in Section 3, we investigate in Section 4 the distribution of $[y', y]$ -smooth numbers in short intervals and the results we prove here may be of independent interest. The results of this section motivate the definition of the weighted function $g_{[y', y]}(n)$ as well as an auxiliary weighted function $h_{[y', y]}(n)$, and we prove that these functions are equidistributed in *short intervals*. After extending in Section 5.1 work of Harper on the distribution of y -smooth numbers in progressions, we deduce in Section 5.2 that a W -tricked version of $g_{[y', y]}(n)$ is equidistributed in *short arithmetic progressions*. The work in Sections 4 and 5 (i.e., the fact that $g_{[y', y]}(n)$ is equidistributed in short arithmetic progressions) prepares the ground for the proof of the noncorrelation estimate (1-1). It allows us in Section 8 to reduce the task of proving (1-1) to the case in which the polynomial sequence g is “equidistributed”. Using a Montgomery–Vaughan-type decomposition, this task is further reduced in Section 9.2 to a Type II sum estimate. More precisely, we

reduce our task to bounding bilinear sums of the (slightly simplified) form

$$\sum_{m \in [y', y]} \sum_{N/m < n \leq (N+N_1)/m} \mathbf{1}_{S([y', y])}(n) \Lambda(m) F(g(mn)\Gamma).$$

Swapping the order of summation reduces this problem to that of bounding the correlation of the (W -tricked) von Mangoldt functions with equidistributed nilsequences, *provided* we can show an implication of the form

$$(g(m)\Gamma)_{m \leq N} \text{ is equidistributed} \implies (g(mn)\Gamma)_{m \leq N/n} \text{ is equidistributed for most } n \in S(N, [y', y]).$$

A precise version of this implication will be established in Section 8.2. The proof of this implication builds on the material from Sections 5, 6 and 7. More precisely, it relies on a strong recurrence result for polynomial sequences over $[y', y]$ -numbers that is proved in Section 7. The recurrence result in turn relies on Weyl sum estimates for $[y', y]$ -smooth numbers that we establish in Section 6 by extending [Drappeau and Shao 2016]. As the Weyl sum estimates by themselves are in fact too weak for our purposes, we need to combine them with a bootstrapping argument, which requires good bounds on the distribution of smooth numbers in short arithmetic progressions (as established in Section 5) as input. The final section contains the proof of Theorem 1.3.

The most difficult part of this work is to ensure that y can be chosen as small as $(\log N)^K$ in our main result. To achieve this, we need to run the bootstrapping argument in Section 7 in four different stages of iterations, using the full scale of results proved in Sections 4 and 5, and we have to choose our parameters in Sections 8 and 9 very carefully. Similarly, working with y as small as $(\log N)^K$ requires us to use a Montgomery–Vaughan-type reduction to a Type II sum estimate in Section 9 instead of working with the intrinsic decomposition that was used, e.g., in [Harper 2016, Section 3] to establish minor arc estimates for exponential sums over smooth numbers. (The intrinsic decomposition appears in the proofs of Lemmas 6.2 and 6.4.) Harper’s approach, which in principle generalises to the nilsequences setting, involves an application of the Cauchy–Schwarz inequality and reduces the minor arc estimate to bounds on degree-1 Weyl sums $\sum_{n \leq N} e(n\theta)$ for irrational θ . The losses of the Cauchy–Schwarz application in this approach can be compensated for by the strong bounds available for degree-1 Weyl sums for irrational θ . When working with an irrational (=equidistributed) nilsequence $n \mapsto F(g(n)\Gamma)$ instead of $n \mapsto e(n\theta)$ for irrational θ , the savings on the corresponding sum $\sum_{n \leq N} F(g(n)\Gamma)$ are usually much weaker, and in our case too weak in order to compensate for the loss of the Cauchy–Schwarz application when $y = (\log N)^K$.

Smooth numbers and the “ W -trick”. A result of the form of Theorem 1.1 requires as a necessary condition that the function it involves, here $g_{[y', y]}^{(W, A)}$, is equidistributed in arithmetic progressions to small moduli. The reason for this is that additive characters $n \mapsto e(an/q)$, where $a, q \in \mathbb{N}$ and $e(x) = e^{2\pi i x}$, form a special case of the nilsequences $n \mapsto F(g(n)\Gamma)$ that appear in the statement. The function $g_{[y', y]}^{(W, A)}$ is a weighted version of the characteristic function of $[y', y]$ -smooth numbers, restricted to a reduced residue class A modulo W . Both the use of a W -trick, i.e., the restriction to integers of the form $n = Wm + A$, as

well as the restriction to the subset of y -smooth numbers that are coprime to $P(y') := \prod_{p < y'} p$ are a means to ensure equidistribution in progressions. The weight is introduced in order to guarantee equidistribution in short intervals and progressions.

We proceed to explain why the W -trick is used. Combining work of Soundararajan [2008] and Harper [2012b] on the distribution of smooth numbers in progressions with work of de la Bretèche and Tenenbaum [2005] one may deduce that

$$\Psi(x, y; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{S(y)}(n) \sim \frac{1}{\phi(q)} \sum_{\substack{n \leq x \\ \gcd(n, q) = 1}} \mathbf{1}_{S(y)}(n) \sim \frac{\Psi(x, y)}{q} \prod_{p | q} \frac{1 - p^{-\alpha(x, y)}}{1 - p^{-1}}$$

for $(\log x)^2 < y \leq x$, $2 \leq q \leq y^2$, and $(a, q) = 1$, as $\log x / \log q \rightarrow \infty$.

When $\alpha(x, y)$ is sufficiently close to 1, which happens when y is sufficiently close to x , the final product over prime divisors $p | q$ will be approximately 1 and the above asymptotic implies that $S(y)$ is equidistributed in *all* residue classes (reduced and nonreduced) modulo q for all small values of q . The necessary condition for the validity of Theorem 1.1 is met in this situation and no W -trick is required.

Once $\alpha(x, y)$ is no longer close to 1, the product over prime factors $p | q$ in the asymptotic formula above genuinely depends on q . In this case, $S(y)$ is seen to be equidistributed in the *reduced* residue classes modulo a fixed integer q . However, the density of $S(y)$ within a reduced class will differ from that in a nonreduced class modulo q , the latter being obtained by dividing out the common factor and applying the asymptotic formula with q replaced by a suitable divisor of q . To remove this discrepancy between reduced and nonreduced residue classes, one may use a W -trick with a slowly growing function $w(N)$. Fixing a reduced residue $A \pmod{W}$, one is then interested in the count of y -smooth integers of the form $n = W(qm + a) + A$, which satisfies

$$\Psi(x, y; Wq, Wa + A) \sim \frac{\Psi(x, y)}{Wq} \left(\prod_{p | W} \frac{1 - p^{-\alpha(x, y)}}{1 - p^{-1}} \right) \prod_{p' | q, p' \nmid W} \frac{1 - p'^{-\alpha(x, y)}}{1 - p'^{-1}}.$$

If $W = P(w(N))$ is the product of all primes $p < w(N)$, then all prime factors p' that appear in the final product satisfy $p' \geq w(N)$, i.e., are large. If q is not too large, this allows one to show that the final product is asymptotically equal to 1, with an error term that depends on $w(N)$. This ensures, within any fixed reduced residue class $A \pmod{W}$, that the necessary condition for the validity of Theorem 1.1 holds.

3. Smooth numbers

In this section we collect general properties of smooth numbers that we will frequently make use of within this paper. Recall the definitions of the sets $S(y)$, $S([y', y])$, $S(y, x)$ and $S([y', y], x)$ from the introduction. The relative quantity

$$u := \frac{\log x}{\log y} \tag{3-1}$$

frequently appears when describing properties of $S(x, y)$ and $S(x, [y', y])$.

3.1. The summatory function of $\mathbf{1}_{S(y)}$. Let

$$\Psi(x, y) = |S(x, y)| = \sum_{n \leq x} \mathbf{1}_{S(y)}(n)$$

denote the number of y -smooth numbers below x . As a function of x , $\Psi(x, y)$ may be viewed as the summatory function of the characteristic function $\mathbf{1}_{S(y)}$ of y -smooth numbers. The latter function is completely multiplicative and the associated Dirichlet series is given by

$$\zeta(s, y) = \sum_{n \in S(y)} n^{-s} = \prod_{p \leq y} (1 - p^{-s})^{-1} \quad (\Re(s) > 0).$$

Rankin's trick shows that $\Psi(x, y) \leq x^\sigma \zeta(\sigma, y)$ for all $\sigma > 0$. The (unique) saddle point of the function $\sigma \mapsto x^\sigma \zeta(\sigma, y)$ that appears here is usually denoted by $\alpha(x, y)$ and we have $\alpha(x, y) \in (0, 1)$. Hildebrand [1986a, Lemma 4] (see also [Hildebrand and Tenenbaum 1986, Lemma 2]) showed that

$$\alpha(x, y) = 1 - \frac{\log(u \log(u + 1))}{\log y} + O\left(\frac{1}{\log y}\right) \quad (\log x < y \leq x). \quad (3-2)$$

This saddle point can be used in order to asymptotically describe $\Psi(x, y)$. More precisely, Hildebrand and Tenenbaum [1986, Theorems 1, 2] proved that, uniformly for $y/\log x \rightarrow \infty$,

$$\Psi(x, y) = \frac{x^\alpha \zeta(\alpha, y)}{\alpha \sqrt{2\pi \log x \log y}} \left(1 + O\left(\frac{1}{\log(u + 1)} + \frac{1}{\log y}\right)\right), \quad (3-3)$$

where $\alpha = \alpha(x, y) > 0$ and $u = (\log x)/\log y$. In applications, it is frequently necessary to understand the relation between $\Psi(cx, y)$ and $\Psi(x, y)$ as $c > 0$ varies. In this direction, Theorem 3 of [Hildebrand and Tenenbaum 1986] shows that

$$\Psi(cx, y) = \Psi(x, y) c^{\alpha(x, y)} \left(1 + O\left(\frac{1}{u} + \frac{\log y}{u}\right)\right), \quad (3-4)$$

uniformly for $x \geq y \geq 2$ and $1 \leq c \leq y$.

The following related lemma is a consequence of (3-2).

Lemma 3.1. *Suppose that $x > 2$ is sufficiently large and that $\log x < y \leq x$. Then*

$$\alpha(cx, y) - \alpha(x, y) \ll 1/\log y$$

uniformly for all $c \in [1, 2]$.

Proof. We write $cx = x + x'$ and expand out the expression that (3-2) provides for $\alpha(x + x', y)$ when taking into account the definition (3-1) of u . The lemma follows provided we can show that $\log_2(x + x') = \log_2 x + O(1)$ and $\log_3(x + x') = \log_3 x + O(1)$. Both of these estimates may be deduced by repeated application of the identity $\log(t + t') = \log t + \log(1 + t'/t)$ combined with the observation that for $t > 2$ (in particular, for large t) the inequality $t \geq t'$ is preserved in the sense that $T > T'$ if $T := \log t$ and $T' := \log(1 + t'/t)$. Finally, the bound $\log(1 + t'/t) \ll 1$ provides the shape of the error term. \square

3.2. The truncated Euler product. On the positive real line, the Dirichlet series $\zeta(\sigma, y)$ can be estimated and we will frequently use the following lemma which is [Montgomery and Vaughan 2007, Lemma 7.5].

Lemma 3.2. *Suppose that $y \geq 2$. If $\max\{2/\log y, 1 - 4/\log y\} \leq \sigma \leq 1$, then*

$$\prod_{p \leq y} (1 - p^{-\sigma})^{-1} \asymp \log y.$$

If $2/\log y \leq \sigma \leq 1 - 4/\log y$, then

$$\prod_{p \leq y} (1 - p^{-\sigma})^{-1} = \frac{1}{1 - \sigma} \exp\left(\frac{y^{1-\sigma}}{(1 - \sigma) \log y} \left\{1 + O\left(\frac{1}{(1 - \sigma) \log y}\right) + O(y^{-\sigma})\right\}\right).$$

3.3. The summatory function of $1_{S([y', y])}$. De la Bretèche and Tenenbaum [2005] analysed the quantities

$$\Psi_m(x, y) = \sum_{\substack{n \in S(x, y) \\ (n, m) = 1}} 1 \quad \text{and} \quad \frac{\Psi_m(x/d, y)}{\Psi(x, y)}$$

and established asymptotic estimates for these expressions under certain assumptions on m and d . Observe that

$$\Psi_{P(y')}(x, y) = \Psi(x, [y', y])$$

for $P(y') = \prod_{p < y'} p$. We shall require versions of two of the results from [de la Bretèche and Tenenbaum 2005] that are restricted to the case $m = P(y')$ but come with sufficiently good explicit error terms. To state these results, define for any given integer $m \in \mathbb{N}$ the restricted Euler product

$$g_m(s) = \prod_{p | m} (1 - p^{-s}) \quad (s \in \mathbb{C}). \tag{3-5}$$

The following lemma is a consequence of [de la Bretèche and Tenenbaum 2005, Théorème 2.1].

Lemma 3.3. *Let $K' > 0$ and $K > \max(2K', 1)$ be constants, let $x \geq 2$ and $y' \leq (\log x)^{K'}$. Then*

$$\Psi(x, [y', y]) = g_{P(y')}(\alpha(x, y)) \Psi(x, y) \left(1 + O\left(\frac{\log_2 x}{\log x} + \frac{1}{\log y}\right)\right)$$

uniformly for all $(\log x)^K < y \leq x$.

Proof. This result follows from [de la Bretèche and Tenenbaum 2005, Théorème 2.1 and Corollaire 2.2]. We need to verify that the error term is of the shape claimed in the statement above. Let $m = P(y') = \prod_{p < y'} p$. Since $y' < y^{1/2}$ and $\pi(y') \ll y^{1/2}/(\log y)$, we are either in the situation C_1 or C_2 of [loc. cit., Corollaire 2.2].

Observe that in our case $\bar{u} := \min(u, y/\log y) = u = \log x/\log y$ and

$$W_m = W_{P(y')} = \log p_{\pi(y')} \asymp \log y' \ll \log \log x.$$

Hence, $\vartheta_m \ll (\log \log x)/\log y (\ll 1)$.

If C_1 holds, then the error term in the expression for $\Psi_{P(y')}(x, y)$ is bounded by

$$E^* := E_m^*(x, y) \ll \frac{\vartheta_m \log(u + 2)}{u} \ll \frac{(\log_2 x)^2}{\log x}.$$

If condition C_2 holds, then $y^{1/\log(u+2)} \ll \omega(P(y')) < y' \leq (\log x)^{K'}$ and, hence,

$$\log y \ll (\log_2 x) \log(u + 2) \ll (\log_2 x)^2.$$

This shows that $\log(u+2) \asymp \log_2 x$ and

$$E^* \ll \frac{\vartheta_{m_x}^2}{\log(u+2)} \ll \frac{(\log_2 x)^2}{(\log y)^2 \log(u+2)} \ll \frac{1}{\log y},$$

as required. \square

By restricting [de la Bretèche and Tenenbaum 2005, Théorème 2.4] to the situation of the present paper, we obtain:

Lemma 3.4. (i) *If $y' \leq (\log x)^{K'}$ for some constant $K' > 0$, then we have*

$$\Psi_{P(y')}(x/d, y) \ll d^{-\alpha} \Psi_{P(y')}(x, y)$$

uniformly for all $\max(y'^2, (\log x)^2) < y \leq x$, $1 \leq d \leq x/y$.

(ii) *If, moreover, $1 \leq d \leq \exp((\log x)^{1/3})$, the following more precise statement holds. We have*

$$\Psi_{P(y')}(x/d, y) = d^{-\alpha} \Psi_{P(y')}(x, y) \left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right) \right)$$

uniformly for all $\max(y'^2, (\log x)^2) < y \leq x$ and $1 \leq d \leq \exp((\log x)^{1/3})$, where $y' \leq (\log x)^{K'}$ and $K' > 0$ as before.

Proof. Part (i) is a direct application of [de la Bretèche and Tenenbaum 2005, Théorème 2.4(i)] combined with Lemma 3.3, when taking into account that $\delta = 0$ according to the remark following (2.8) of that paper, since $y > \log^2 x$.

Concerning part (ii), Lemma 3.3 implies that the asymptotic formula given in the statement is equivalent to

$$\Psi_{P(y')}(x/d, y) = d^{-\alpha} g_{P(y')}(\alpha(x, y)) \Psi(x, y) \left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right) \right).$$

The statement thus follows from [de la Bretèche and Tenenbaum 2005, Théorème 2.4(ii)] provided we can show that under our assumptions the error term in that result is in fact of the above-stated shape. In the notation of [loc. cit., Théorème 2.4(ii)] we therefore need to show that $h_m = o(1)$ and $(1 - t^2/(2u^2))^{bu} = 1 + o(1)$ as $x \rightarrow \infty$, where $o(1)$ needs to be made explicit in both cases. In the former case, we have

$$h_m = \frac{1}{u_y} + t \frac{(1 + E_m)}{u} + E_m^*(x, y) \ll \frac{\log_2(2y)}{\log y} + \frac{1}{\log_2 x} + \frac{\log d}{\log x} + \frac{\log d}{\log y} \frac{E_m}{u},$$

where we used that $u_y = u + (\log y)/\log(u+2) \gg (\log y)/\log_2(2y)$ (cf. [de la Bretèche and Tenenbaum 2005, (2.18)]) and that $E_m^*(x, y) \ll 1/(\log_2 x)$ by the previous proof. For the term $(1 + E_m)/u$, we obtain

$$E_m = \frac{\vartheta_m(u \log(u+2))^{\vartheta_m}}{1 + \vartheta_m \log(u+2)}.$$

Note that $\vartheta_m = (\log p_{\omega(m)})/\log y = (\log y')/\log y \leq \frac{1}{2}$ for $y \geq y'^2$. Further, $u \log(u+2) \geq 1$. Hence,

$$\frac{E_m}{u} \ll \frac{\vartheta_m \log^{1/2}(u+2)}{u^{1/2}(1 + \vartheta_m \log(u+2))}$$

If $\log y > (\log_2 x)^3$, then

$$\frac{\log d}{\log y} \frac{E_m}{u} \ll \frac{\log d \log^{1/2}(u+2)}{\log y u^{1/2}(\log_2 x)^2} \ll \frac{(\log d)(\log_2 x)^{1/2}}{(\log x)^{1/2}(\log y)^{1/2} \log_2 x} \ll \frac{\log d}{(\log x)^{1/2}(\log_2 x)^2}.$$

If $\log y \leq (\log_2 x)^3$ and since $\log_2 x \ll \log y$, we have

$$\begin{aligned} t \frac{E_m}{u} &\ll t \frac{\vartheta_m \log^{1/2}(u+2)}{u^{1/2}(1 + \vartheta_m \log(u+2))} \ll t \frac{\vartheta_m (\log_2 x)^{1/2} (\log y)^{1/2}}{(\log x)^{1/2} \left(1 + \frac{1}{(\log_2 x)^2} \log(u+2)\right)} \\ &\ll t \frac{(\log_2 x)^2}{(\log x)^{1/2}} \ll \frac{\log d}{\log y} \frac{(\log_2 x)^2}{(\log x)^{1/2}} \ll \frac{(\log d) \log_2 x}{(\log x)^{1/2}}. \end{aligned}$$

It remains to analyse $(1 - t^2/(2u^2))^{bu}$. Since

$$\frac{t}{u} = \frac{\log d}{\log x} = o(1),$$

we have

$$\left(1 - \frac{t^2}{2u^2}\right)^{bu} = \exp\left(-C \frac{t^2}{u}\right) = 1 + O\left(\frac{1}{\log y}\right)$$

for some constant $C > 0$. The final step above follows since

$$\frac{t^2}{u} = \frac{(\log d)^2}{(\log y) \log x} \ll \frac{1}{\log y}$$

by our assumption on d . □

4. Smooth numbers in short intervals

Our aim in this section is to show that a suitably weighted version of the function $\mathbf{1}_{S([y', y])}$ is equidistributed in short intervals. More precisely, we will consider intervals contained in $[x, 2x]$ that are of length at least $x/(\log x)^c$. The error terms in Lemmas 3.3 and 3.4 are too weak in order to allow one to work with intervals as short as $x/(\log x)^c$. Better error terms are, however, available when considering the quantity

$$\frac{\Psi_{P(y')}(x/d, y)}{\Psi_{P(y')}(x, y)} = \frac{\Psi(x/d, [y', y])}{\Psi(x, [y', y])} \quad \text{instead of} \quad \frac{\Psi_{P(y')}(x/d, y)}{\Psi(x, y)}.$$

In the first subsection below we prove asymptotic formulas for the former quotient. In the second subsection we introduce a suitable smooth weight for $\mathbf{1}_{S([y', y])}$ and prove that the correspondingly weighted $[y', y]$ -smooth numbers are equidistributed in short intervals of the above form.

4.1. The local behaviour of $\Psi(x, [y', y])$. In order to analyse the distribution of $S([y', y])$ in short intervals, we need to compare $\Psi(x, [y', y])$ to $\Psi(x(1 + 1/z), [y', y])$. In the notation of Lemma 3.4 this means that our d is very close to 1. Recall that $u = (\log x)/\log y$. We split our analysis below into two cases according to whether u is large or small.

Lemma 4.1 (local behaviour of $\Psi_{P(y')}$ for large u). *Let $x > 2$ be a parameter, let $K' > 0$ and $K \geq \max(2K', 2)$ be constants and suppose that $y' \leq (\log x)^{K'}$ and $(\log x)^K < y < \exp((\log x)^{1/3})$. Let $\alpha = \alpha(x, y)$ denote the saddle point associated to $S(x, y)$. Then*

$$|\Psi(x, [y', y]) - (1 + z^{-1})^{-\alpha} \Psi((1 + z^{-1})x, [y', y])| \ll \Psi(x, [y', y]) \left(\frac{\sqrt{\log y}}{z \log^{1/6} x} + \frac{O_A(1)}{\log^A x} \right)$$

for all $z > 1$.

Proof. We will follow the strategy via Perron's formula employed in [Hildebrand and Tenenbaum 1986, Lemma 8; de la Bretèche and Tenenbaum 2005, Section 4.2], and start by bounding Euler factors from the relevant Dirichlet series. Suppose that $s = \alpha + i\tau$ and $(\log x)^{-1/4} \leq \tau \leq (\log y)^{-1}$. Then⁴ for any $y' \leq p \leq y$, we have

$$\frac{|1 - p^{-\alpha}|}{|1 - p^{-s}|} = \left(1 + \frac{2(1 - \cos(\tau \log p))}{p^\alpha (1 - p^{-\alpha})^2} \right)^{-1/2} \leq \exp\left(-c_1 \frac{1 - \cos(\tau \log p)}{p^\alpha (1 - p^{-\alpha})^2}\right) \leq \exp\left(-c_2 \frac{(\tau \log p)^2}{p^\alpha}\right),$$

where we took advantage of the fact that $1 - p^{-\alpha} \asymp 1$ for all $p \leq y$ since $\alpha > \frac{1}{2}$. Hence,

$$\begin{aligned} \prod_{y' \leq p \leq y} \frac{|1 - p^{-\alpha}|}{|1 - p^{-s}|} &\leq \exp\left(-c_2 \frac{\tau^2}{y^\alpha} \sum_{y' \leq p \leq y} (\log p)^2\right) \leq \exp\left(-c_2 \frac{\tau^2}{y^\alpha} \sum_{y/2 \leq p \leq y} (\log p)^2\right) \\ &\leq \exp(-c_3 \tau^2 y^{1-\alpha} \log y) \leq \exp(-c_4 \tau^2 u \log(u+1) \log y) \\ &\leq \exp(-c_4 \tau^2 \log x). \end{aligned} \tag{4-1}$$

By invoking bounds of the above Euler factors in different regimes of τ , de la Bretèche and Tenenbaum [2005, (4.50)] deduced from Perron's formula that

$$\Psi(x, [y', y]) = \frac{1}{2\pi i} \int_{\alpha-i/\log y}^{\alpha+i/\log y} \zeta_m(s, y) x^s \frac{ds}{s} + O(x^\alpha \zeta_m(\alpha, y) R), \tag{4-2}$$

where

$$R = \exp\left(\frac{-c_5 u}{(\log 2u)^2}\right) + \exp(-(\log y)^{4/3}), \quad m = P(y') = \prod_{p < y'} p,$$

and

$$\zeta_m(s, y) = \prod_{p \leq y, p \nmid m} (1 - p^{-s})^{-1}.$$

The approximation (4-2) applies in our situation since $u > (\log x)^{3/2} > (\log y)^{3/2}$ and $\pi(y') = \omega(P(y')) < (\log x)^{K'} < \sqrt{y}$, which ensures that the conditions [de la Bretèche and Tenenbaum 2005, (4.40)] are satisfied. Under our assumption that $\log x < y < \exp((\log x)^{1/3})$, we have

$$R \ll \exp(-c_5 (\log x)^{1/2}) + \exp(-(\log \log x)^{4/3}) \ll_A (\log x)^{-A}.$$

⁴See the proof of [Hildebrand and Tenenbaum 1986, Lemma 8] for more details.

The bound (4-1) allows us to further truncate the integral in (4-2). More precisely,

$$\begin{aligned} \left| \int_{\alpha \pm i/(\log x)^{1/3}}^{\alpha \pm i/\log y} \zeta_m(s, y) x^s \frac{ds}{s} \right| &\leq \zeta_m(\alpha, y) \frac{x^\alpha}{\alpha} \int_{1/(\log x)^{1/3}}^{1/\log y} \exp(-c_4 \tau^2 \log x) \, d\tau \\ &\leq \frac{\zeta_m(\alpha, y) x^\alpha}{\alpha \log y} \exp(-c_4 (\log x)^{-1/3}) \\ &\leq \Psi(x, [y', y]) \exp(-c_6 (\log x)^{-1/3}), \end{aligned}$$

which implies that

$$\Psi(x, [y', y]) = \frac{1}{2\pi i} \int_{\alpha - i/(\log x)^{1/3}}^{\alpha + i/(\log x)^{1/3}} \zeta_m(s, y) x^s \frac{ds}{s} + O_A \left(\frac{\Psi(x, [y', y])}{(\log x)^A} \right). \tag{4-3}$$

The latter approximation then yields

$$\begin{aligned} \Psi(x, [y', y]) - (1 + z^{-1})^{-\alpha} \Psi(x(1 + z^{-1}), [y', y]) &= \frac{1}{2\pi i} \int_{\alpha - i/(\log x)^{1/3}}^{\alpha + i/(\log x)^{1/3}} \zeta_m(s, y) x^s (1 - (1 + z^{-1})^{s-\alpha}) \frac{ds}{s} + O_A \left(\frac{\Psi(x, [y', y])}{(\log x)^A} \right) \\ &\ll \zeta_m(\alpha, y) x^\alpha \int_{-1/(\log x)^{1/3}}^{1/(\log x)^{1/3}} |1 - (1 + z^{-1})^{i\tau}| \frac{d\tau}{\alpha} + O_A \left(\frac{\Psi(x, [y', y])}{(\log x)^A} \right) \\ &\ll \frac{\zeta_m(\alpha, y) x^\alpha}{z} \int_{-1/(\log x)^{1/3}}^{1/(\log x)^{1/3}} |\tau| \, d\tau + O_A \left(\frac{\Psi(x, [y', y])}{(\log x)^A} \right) \\ &\ll \frac{\zeta_m(\alpha, y) x^\alpha}{z(\log x)^{2/3}} + O_A \left(\frac{\Psi(x, [y', y])}{(\log x)^A} \right) \\ &\ll \Psi(x, [y', y]) \left(\frac{\sqrt{\log y}}{z(\log x)^{1/6}} + O_A((\log x)^{-A}) \right), \end{aligned}$$

as claimed. □

The previous lemma applies to $y < \exp((\log x)^{1/3})$. Below, we establish an analogous result in the complementary range where $\exp((\log x)^{1/4}) \leq y \leq x$. In this case, $1 - \alpha(x', y)$ is very small as x' ranges over $[x, 2x]$, and the error terms in Lemma 3.3 are also very good. This allows us to reduce the problem of asymptotically evaluating $\Psi(x(1 + z^{-1}), [y', y])$ to that of bounding $\Psi(x(1 + z^{-1}), y) - \Psi(x, y)$. Bounds on the latter difference have been established by [Hildebrand 1986b, Theorem 3] in the case where $u = (\log x)/\log y$ is small, which applies to our current situation.⁵

Lemma 4.2 (local behaviour of $\Psi_{P(y')}$ for small u). *Let $x > 2$ be a parameter, suppose that $y' \leq (\log x)^{K'}$ for some fixed $K' > 0$ and that $\exp((\log x)^{1/4}) \leq y \leq x$, and let $\alpha = \alpha(x, y)$ denote the saddle point associated to $S(x, y)$. Then*

$$\left| \Psi(x(1 + z^{-1}), [y', y]) - (1 + z^{-1})^\alpha \Psi(x, [y', y]) \right| \ll \Psi(x, [y', y]) \frac{(\log_2 x)^2}{\log y} \tag{4-4}$$

holds uniformly for $1 \leq z \leq y^{5/12}$.

⁵Hildebrand's theorem has been extended by several authors and we refer to the survey paper [Granville 2008, Section 4.1] for a discussion of these extensions and references.

Proof. Under the assumptions on y it follows from (3-2) that the saddle point $\alpha(x, y)$ is very close to 1. More precisely, writing $u' = (\log x')/\log y$, we have

$$\alpha(x', y) = 1 - \frac{\log(u' \log(u' + 1))}{\log y} + O\left(\frac{1}{\log y}\right) = 1 + O\left(\frac{\log_2 x'}{\log y}\right) = 1 + O\left(\frac{\log_2 x}{(\log x)^{1/4}}\right) \quad (4-5)$$

uniformly for all $x' \in [x, 2x]$ and $y \geq \exp((\log x)^{1/4})$. This implies, in particular, that

$$(1 + z^{-1})^\alpha = (1 + z^{-1})(1 + z^{-1})^{O((\log_2 x)/\log y)} = (1 + z^{-1})\left(1 + O\left(\frac{\log_2 x}{z \log y}\right)\right)$$

for $\alpha = \alpha(x, y)$. Substituting this formula into (4-4) and rearranging the resulting expression reduces our task to that of establishing

$$\Psi(x(1 + z^{-1}), [y', y]) - \Psi(x, [y', y]) = \frac{\Psi(x, [y', y])}{z} \left(1 + O\left(\frac{z(\log_2 x)^2}{\log y}\right)\right).$$

By Lemma 3.3, the difference on the left-hand side satisfies

$$\begin{aligned} & \Psi(x(1 + z^{-1}), [y', y]) - \Psi(x, [y', y]) \\ &= g_{P(y')}(\alpha_z) \Psi(x(1 + z^{-1}), y) - g_{P(y')}(\alpha) \Psi(x, y) + O\left(\Psi(x, [y', y]) \left(\frac{\log_2 x}{\log x} + \frac{1}{\log y}\right)\right), \end{aligned} \quad (4-6)$$

where $\alpha_z = \alpha(x(1 + z^{-1}), y)$ is the saddle point associated to $S(x(1 + z^{-1}), y)$. In order to simplify the main term above, we seek to relate $g_{P(y')}(\alpha_z)$ to $g_{P(y')}(\alpha)$. Following [de la Bretèche and Tenenbaum 2005, Section 3.6], define $\gamma_m(s) := \log g_m(s)$ and let $\gamma'_m(s)$ denote the derivative with respect to s . Since $\alpha_z \leq \alpha$, we then have

$$\log\left(\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)}\right) = \int_{\alpha}^{\alpha_z} \gamma'_{P(y')}(\sigma) d\sigma \leq (\alpha - \alpha_z) \sup_{\alpha_z \leq \sigma \leq \alpha} \gamma'_{P(y')}(\sigma)$$

and shall estimate the latter two factors in turn. Since $1 < 1 + z^{-1} < 2$ for $z > 1$, it follows from (4-5) that

$$\alpha - \alpha_z \ll \frac{\log_2 x}{\log^{1/4} x}.$$

Concerning the second factor, we have

$$\gamma'_{P(y')}(\sigma) = \sum_{p \leq y'} \frac{\log p}{p^\sigma - 1} \ll \sum_{n \leq y'} \frac{\Lambda(n)}{n^\sigma} \ll y'^{1-\sigma} \sum_{n \leq y'} \frac{\Lambda(n)}{n} \ll \log y' \ll \log_2 x$$

since $\log y' \ll \log_2 x$ and $1 - \sigma \ll_\varepsilon \log^{-1/4+\varepsilon} x$ by (4-5). Hence

$$\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)} = \exp\left(\log\left(\frac{g_{P(y')}(\alpha_z)}{g_{P(y')}(\alpha)}\right)\right) = \exp\left(O\left(\frac{(\log_2 x)^2}{\log^{1/4} x}\right)\right) = 1 + O\left(\frac{(\log_2 x)^2}{\log^{1/4} x}\right).$$

Substituting this expression into (4-6), we obtain

$$\begin{aligned} & \Psi(x(1 + z^{-1}), [y', y]) - \Psi(x, [y', y]) \\ &= g_{P(y')}(\alpha) \left\{ \Psi(x(1 + z^{-1}), y) - \Psi(x, y) \right\} + O\left(\Psi(x, [y', y]) \frac{(\log_2 x)^2}{\log y}\right), \end{aligned}$$

where we used Lemma 3.3 and (3-3) to simplify the error term. To the right-hand side, we may now apply Theorems 1 and 3 of [Hildebrand 1986b], which assert that

$$\Psi(x(1+z^{-1}), y) - \Psi(x, y) = \frac{\Psi(x, y)}{z} \left\{ 1 + O_\varepsilon\left(\frac{\log(u+1)}{\log y}\right) \right\}$$

holds uniformly for $\exp((\log_2 x)^{5/3+\varepsilon}) \leq y \leq x$ and $1 \leq z \leq y^{5/12}$. Thus, under our assumptions on z and y ,

$$\begin{aligned} &\Psi(x(1+z^{-1}), [y', y]) - \Psi(x, [y', y]) \\ &= \frac{g_{P(y')}(\alpha)\Psi(x, y)}{z} \left\{ 1 + O\left(\frac{\log(u+1)}{\log y}\right) \right\} + O\left(\Psi(x, [y', y]) \frac{(\log_2 x)^2}{\log y}\right). \end{aligned}$$

Another application of Lemma 3.3 to the right-hand side yields the desired result. □

4.2. Equidistribution of weighted $[y', y]$ -smooth numbers in short intervals. The correction factors in Lemmas 4.1 and 4.2 suggest to consider the smoothly weighted version

$$n \mapsto n^{1-\alpha(x,y)} \mathbf{1}_{S([y',y])}(n)$$

of the characteristic function of $[y', y]$ -smooth numbers. This choice of weight allows us to deduce from these lemmas that the new weighted function is equidistributed in all short intervals of length at least $x(\log x)^{-c}$, a property that is required in Section 7 and, indirectly, in Section 8. For simplicity, we further normalise our weighted function to mean value 1 in the following lemma.

Lemma 4.3 (equidistribution in short intervals). *Let $N > 2$ be a parameter and define the weighted and normalised function*

$$h_{[y',y]}(n) := \frac{N^\alpha}{\Psi(N, [y', y])} \frac{n^{1-\alpha}}{\alpha} \mathbf{1}_{S([y',y])}(n) \quad (N \leq n \leq 2N), \tag{4-7}$$

where $\alpha = \alpha(N, y)$. Let $y' \leq (\log N)^{K'}$ for some fixed $K' > 0$, let $K = \max(2, 2K')$, and suppose that $(\log N)^K < y \leq N$. Then

$$\sum_{N_0 < n \leq N_0 + N_1} h_{[y',y]}(n) = N_1 \left\{ 1 + O\left(\frac{\log_3 N}{\log_2 N}\right) \right\} + O\left(\frac{N}{\log^{1/24} N}\right)$$

uniformly for all $N \leq N_0 < N_0 + N_1 \leq 2N$ such that $N_1 \gg N \exp(-(\log N)^{1/4}/4)$.

Proof. By partial summation,

$$\begin{aligned} \sum_{\substack{n \in S([y',y]) \\ N_0 < n \leq N_0 + N_1}} n^{1-\alpha} &= (N_0 + N_1)^{1-\alpha} \Psi(N_0 + N_1, [y', y]) \\ &\quad - N_0^{1-\alpha} \Psi(N_0, [y', y]) - (1-\alpha) \int_{N_0}^{N_0 + N_1} \frac{\Psi(t, [y', y])}{t^\alpha} dt. \end{aligned} \tag{4-8}$$

We seek to bound the right-hand side with the help of Lemmas 4.1 and 4.2. To start with, suppose that $y < \exp((\log N)^{1/4})$. For $N \leq N_0 < t \leq N_0 + N_1 \leq 2N$, we have $t = N(1 + 1/z)$, where $z > 1$ and

$(t/N)^{-\alpha} \asymp 1$. Hence,

$$\begin{aligned} \Psi(t, [y', y]) &= (t/N)^\alpha \Psi(N, [y', y]) \left(1 + O\left(\frac{\sqrt{\log y}}{z(\log N)^{1/6}} + O_A((\log N)^{-A}) \right) \right) \\ &= (t/N)^\alpha \Psi(N, [y', y]) (1 + O((\log N)^{-1/24})) \end{aligned}$$

for $y < \exp((\log N)^{1/4})$ by Lemma 4.1. Thus,

$$\begin{aligned} \frac{N^\alpha}{\Psi(N, [y', y])} \sum_{\substack{n \in S([y', y]) \\ N_0 < n \leq N_0 + N_1}} n^{1-\alpha} &= (N_0 + N_1) - N_0 - (1 - \alpha) \int_{N_0}^{N_0 + N_1} 1 \, dt + O\left(\frac{N}{(\log N)^{1/24}} \right) \\ &= \alpha N_1 + O\left(\frac{N}{(\log N)^{1/24}} \right), \end{aligned}$$

which establishes the lemma when $y < \exp((\log N)^{1/4})$.

Suppose next that $\exp((\log N)^{1/4}) \leq y \leq N$. In view of the restriction on the size of z in Lemma 4.2, we cannot apply this lemma in the current situation with $x = N$ (which would correspond to the choice in the first part of the proof), but need to choose $x = N_0$ instead. For this purpose, let $\alpha' = \alpha(N_0, y)$ denote the saddle point associated to $S(N_0, y)$ and note that

$$|\alpha' - \alpha| = |\alpha(N_0, y) - \alpha(N, y)| \ll (\log y)^{-1} \ll (\log N)^{-1/4}$$

by Lemma 3.1 and the lower bound on y .

For $N \leq N_0 < t \leq N_0 + N_1$ with $N \exp(-(\log N)^{1/4}/4) \leq N_1 \leq N$, we have

$$t = N_0(1 + 1/z), \quad \text{where } 1 < z \ll \exp((\log N)^{1/4}/4) = o(y^{5/12}),$$

and $(t/N_0)^{-\alpha'} \asymp 1$. Hence, Lemma 4.2 applies and yields

$$\Psi(t, [y', y]) = (t/N_0)^{\alpha'} \Psi(N_0, [y', y]) \left(1 + O\left(\frac{(\log_2 N)^2}{\log^{1/4} N} \right) \right). \tag{4-9}$$

Each of the three terms arising in the partial summation expression (4-8) involves a weighted count of the form $\Psi(t, [y', y])/t^\alpha$. With this and (4-9) in mind, observe that

$$\begin{aligned} (t/N_0)^{\alpha'} t^{-\alpha} &= (t/N_0)^{\alpha' - \alpha} N_0^{-\alpha} = N_0^{-\alpha} (1 + 1/z)^{\alpha' - \alpha} = N_0^{-\alpha} (1 + O(|\alpha - \alpha'|)) \\ &= N_0^{-\alpha} (1 + O((\log N)^{-1/4})), \end{aligned}$$

since

$$|(1 + 1/z)^a - 1^a| \leq \int_1^{1+1/z} |a| t^{a-1} \, dt \leq \frac{|a|}{z} \leq |a|$$

for $0 < |a| < 1$, which we applied with $a = \alpha' - \alpha \ll (\log N)^{-1/4}$.

Hence, it follows from (4-8) and (4-9) that

$$\begin{aligned} \frac{N_0^\alpha}{\Psi(N_0, [y', y])} \sum_{\substack{n \in S([y', y]) \\ N_0 < n \leq N_0 + N_1}} n^{1-\alpha} &= (N_0 + N_1) - N_0 - (1 - \alpha) \int_{N_0}^{N_0 + N_1} 1 \, dt + O\left(\frac{N(\log_2 N)^2}{(\log N)^{1/4}} \right) \\ &= \alpha(N, y) N_1 + O\left(\frac{N(\log_2 N)^2}{(\log N)^{1/4}} \right). \end{aligned}$$

The lemma then follows since

$$\frac{(N_0/N)^{-\alpha} \Psi(N_0, [y', y])}{\Psi(N, [y', y])} = 1 + O\left(\frac{\log_3 N}{\log_2 N}\right)$$

by Lemma 3.3 and part (ii) of Lemma 3.4. □

Working with functions $h_{[y', y]} : [N, 2N] \rightarrow \mathbb{R}$ whose support is restricted to a dyadic interval is too restrictive for the purposes of our main theorem. The following lemma provides an intrinsically defined function $g_{[y', y]} : \mathbb{N} \rightarrow \mathbb{R}$ that approximates any function $h_{[y', y]}$ on the interval $[N, 2N]$ where the latter function is defined.

Lemma 4.4. *Define*

$$g_{[y', y]}(n) = \frac{n}{\alpha(n, y) \Psi(n, [y', y])} \mathbf{1}_{S([y', y])} \quad (n \in \mathbb{N}), \tag{4-10}$$

and suppose that $N \leq n < 2N$ and that $h_{[y', y]}$ is defined on $[N, 2N]$. Then

$$g_{[y', y]}(n) = h_{[y', y]}(n) \left(1 + O\left(\frac{\log_3 N}{\log_2 N}\right)\right),$$

provided that N is sufficiently large.

Proof. This follows immediately from Lemmas 3.1 and 3.4(ii). □

5. Major arc analysis: smooth numbers in arithmetic progressions and a W -trick

Extending previous work of Soundararajan [2008], Harper [2012b] proved that for $y \leq x$, $\varepsilon > 0$, $2 \leq q \leq y^{4\sqrt{e}-\varepsilon}$, and $(a, q) = 1$,

$$\Psi(x, y; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{S(y)}(n) \sim \frac{1}{\phi(q)} \Psi_q(x, y)$$

as $\log x / \log q \rightarrow \infty$ provided that y is sufficiently large in terms of ε . This shows that, for $y \in (\log x, x]$, the y -smooth numbers up to x are equidistributed within the reduced residue classes to a given modulus q provided x is sufficiently large. Our aim in this section is to construct a subset of the y -smooth numbers with the property that its elements are equidistributed in the reduced residue classes of *any* modulus $q \leq (\log x)^c$. More precisely, we will show that, after applying a W -trick, the $[y', y]$ -smooth numbers for any $y' \leq (\log N)^{K'}$, $K' > 0$, have this property. In addition, we will construct a suitable weight for this subset which allows us to retain equidistribution when we restrict any of the above progressions to a shorter interval of suitable length. These equidistribution properties will be required in order to reduce in Section 8 the task of establishing a noncorrelation estimate with nilsequences to the case where the nilsequence is highly equidistributed.

5.1. Smooth numbers in arithmetic progressions. In this subsection we extend [Harper 2012b, Theorem 1] and show that numbers without small and large prime factors are equidistributed in progressions. More

precisely, defining

$$\Psi(x, [y', y]; q, a) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{S([y', y])}(n),$$

we show the following:

Theorem 5.1. *Let $K', K'' > 0$ and $K > \max(2K', 2)$ be constants, let $x \geq 2$ be a parameter and let $1 \leq y' \leq (\log x)^{K'}$ and $(\log x)^K < y \leq x$. If $q < (\log x)^{K''}$, $p \mid q \Rightarrow p < y'$ and $(a, q) = 1$, then*

$$\Psi(x, [y', y]; q, a) = \frac{\Psi(x, [y', y])}{\phi(q)} (1 + O(\log^{-1/5} x)),$$

provided that K and K/K'' are sufficiently large.

Remark. The theorem generalises to the case where $q \leq (\log x)^{K''}$ but $p \mid q \not\Rightarrow p < y'$. In this case, $\Psi(x, [y', y])$ needs to be replaced by $\Psi_q(x, [y', y])$.

As we require explicit error terms, we do not follow the proof strategy from [Harper 2012b], but instead start by establishing a version of the Perron-type bound given in [Harper 2012a, Proposition 1] that applies to

$$\Psi(x, [y', y], \chi) := \sum_{n \leq x} \chi(n) \mathbf{1}_{S([y', y])}.$$

Proposition 5.2. *There exist a small absolute constant $\varrho \in (0, 1)$ and a large absolute constant $C > 1$ such that the following is true. Let $K' > 0$ and $K > \max(2K', 2)$ be constants, let $1 \leq y' \leq (\log x)^{K'}$ and $(\log x)^K < y \leq x$, and suppose that x is large. Suppose that χ is a nonprincipal Dirichlet character with conductor $r \leq x^\varrho$, and to modulus $q \leq x$, such that $L(s, \chi) = \sum_n \chi(n)n^{-s}$ has no zeros in the region*

$$\Re(s) > 1 - \varepsilon, \quad |\Im(s)| \leq H,$$

where $C/\log y < \varepsilon \leq \min(\alpha/2, 1/(2K'))$ and $y^{0.9\varepsilon} \log^2 x \leq H \leq \min(x^\varrho, y^{5/6})$. Suppose, moreover, that at least one of the following holds:

- (i) $y \geq (Hr)^C$.
- (ii) $\varepsilon \geq 40 \log \log(qyH)/\log y$.

Then we have the bound

$$\Psi(x, \chi; [y', y]) \ll \Psi(x, [y', y]) \left((\log x)^{-1/5} + \sqrt{\log x \log y} (H^{-1/2} + x^{-0.4\varepsilon} \log H) \right).$$

Remark 5.3. The $(\log x)^{-1/5}$ term in the bound can be omitted when $y < \exp(\log^{1/4} x)$, i.e., when the proof involves an application of Lemma 4.1 and avoids Lemma 4.2.

The proof of Proposition 5.2 presented below follows the proof of Harper's original result [2012a, Proposition 1] very closely. It involves an application of Perron's formula to relate $\Psi(x, [y', y]; \chi)$ to its Dirichlet L -function, defined as

$$L(s, \chi; [y', y]) = \sum_{n \in S([y', y])} \frac{\chi(n)}{n^s} = \prod_{y' \leq p \leq y} (1 - \chi(p)p^{-s})^{-1} \quad (\Re s > 0),$$

which is followed by a contour shift in the resulting contour integral. The following lemma will be used in order to estimate the integrals on the new contour.

Lemma 5.4. *Suppose that $\alpha = \alpha(x, y)$ is the saddle point associated to $S(x, y)$. Then under the assumptions of the proposition, we have*

$$|\log L(\sigma + it, \chi; [y', y]) - \log L(\alpha + it, \chi; [y', y])| \leq \frac{(\alpha - \sigma) \log x}{2}$$

for all $\alpha - 0.8\varepsilon \leq \sigma \leq \alpha$ and $|t| \leq H/2$.

Before proving this lemma, we complete the proof of the proposition.

Proof of Proposition 5.2 assuming Lemma 5.4. The truncated version of Perron’s formula as given in [Montgomery and Vaughan 2007, Theorems 5.2 and 5.3] yields

$$\begin{aligned} \Psi(x, [y', y]; \chi) &= \frac{1}{2\pi i} \int_{\alpha-iH/2}^{\alpha+iH/2} L(s, \chi; [y', y]) \frac{x^s}{s} ds \\ &\quad + O\left(\frac{x^\alpha}{H} \sum_{n \in \mathcal{S}([y', y])} \frac{\chi_0(n)}{n^\alpha} + \sum_{\substack{x/2 < n < 2x \\ n \in \mathcal{S}([y', y])}} \chi_0(n) \min\left\{1, \frac{x}{H|x-n|}\right\}\right), \end{aligned} \tag{5-1}$$

where $\alpha = \alpha(x, y)$. Since $x/(H|x-n|) \leq H^{-1/2}$ whenever $|x-n| > xH^{-1/2}$, we have

$$\min\left\{1, \frac{x}{H|x-n|}\right\} \leq \begin{cases} 1 & \text{if } |x-n| \leq xH^{-1/2}, \\ H^{-1/2} & \text{if } |x-n| > xH^{-1/2}. \end{cases}$$

Thus, the second error term in (5-1) is bounded above by

$$\sum_{\substack{n \in \mathcal{S}([y', y]) \\ x/2 < n < 2x}} H^{-1/2} \chi_0(n) + \sum_{\substack{|x-n| \leq xH^{-1/2} \\ n \in \mathcal{S}([y', y])}} 1 \ll \frac{x^\alpha}{H^{1/2}} L(\alpha, \chi_0; [y', y]) + |\Psi(x + xH^{-1/2}, [y', y]) - \Psi(x - xH^{-1/2}, [y', y])|.$$

Since the first error term in (5-1) is smaller than the first term in the preceding line, it suffices to estimate the two terms in the bound above in order to bound the error in (5-1).

Concerning the first term, Lemma 3.3, combined with formulas (3-3) and (3-5), shows that

$$\begin{aligned} x^\alpha L(\alpha, \chi_0; [y', y]) &\ll x^\alpha \sum_{n \in \mathcal{S}([y', y])} n^{-\alpha} \ll x^\alpha \zeta(\alpha, y) \prod_{p < y'} (1 - p^{-\alpha}) \\ &\ll \Psi(x, [y', y]) \sqrt{\log x \log y}. \end{aligned} \tag{5-2}$$

Concerning the second term, applying Lemma 4.1 if $y < \exp(\log^{1/4} x)$ and Lemma 4.2 if $\exp(\log^{1/4} x) \leq y \leq x$ shows that

$$\begin{aligned} \Psi(x(1 + H^{-1/2}), [y', y]) - \Psi(x(1 - H^{-1/2}), [y', y]) &\ll ((1 + H^{-1/2})^\alpha - (1 - H^{-1/2})^\alpha + H^{-1/2} + (\log x)^{-1/5}) \Psi(x, [y', y]) \\ &\ll (H^{-1/2} + (\log x)^{-1/5}) \Psi(x, [y', y]), \end{aligned}$$

provided that $H \leq y^{5/6}$ so that $z = H^{1/2} \leq y^{5/12}$ in the application of Lemma 4.2, and where we used an integral to bound the difference. This shows that

$$\Psi(x, [y', y]; \chi) = \frac{1}{2\pi i} \int_{\alpha-iH/2}^{\alpha+iH/2} L(s, \chi; [y', y]) \frac{x^s}{s} ds + O(\Psi(x, [y', y])\{(\log x)^{-1/5} + H^{-1/2}\sqrt{\log x \log y}\}). \quad (5-3)$$

Since $L(s, \chi; [y', y])$ is defined and has no singularities in $\Re s > 0$, shifting the line of integration to $\Re s = \alpha - 0.8\epsilon$ shows that the integral above is equal to

$$\int_{-H/2}^{H/2} L(\alpha - 0.8\epsilon + it, \chi; [y', y]) \frac{x^{\alpha-0.8\epsilon+it}}{\alpha - 0.8\epsilon + it} dt \pm \int_{\alpha-0.8\epsilon}^{\alpha} L\left(\sigma \pm \frac{iH}{2}, \chi; [y', y]\right) \frac{x^{\sigma \pm iH/2}}{\sigma \pm \frac{iH}{2}} d\sigma,$$

where the final expression indicates the sum over the two horizontal pieces. Multiplying each of their integrands by a trivial factor of the form $L(\alpha \pm iH/2, \chi; [y', y])^{+1-1}$ and applying Lemma 5.4, each of the horizontal integrals is seen to be bounded by

$$\frac{L(\alpha, \chi_0; [y', y])}{H} \int_{\alpha-0.8\epsilon}^{\alpha} x^{(\alpha-\sigma)/2} x^\sigma d\sigma \ll \frac{x^\alpha L(\alpha, \chi_0; [y', y])}{H}.$$

A similar argument shows that the vertical integral is bounded by

$$\begin{aligned} L(\alpha, \chi_0; [y', y]) x^{\alpha-0.8\epsilon+0.4\epsilon} \int_{-H/2}^{H/2} \frac{dt}{\alpha + |t|} &\ll x^\alpha L(\alpha, \chi_0; [y', y]) x^{-0.4\epsilon} \left(\frac{1}{\alpha} + \log H\right) \\ &\ll \sqrt{\log x \log y} \Psi(x, [y', y]) x^{-0.4\epsilon} \log H, \end{aligned}$$

where we applied (5-2). Putting everything together,

$$\Psi(x, \chi; [y', y]) \ll \Psi(x, [y', y]) \left((\log x)^{-1/5} + \sqrt{\log x \log y} (H^{-1/2} + x^{-0.4\epsilon} \log H) \right),$$

as required. It remains to prove the lemma. □

Proof of Lemma 5.4. Reinterpreting the given difference as an integral shows that it is bounded above by

$$(\alpha - \sigma) \sup_{\sigma \leq \sigma' \leq \alpha} \left| \frac{L'(\sigma' + it, \chi; [y', y])}{L(\sigma' + it, \chi; [y', y])} \right| = (\alpha - \sigma) \sup_{\sigma \leq \sigma' \leq \alpha} \left| \sum_{n \in \mathcal{S}([y', y])} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right|.$$

We may replace the summation condition in the final sum by $n \in [y', y]$ when bounding the contribution from proper prime powers $p^k \in [y', y]$ with $p < y'$ and $p^k > y$ with $p \leq y$ separately. Bounding this contribution from proper prime powers trivially, the expression above is seen to be

$$\begin{aligned} &\ll (\alpha - \sigma) \sup_{\sigma \leq \sigma' \leq \alpha} \left| \sum_{y' \leq p \leq y} \frac{\log p \chi(n)}{p^{\sigma'+it}} \right| + (\alpha - \sigma) \sum_{p \leq y} \frac{\log p}{p^{2(\alpha-0.8\epsilon)}} \\ &\ll (\alpha - \sigma) \left(\sup_{\alpha-0.8\epsilon \leq \sigma' \leq \alpha} \left| \sum_{n \leq y} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| + \sup_{\alpha-0.8\epsilon \leq \sigma' \leq \alpha} \left| \sum_{n \leq y'} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| + \frac{y^{1-\alpha-0.1\epsilon}}{1-\alpha} + \frac{1}{\epsilon} \right), \end{aligned}$$

as $2(\alpha - 0.8\varepsilon) \geq \alpha + 0.4\varepsilon$ if $\varepsilon \leq \alpha/2$. In the case that condition (i) of Proposition 5.2 holds, that is $y \geq (Hr)^C$, it follows from [Harper 2012a, Lemma 1] that

$$\sup_{\alpha - 0.8\varepsilon \leq \sigma' \leq \alpha} \left| \sum_{n \leq y} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| \ll \frac{y^{1-\alpha-0.1\varepsilon}}{1-\alpha} + \log(rH) + \log^{0.9} q + \frac{1}{\varepsilon}$$

since $\log^2(qyH)/H \ll y^{-0.9\varepsilon}$ whenever $y^{0.9\varepsilon} \log^2 x \leq H$. Otherwise, condition (ii) of Proposition 5.2 holds by assumption, and we have $\log^2(qyH) < y^{2\varepsilon/40} = y^{0.05\varepsilon}$. Thus it follows from [Harper 2012a, Lemma 2] that the previous estimate holds in this case as well.

Concerning the sum over $n \leq y'$, we shall show that, for all sufficiently large x ,

$$\sup_{\alpha - 0.8\varepsilon \leq \sigma' \leq \alpha} \left| \sum_{n \leq y'} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| < \frac{\log x}{4}.$$

To see this, note that if $1 - 1/\log y' < \sigma' \leq 1$, then

$$\left| \sum_{n \leq y'} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| \leq \sum_{n \leq y'} \frac{\Lambda(n)}{n^{\sigma'}} \ll \log y' \ll K' \log_2 x = o(\log x).$$

If $\alpha - 0.8\varepsilon \leq \sigma' < 1 - 1/\log y'$, then partial summation and the prime number theorem imply that

$$\left| \sum_{n \leq y'} \frac{\Lambda(n)\chi(n)}{n^{\sigma'+it}} \right| \ll \frac{y'^{1-\sigma'}}{1-\sigma'} \ll y'^{1-\sigma'} \log y' \ll y'^{1-\alpha} y'^{0.8\varepsilon} \log y' \ll y^{(1-\alpha)/2} (\log x)^{0.4} \log y',$$

where we also used that $\varepsilon \leq 1/(2K')$ and $y' < y^{1/2}$. On recalling the estimate (3-2), that is,

$$1 - \alpha = \frac{\log(u \log(u + 1)) + O(1)}{\log y},$$

and noting that $\log(u + 1) \ll \log_2 x$, the above is seen to be bounded by

$$\ll (u \log(u + 1))^{1/2} (\log x)^{0.4} \log y' \ll \frac{K' (\log x)^{9/10} (\log_2 x)^{3/2}}{(\log y)^{1/2}} = o(\log x),$$

as required.

Collecting everything together, we conclude that

$$\begin{aligned} & \left| \log L(\sigma + it, \chi; [y', y]) - \log L(\alpha + it, \chi; [y', y]) \right| \\ & \ll (\alpha - \sigma) \left(\left(y^{-0.1\varepsilon} + \frac{1}{4} \right) \log x + \log(rH) + \log^{0.9} q + \frac{1}{\varepsilon} \right). \end{aligned}$$

Since $\varepsilon > C/\log y$, $q < x$ and $r, H \leq x^\varrho$, we obtain

$$\begin{aligned} & \left| \log L(\sigma + it, \chi; [y', y]) - \log L(\alpha + it, \chi; [y', y]) \right| \\ & \ll (\alpha - \sigma) \left((e^{-0.1C} + 4^{-1} + 2\varrho + C^{-1}) \log x + \log^{0.9} x \right). \end{aligned}$$

The claimed bound thus follows provided $C > 1$ is sufficiently large, $0 < \varrho < 1$ sufficiently small and x is sufficiently large. □

Proof of Theorem 5.1. Since $(a, q) = 1$, character orthogonality implies that

$$\begin{aligned}\Psi(x, [y', y]; q, a) &= \frac{1}{\phi(q)} \sum_{\chi \pmod{q}} \bar{\chi}(a) \sum_{n \in S(x, [y', y])} \chi(n) \\ &= \frac{1}{\phi(q)} \Psi(x, [y', y]; \chi_0) + \frac{1}{\phi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \bar{\chi}(a) \Psi(x, [y', y]; \chi).\end{aligned}$$

If $p \mid q \Rightarrow p < y'$, then

$$\Psi(x, [y', y], \chi_0) = \sum_{n \leq x, (n, q)=1} 1_{S([y', y])}(n) = \Psi(x, [y', y]).$$

To complete the proof, it thus suffices to show that

$$\Psi(x, [y', y]; \chi) \ll \Psi(x, [y', y]) \log^{-1/5} x$$

uniformly for all nonprincipal characters $\chi \pmod{q}$.

The classical zero-free region (cf. [Montgomery and Vaughan 2007, Theorem 11.3]) implies that there is an absolute constant $0 < \kappa \leq 1$ such that $\prod_{\chi \neq \chi_0} L(\sigma + it, \chi)$ has at most one, necessarily simple, zero in the region

$$\{s : \sigma > 1 - \kappa / \log(qH), |t| < H\}.$$

If such an exceptional zero exists, let χ_{Siegel} denote the corresponding character. We consider the cases $\chi \neq \chi_{\text{Siegel}}$ and $\chi = \chi_{\text{Siegel}}$ separately.

Suppose first that $\chi \neq \chi_{\text{Siegel}}$ and let $r \leq q \leq (\log x)^{K''}$ denote the conductor of χ . We claim that the conditions of Proposition 5.2 are satisfied if we choose

$$H = \min\{y^{\kappa/(2C)}, \exp(\log^{2/3} x), y^{5/6}, x^\varepsilon\} \quad \text{and} \quad \varepsilon = \kappa / \log(qH).$$

Suppose that x is sufficiently large to ensure that $\exp(\log^{2/3} x) < x^\varepsilon$ and suppose that $C > 1 \geq \kappa$. If $K > 2CK''/\kappa$ (or, in other words, if K/K'' is sufficiently large), then

$$q \leq (\log x)^{K''} < (\log x)^{\kappa K/(2C)} < y^{\kappa/(2C)}.$$

We thus have $qH < y^{\kappa/C}$ and, hence, $\varepsilon = \kappa / \log(qH) > C / \log y$. The upper bound $\varepsilon < \min(\alpha/2, 1/(2K''))$ holds as soon as x , and hence H , is sufficiently large. Concerning the conditions on H , suppose first that $H = \exp(\log^{2/3} x)$. In this case, $\varepsilon \asymp \log^{-2/3} x$ and $y^{0.9\varepsilon} \ll \exp(\log^{1/3} x)$. Hence, $H > y^{0.9\varepsilon} \log^2 x$ holds as soon as x is sufficiently large. Next, suppose that $H \neq \exp(\log^{2/3} x)$. In this case, $\min(y^{\kappa/(2C)}, y^{5/6}) \leq \exp(\log^{2/3} x)$ and $\varepsilon \asymp 1/\log y \gg 1/\log^{2/3} x$. From $\varepsilon \asymp 1/\log y$ we obtain $y^{0.9\varepsilon} \ll 1$, and thus $H > y^{0.9\varepsilon} \log^2 x$ holds provided that the exponent K in $y > (\log x)^K$ satisfies $\min(5K/6, K\kappa/(2C)) > 2$ and provided that x is sufficiently large. With the below application of Proposition 5.2 in mind, observe that the lower bound $\varepsilon \gg 1/\log^{2/3} x$, which holds in either of the above two cases, implies that

$$x^{-0.4\varepsilon} \ll \exp(-c \log^{1/3} x) \quad (c > 0).$$

Finally, observe that condition (i) of Proposition 5.2 is satisfied since $rH \leq qH \leq y^{\kappa/C} \leq y^{1/C}$ in view of $\kappa \leq 1$.

This shows that Proposition 5.2 applies to all $\chi \pmod q$, $\chi \neq \chi_{\text{Siegel}}$ and yields

$$\begin{aligned} \Psi(x, [y', y]; \chi) &\ll \Psi(x, [y', y])((\log x)^{-1/5} + \sqrt{\log x \log y}(H^{-1/2} + x^{-0.4\epsilon} \log H)) \\ &\ll \Psi(x, [y', y])(\log x)^{-1/5}. \end{aligned}$$

To see that the latter bound holds, we shall show that the second and third terms in the bound above make a negligible contribution. Indeed, the third term is bounded as follows:

$$\begin{aligned} \sqrt{\log x \log y} x^{-0.4\epsilon} \log H &\ll \sqrt{\log x \log y} \exp(-c \log^{1/3} x)(\log y + \log^{2/3} x) \\ &\ll \exp(-c' \log^{1/3} x), \end{aligned}$$

where $c, c' > 0$ are positive constants. Further, if $H = \exp(\log^{2/3} x)$, then the second term is bounded by

$$\sqrt{\log x \log y} H^{-1/2} \ll \exp(-(\log x)^{2/3}/2 + \log_2 x),$$

which is negligible. If $H \neq \exp(\log^{2/3} x)$, then $\log y \ll \log^{2/3} x$ and $\epsilon \asymp 1/(\log y)$, and the second term satisfies

$$\sqrt{\log x \log y} H^{-1/2} \ll \sqrt{\log x \log y} \frac{y^{-9\epsilon/20}}{\log x} \ll \left(\frac{\log y}{\log x}\right)^{1/2} \ll (\log x)^{-1/3},$$

which is also negligible. This concludes the case of unexceptional characters.

We now consider the contribution of the potential exceptional character $\chi = \chi_{\text{Siegel}}$. Following [Harper 2012a, Section 3.3], we split the analysis into two cases according to the size of y . Suppose first that $y^2 \leq y \leq x^{1/(\log \log x)^2}$. Applying the truncated Perron formula (5-3) with $H = y^{5/6}$, we obtain

$$\begin{aligned} \Psi(x, [y', y]; \chi) &= \frac{1}{2\pi i} \int_{\alpha - iy^{5/6}/2}^{\alpha + iy^{5/6}/2} L(s, \chi; [y', y]) x^s \frac{ds}{s} \\ &\quad + O(\Psi(x, [y', y])\{(\log x)^{-1/5} + y^{-5/12} \sqrt{\log x \log y}\}). \end{aligned} \tag{5-4}$$

As in [Harper 2012a], we proceed by bounding the integrand in absolute value from above. The argument used in Section 3.3 of that work is partly based on ideas from [Soundararajan 2008]. In our case, only small modifications are required. To start with, we have

$$\begin{aligned} \left| \frac{L(\alpha + it, \chi; [y', y])}{L(\alpha, \chi_0; [y', y])} \right| &= \prod_{y' \leq p \leq y} \left| \frac{1 - \chi(p)p^{-\alpha - it}}{1 - \chi_0(p)p^{-\alpha}} \right|^{-1} \leq \prod_{\substack{y' \leq p \leq y \\ p \nmid q}} \left| 1 + \sum_{k \geq 1} \frac{1 - \Re(\chi(p)p^{-it})}{p^{k\alpha}} \right|^{-1} \\ &\leq \exp \left\{ - \sum_{\substack{y' \leq p \leq y \\ p \nmid q}} \frac{1 - \Re(\chi(p)p^{-it})}{p^\alpha} \right\} \leq \exp \left\{ - \sum_{\substack{y^{1/2} \leq p \leq y \\ p \nmid q}} \frac{1 - \Re(\chi(p)p^{-it})}{p^\alpha} \right\}, \end{aligned}$$

since $y'^2 < y$. The next step is to show that in the final expression the argument of the exponential function satisfies

$$\sum_{\substack{y^{1/2} \leq p \leq y \\ p \nmid q}} \frac{1 - \Re \chi(p)p^{-it}}{p^\alpha} \gg \frac{u}{\log^2(u+1)},$$

provided that $|t| \leq y/2$ and $y \leq x^{1/(\log \log x)^2}$. The proof of this bound proceeds by splitting the range of t into $|t| \leq 1/(2 \log y)$ and $1/(2 \log y) \leq |t| \leq y/2$. On the latter range, the proof of the second part of [Soundararajan 2008, Lemma 5.2] is employed. In that proof, all sums over primes can be restricted to the range $y^{1/2} \leq p \leq y$, $p \nmid q$. On the former range, Harper’s argument [2012a, p. 17] applies directly. For sake of completeness, we summarise this argument here. Since $p \leq y$ and $t \leq 1/(2 \log y)$, we have $|t \log p| \leq \frac{1}{2}$ and, by analysing the cases $\chi(p) = 1$ and $\chi(p) = -1$ for the quadratic character $\chi = \chi_{\text{Siegel}}$ separately, one obtains

$$\begin{aligned} \sum_{y^{1/2} \leq p \leq y, p \nmid q} \frac{1 - \Re \chi(p) p^{-it}}{p^\alpha} &\gg \sum_{y^{1/2} \leq p \leq y, p \nmid q} \frac{1 - \chi(p)}{p^\alpha} \\ &\geq \frac{1}{\log y} \left(\sum_{y^{1/2} \leq p \leq y, p \nmid q} \frac{\log p}{p^\alpha} - \sum_{y^{1/2} \leq p \leq y, p \nmid q} \frac{\chi_{\text{Siegel}}^*(p) \log p}{p^\alpha} \right), \end{aligned}$$

where χ_{Siegel}^* is the primitive character that induces χ_{Siegel} . Harper’s argument is based on a combination of partial summation and the asymptotic evaluation of $\sum_{n \leq x} \Lambda(n) \chi(n)$ together with the observation that a Siegel zero has a negative contribution in that asymptotic evaluation. The error terms in the asymptotic evaluation in [Montgomery and Vaughan 2007, Theorem 11.16 and Exercise 11.3.1.2] are small when $q < (\log x)^K$. Thanks to the minus sign in the final expression above, the contribution from the Siegel zero to this expression is positive and can therefore be ignored when seeking a lower bound. See [Harper 2012a, Section 3.3] for the remaining details.

Returning to (5-4), the above bounds and an application of (5-2) yield

$$\begin{aligned} \Psi(x, [y', y]; \chi) &\ll L(\alpha, \chi_0; [y', y]) e^{-c_3 u (\log u)^{-2}} x^\alpha \log y + \frac{\Psi(x, [y', y])}{\log^{1/5} x} \\ &\ll \Psi(x, [y', y]) (\sqrt{\log x} (\log y)^{3/2} e^{-c_3 u (\log u)^{-2}} + \log^{-1/5} x). \end{aligned}$$

Since $u = (\log x)/\log y \gg \log \log^2 x$ when $y \leq \exp((\log x)/\log \log^2 x)$, the first term in the bound is $O_A(\Psi(x, [y', y]) (\log x)^{-A})$ and we thus have

$$\Psi(x, [y', y]; \chi) \ll (\log x)^{-1/5} \Psi(x, [y', y])$$

if $\chi = \chi_{\text{Siegel}}$ and $y \leq x^{1/(\log \log x)^2}$.

It remains to analyse the range $x^{1/(\log \log x)^2} \leq y \leq x$ of y . In this case too, we follow the overall strategy used in [Harper 2012a] but need to make a number of technical changes arising from our slightly different situation. Observe that

$$\Psi(x, [y', y]; \chi) = \sum_{n \in S(x, [y', y])} \chi(n) = \sum_{\substack{n \in S(x, y) \\ (n, P(y'))=1}} \chi(n) \leq \sum_{\substack{d \leq x \\ d | P(y')}} \left| \sum_{n \in S(x/d, y)} \chi(n) \right|,$$

where we applied Möbius inversion in order to remove the coprimality condition. Our next aim is to show that, at the expense of an acceptable error term, the sum over d can be further truncated in such a way that $\log(x/d) \asymp \log x$ uniformly for all values of d that remain in the sum. Recall that $\alpha(x, y) > 1 - 1/K + o(1)$

and that $\Psi(x, y') \leq x^{1-1/K'+o(1)}$. Let $K^* > K$ be sufficiently large so that $\alpha(x, y)(1 - 1/K^*) > 1 - 1/K'$. Then the $m = 1$ case of [de la Bretèche and Tenenbaum 2005, Theorem 2.4(i)], which applies uniformly for $1 \leq d \leq x$, implies that

$$\begin{aligned} \sum_{\substack{x^{1-1/K^*} \leq d \leq x \\ d \mid P(y')}} \Psi(x/d, y) &\ll \Psi(x, y)x^{-(1-1/K^*)\alpha} \sum_{x^{1-1/K^*} \leq d \leq x} \mathbf{1}_{S(y')}(d) \\ &\ll \Psi(x, y)x^{-(1-1/K^*)\alpha} x^{1-1/K'+o(1)} \ll \Psi(x, y)x^{-c_4} \end{aligned} \tag{5-5}$$

for some positive constant $c_4 > 0$. Assuming for the moment that this bound is sufficiently good, it remains to analyse the expression

$$\sum_{\substack{d \leq x^{1-1/K^*} \\ d \mid P(y')}} \left| \sum_{n \in S(x/d, y)} \chi(n) \right|,$$

where $\log(x/d) \asymp \log x$. Fouvry and Tenenbaum [1996, Lemme 2.1] (see also their Lemme 2.2 and its deduction) studied bounds on character sums such as the inner sum above in a larger range for y than the one under current consideration. In view of the truncation of the sum over d and the bound $q \leq (\log x)^{K''}$ on the modulus of χ , [loc. cit., Lemme 2.1(i)] applies uniformly to the above character sum for each d . By invoking [de la Bretèche and Tenenbaum 2005, Theorem 2.4(i)] in the case $m = 1$, we obtain

$$\begin{aligned} \sum_{\substack{d \leq x^{1-1/K^*} \\ d \mid P(y')}} \left| \sum_{n \in S(x/d, y)} \chi(n) \right| &\ll \sum_{\substack{d \leq x^{1-1/K^*} \\ d \mid P(y')}} d^{-\alpha} \Psi(x, y) \exp(-c\sqrt{\log y}) \\ &\ll \exp(-c\sqrt{\log y}) \Psi(x, y) \prod_{p \leq y'} (1 - p^{-\alpha})^{-1}. \end{aligned} \tag{5-6}$$

It remains to express the bounds (5-6) and (5-5) in terms of $\Psi(x, [y', y])$. By Lemma 3.3 and the first part of Lemma 3.2, we have

$$\Psi(x, y) \ll \Psi(x, [y', y]) \prod_{p \leq y'} (1 - p^{-\alpha})^{-1}.$$

Since $\log y' \asymp \log_2 x$ and $1 - \alpha \ll (\log_2 x)^{2+\varepsilon} / \log x$ for $x^{1/(\log \log x)^2} \leq y \leq x$, it follows from the first part of Lemma 3.2 (or direct computation) that $\prod_{p \leq y'} (1 - p^{-\alpha})^{-1} \ll \log y'$. Hence, it follows from (5-6) and (5-5) that

$$\Psi(x, [y', y]; \chi) \ll_A (\log x)^{-A} \Psi(x, [y', y])$$

if $\chi = \chi_{\text{Siegel}}$ and $x^{1/(\log \log x)^2} \leq y \leq x$. Taking $A = \frac{1}{5}$ completes the proof. □

5.2. *W-trick and equidistribution of weighted smooth numbers in short APs.* The choice of our weight factor $n^{1-\alpha}$ ensures that the weighted version $h_{[y', y]}(n)$ of $\mathbf{1}_{S([y', y])}$ is equidistributed in short intervals and, by invoking Theorem 5.1, also in short progressions $n \equiv a \pmod q$, provided the residue class a is coprime to the modulus q of the progression and $q < y'$. Establishing noncorrelation with nilsequences later requires equidistribution in almost all residue classes $a \pmod q$. Given any $Q \in \mathbb{N}$ and $1 \leq A < Q$,

$\gcd(Q, A) = 1$, we consider the following renormalised restrictions:

$$h_{[y', y]}^{(Q, A)}(m) = \frac{\phi(Q)}{Q} h_{[y', y]}(Qm + A) \quad (m \sim (N - A)/Q), \quad (5-7)$$

$$g_{[y', y]}^{(Q, A)}(m) = \frac{\phi(Q)}{Q} g_{[y', y]}(Qm + A) \quad (m \in \mathbb{N}). \quad (5-8)$$

Let $w(N) = \frac{1}{2} \log_3 N$, let $W = W(N) = \prod_{p \leq w(N)} p$ and observe for later use that $w(N) = (\log_3 N')/2 + o(1)$ for any $N' \in (N/\log N, N]$. We will mainly be interested in the case where

- $Q = W(N)$, or
- $Q = \tilde{W}$ for some $\tilde{W} \leq (\log N)^{K''}$ such that $W(N) \mid \tilde{W}$ and $p \mid \tilde{W} \Rightarrow p \leq w(N)$

in (5-7) and (5-8).

The following lemma is an immediate consequence of Theorem 5.1 and Lemma 4.3.

Lemma 5.5 (equidistribution in short progression). *Let $K', K'' > 0$ and $K > \max(2K', 2)$ be constants and let $N \geq 2$ be a parameter. Suppose that $y' \leq (\log N)^{K'}$, and that $(\log N)^K < y \leq N$. Then, provided that K and K/K'' are sufficiently large, the following estimate holds uniformly for all N_0, N_1 such that*

$$N < N_0 < N_0 + N_1 \leq 2N$$

and $N_1 \gg N \exp(-(\log^{1/4} N)/4)$, for all $q \leq (\log N)^{K''}$ such that $p \mid q \Rightarrow p < y'$ and all $a \pmod{q}$ such that $\gcd(a, q) = 1$:

$$\sum_{\substack{N_0 < m \leq N_0 + N_1 \\ m \equiv a \pmod{q}}} h_{[y', y]}(m) - \frac{N_1}{\phi(q)} \ll \frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q)} \frac{1}{\log^{1/24} N} + \frac{N}{\log^{1/5} N}.$$

Proof. Recall the definition (4-7) of $h_{[y', y]}(n)$. On omitting the weight factor $N^\alpha/(\alpha\Psi(N, [y', y]))$ for the moment, partial summation yields

$$\begin{aligned} & \sum_{N_0 < n \leq N_0 + N_1} n^{1-\alpha} \mathbf{1}_{n \equiv a \pmod{q}} \mathbf{1}_{S([y', y])}(n) \\ &= (N_0 + N_1)^{1-\alpha} \Psi(N_0 + N_1, [y', y]; q, a) - (N_0)^{1-\alpha} \Psi(N_0, [y', y]; q, a) - (1-\alpha) \int_{N_0}^{N_0 + N_1} \frac{\Psi(t, [y', y]; q, a)}{t^\alpha} dt. \end{aligned}$$

Theorem 5.1 implies

$$\Psi(t, [y', y]; q, a) = \frac{\Psi(t, [y', y])}{\phi(q)} (1 + O(\log^{-1/5} N)) \quad (5-9)$$

whenever $N \leq t \leq 2N$ and provided that K is sufficiently large. Inserting this expansion into the previous expression and recalling the partial summation application (4-8), we see that the contribution from the main term in (5-9) can be reinterpreted as a sum over $n^{1-\alpha} \mathbf{1}_{S([y', y])}$ that is not restricted to a progression. More precisely,

$$\phi(q) \sum_{N_0 < n \leq N_0 + N_1} n^{1-\alpha} \mathbf{1}_{n \equiv a \pmod{q}} \mathbf{1}_{S([y', y])}(n) = \sum_{N_0 < n \leq N_0 + N_1} n^{1-\alpha} \mathbf{1}_{S([y', y])}(n) + O\left(\frac{N^{1-\alpha} \Psi(N, [y', y])}{\log^{1/5} N}\right). \quad (5-10)$$

The shape of the error term above follows by several applications of Lemmas 3.4(i) and 3.3. On reinserting the weights into (5-10) and invoking Lemma 4.3, we obtain

$$\begin{aligned} \sum_{\substack{N_0 \leq m \leq N_0 + N_1 \\ m \equiv a \pmod{q}}} h_{[y', y]}(m) &= \frac{1}{\phi(q)} \sum_{N_0 \leq n \leq N_0 + N_1} h_{[y', y]}(n) + O\left(\frac{N}{\log^{1/5} N}\right) \\ &= \frac{N_1}{\phi(q)} \left(1 + O\left(\frac{\log_3 N}{\log_2 N}\right)\right) + O\left(\frac{N}{\phi(q)(\log N)^{1/24}}\right) + O\left(\frac{N}{\log^{1/5} N}\right) \end{aligned}$$

provided that $N_1 \gg N \exp(-(\log N)^{1/4}/4)$. □

By analysing the asymptotic order of the weight factor in $h_{[y', y]}$, we may deduce from the previous result an upper bound on the unweighted count of $[y', y]$ -smooth numbers in short intervals. This bound will be integral in the application of an bootstrapping argument in Section 7.

Corollary 5.6 ($[y', y]$ -smooth numbers in short intervals). *Let $K', K'' > 0$ and $K > \max(2K', 2)$ be constants and let $N \geq 2$ be a parameter. Suppose that $1 \leq y' \leq (\log N)^{K'}$ and that $(\log N)^K < y \leq N$. Let $P \subseteq [N, 2N]$ be any progression of length $|P| \geq N \exp(-(\log^{1/4} N)/4)$ and common difference $q \leq (\log N)^{K''}$ such that $p \mid q \Rightarrow p < y'$.*

Then the following assertions hold provided that K and K/K'' are sufficiently large.

(1) *Suppose that $|P| \geq L := N/(\log N)^\ell$ for some constant $\ell > 0$. Then*

$$|S(2N, [y', y]) \cap P| \ll \Delta \frac{\Psi(2N, [y', y]) \cdot |P|}{N},$$

where $\Delta = \log_3 N + (\log N)^{\ell-1/24}$.

(2) *If, instead, $|P| \geq L := N \exp(-C\varpi(N))$ and $q \leq N/L \leq \exp(C\varpi(N)) = o(\log N)$ for some $\varpi(N) = o(\log_2 N)$, then*

$$|S(2N, [y', y]) \cap P| \ll \Delta \frac{\Psi(2N, [y', y]) \cdot |P|}{N},$$

where $\Delta = \log \varpi(N)$.

Proof. Note that if $n \in [N, 2N]$ is an element of $S([y', y])$, then $h_{[y', y]}(n) \neq 0$ and $N^{1-\alpha} \leq n^{1-\alpha} \leq (2N)^{1-\alpha}$, so that

$$h_{[y', y]}(n) = \frac{N^\alpha}{\Psi(N, [y', y])} \frac{n^{1-\alpha}}{\alpha} \asymp \frac{N}{\Psi(N, [y', y])} \quad (n \in [N, 2N] \cap S([y', y])). \tag{5-11}$$

We shall show that in either of the two cases in the statement

$$\frac{q}{\phi(q)} \ll \Delta,$$

which then leaves us with the task of deducing from Lemma 5.5 that

$$\sum_{n \in P} h_{[y', y]}(n) \ll \Delta |P|.$$

In order to bound $q/\phi(q)$, consider the decomposition $q = q_1 q_2$, where q_1 is composed of primes $p \leq \varpi$ and q_2 is composed of primes $p > \varpi$. Then

$$\prod_{p|q_1} (1 - p^{-1})^{-1} \ll \log \varpi.$$

Concerning the product over prime divisors of q_2 , we obtain

$$\prod_{p|q, p > \varpi} (1 - p^{-1})^{-1} \ll \exp\left(\sum_{p|q, p > \varpi} p^{-1}\right) \ll 1 + O\left(\frac{\log q}{\varpi \log \varpi}\right), \tag{5-12}$$

where we used the bound $\omega(q_2) \leq (\log q) / \log \varpi$. Thus,

$$\frac{q}{\phi(q)} \ll \log \varpi + \frac{\log q}{\varpi} \ll \log_2 q,$$

where we chose $\varpi \log \varpi = \log q$. In case (2), we use this bound directly and obtain $q/\phi(q) \ll \log \varpi(N) = \Delta$. In case (1), we have $q/\phi(q) \ll \log_2 q \ll \log_3 N \ll \Delta$.

Turning towards the application of Lemma 5.5, note that for $|P| = N_1/q$ the bound in that lemma satisfies

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \left(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + (N/|P|)(\log N)^{-1/24} \right).$$

If $|P| > N(\log N)^{-\ell}$, we obtain

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \left(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + (\log N)^{\ell-1/24} \right) \ll |P| \Delta,$$

provided $\Delta \gg (\log N)^{\ell-1/24}$, which holds under the assumptions of (1).

If, instead, $\log(N/|P|) \ll \varpi(N) = o(\log_2 N)$, then

$$\frac{N_1}{\phi(q)} \frac{\log_3 N}{\log_2 N} + \frac{N}{\phi(q) \log^{1/24} N} + \frac{N}{\log^{1/5} N} \ll |P| \left(\frac{q}{\phi(q)} \frac{\log_3 N}{\log_2 N} + o(1) \right) = o(|P|),$$

which concludes the proof. □

In addition to the above result, we shall need an analogous bound that is valid on much shorter intervals but may in return have a larger value of Δ . For y -smooth numbers, such a result appears as ‘‘Smooth Numbers Result 3’’ in [Harper 2016]; see also [Drappeau and Shao 2016, Lemma 3.3] for a slight extension. The lemma below will be used in Section 6, where we extend some of the results of [Drappeau and Shao 2016], as well as in Section 7, where it will find application in those situations where the progressions are too short for the previous corollary to be used.

Lemma 5.7. *Let $K' > 0$, $K \geq \max(2K', 2)$, $x > 2$, $1 \leq y' \leq (\log x)^{K'}$ and $(\log x)^K < y \leq x$. Suppose that $P \subseteq [x, 2x]$ is an arithmetic progression. Then, provided that x is sufficiently large, we have*

$$\#\{n \in P : n \in S([y', y])\} \ll (x/|P|)^{1-\alpha} \frac{\Psi(x, [y', y])|P|}{x} \log x.$$

Proof. Let $X \geq x$ denote the smallest element of P . This lemma is a generalisation of [Harper 2016, Smooth Numbers Result 3], which bounds

$$\sum_{\substack{X \leq n \leq X+Z \\ n \equiv a \pmod{q}}} \mathbf{1}_{S(y)}(n)$$

under the assumptions that $q \geq 1$ and $qy \leq Z \leq X$, i.e., concerns progressions of length $|P| \geq y$. The proof of [Harper 2016, Smooth Numbers Result 3] carries over directly to the situation where $S(y)$ is replaced by $S([y', y])$, with one small exception: The application of [de la Bretèche and Tenenbaum 2005, Théorème 2.4(i)] with $m = 1$ needs to be replaced by an application of Lemma 3.4(i), which is [loc. cit., Théorème 2.4(i)] with $m = P(y')$. While in the former case, the bound holds uniformly for $1 \leq d \leq x$, it is only available for $1 \leq d \leq x/y$ in our case.

Replacing the assumption that $qy \leq Z$ by $qy^2 \leq 2Z$ ensures that $d < X/y$ in all applications of [loc. cit., Théorème 2.4(i)]. More precisely, we have

$$\frac{Xqy}{Z2^{j+1}} \leq \frac{X}{y}$$

for all admissible values of j in Harper’s proof.

It remains to consider the case where $|P| < y^2/2$, which corresponds to the case $|P| < y$ addressed in (the proof of) [Drappeau and Shao 2016, Lemma 3.3]. By bounding the left-hand side in our statement trivially by $|P|$, and observing that $|P|^{-1+\alpha} \gg y^{(-1+\alpha)/2}$, it suffices to show that

$$y^{\alpha-1} \frac{\Psi(x, [y', y])}{x^\alpha} \log x \gg 1.$$

Note that $y^{\alpha-1} = (u \log(u + 1) + O(1))^{-1}$ in view of (3-2) and observe that

$$y^{\alpha-1} \frac{\Psi(x, [y', y])}{x^\alpha} \log x \gg y^{\alpha-1} \frac{\prod_{y' \leq p \leq y} (1 - p^{-\alpha})^{-1}}{\sqrt{\log x \log y}} \log x \gg y^{\alpha-1} \prod_{y' \leq p \leq y} (1 - p^{-\alpha})^{-1}$$

by Lemma 3.3 and (3-3). We split the analysis of the product over primes into two cases according to the size of y . When $\exp(\log x / \log_2 x) \leq y \leq x$, noting that $\alpha \leq 1$, we have

$$\prod_{y' \leq p \leq y} (1 - p^{-\alpha})^{-1} \gg \frac{\prod_{p \leq y} (1 - p^{-1})^{-1}}{\prod_{p \leq y'} (1 - p^{-1})^{-1}} \gg \frac{\log y}{1 + \log y'} \gg \frac{\log x}{(\log_2 x)^2}.$$

Hence,

$$y^{\alpha-1} \prod_{y' \leq p \leq y} (1 - p^{-\alpha})^{-1} \gg \frac{\log y}{\log x \log(u + 1)} \frac{\log x}{(\log_2 x)^2} \gg \frac{\log x}{(\log_2 x)^3 \log_3 x} \gg 1.$$

When $y^{1/2} \leq y \leq \exp(\log x / \log_2 x)$, we have $u \geq \log_2 x$ and $\alpha \leq 1 - (\log_3 x) / \log y + O(1 / \log y)$. Hence, the second conclusion of Lemma 3.2 shows that

$$\begin{aligned} y^{-1+\alpha} \prod_{y' \leq p \leq y} (1 - p^{-\alpha})^{-1} &\gg y^{-1+\alpha} \prod_{\sqrt{y} \leq p \leq y} (1 - p^{-\alpha})^{-1} \\ &\gg y^{-1+\alpha} \exp\left(\frac{y^{1-\alpha}(1 + o(1)) - 2y^{(1-\alpha)/2}(1 + o(1))}{(1 - \alpha) \log y}\right) \\ &\gg \exp\left(\frac{y^{1-\alpha}(1 + o(1)) - (\log y^{1-\alpha})^2}{\log y^{1-\alpha}}\right) \gg 1, \end{aligned}$$

where we used that $y^{1-\alpha} \asymp u \log u$, where $u \rightarrow \infty$ as $x \rightarrow \infty$, implies that $y^{(1-\alpha)/2} = o(y^{1-\alpha})$. □

6. Weyl sums for numbers without small and large prime factors

As technical input in the proof of the orthogonality of $(h_{[y',y]}^{(W,A)}(n) - 1)$ to nilsequences, we shall require bounds on Weyl sums where the summation variable is restricted to integers in $S([y', y])$. Drappeau and Shao [2016, Section 5] established bounds on Weyl sums for smooth numbers $S(x, y)$ by extending the methods from Harper’s minor arc analysis [2016]. Our aim in this section is to obtain analogous bounds for Weyl sums over $S(x, [y', y])$, that is, for Weyl sums over numbers free from small and large prime factors.

Following the notation in [Drappeau and Shao 2016], define for any given parameters $Q, x > 0$ and coprime integers $0 < a \leq q \leq Q$ the major arcs

$$\mathfrak{M}(q, a; Q, x) = \{\theta \in [0, 1) : |q\theta - a| \leq Qx^{-k}\}$$

and

$$\mathfrak{M}(Q, x) = \bigcup_{\substack{0 \leq a < q \leq Q \\ (a,q)=1}} \mathfrak{M}(q, a; Q, x). \tag{6-1}$$

Given any positive integer k , define the following smooth Weyl sum of degree k :

$$E_k(x, [y', y]; \theta) := \sum_{n \in S(x, [y', y])} e(\theta n^k).$$

The main objective of this section is to establish the theorem below.

Theorem 6.1 (Weyl sums, x^η -smooth case). *Let $k > 0$ be a fixed positive integer and suppose that $\eta \in (0, 1/(4k)]$. Let $K > 2K' > 2$ and suppose that $y' \leq (\log x)^{K'} < (\log x)^K < y \leq x^\eta$ and that K is sufficiently large. Let $\alpha = \alpha(x, y)$ denote the saddle point associated to $S(x, y)$, and let $\theta \in \mathbb{T} := \mathbb{R}/\mathbb{Z}$ be a frequency. Then:*

(1) *If x is sufficiently large and $\theta \notin \mathfrak{M}(x^{1/12}, x)$, we have*

$$E_k(x, [y', y]; \theta) \ll x^{1-c}$$

for some $c = c(k, \varepsilon) > 0$.

(2) *If x is sufficiently large, $\theta \in \mathfrak{M}(x^{1/12}, x)$ and if $0 < a < q \leq x^{0.1}$ are coprime integers such that $|q\theta - a| \leq q^{-1}$, then*

$$E_k(x, [y', y]; \theta) \ll \mathcal{Q}^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y])$$

for some $c = c(k, \varepsilon) > 0$ and $\mathcal{Q} = q + x^k \|q\theta\|$.

We prove the two parts of this lemma in turn. The former case, in which θ is “highly irrational”, will be handled by extending results of [Wooley 1995]. For this purpose, we first establish an auxiliary lemma generalising [loc. cit., Lemma 2.3], which in turn builds on an argument of [Vaughan 1989].

Lemma 6.2. *Suppose that $\theta \in \mathbb{T}$ is a frequency and $r \in \mathbb{N}$. If $1 \leq y' < y \leq M < x$, then we have*

$$\sum_{\substack{n \in S(x, [y', y]) \\ (n,r)=1}} e(n^k \theta) \ll y(\log x) \max_{y' \leq p \leq y} \sum_{\substack{v \in \mathfrak{B}(M, p, y) \\ (v,r)=1}} \sup_{\beta \in \mathbb{T}} \left| \sum_{\substack{u \in S(x/M, [y', p]) \\ (u,r)=1}} e(u^k v^k \theta + u\beta) \right| + M,$$

where $\mathfrak{B}(M, p, y) = \{M < v \leq Mp : p \mid v, v \in S([p, y])\}$.

Proof. We start by decomposing the elements $n \in S(x, [y', y])$ in a similar fashion as in [Vaughan 1989, Lemma 10.1]. By considering the prime factorisation $n = p_1^{k_1} \cdots p_s^{k_s}$, where the prime factors are ordered in *decreasing* order $y \geq p_1 > \cdots > p_s \geq y'$, it is immediate that, for every given $M \in [y, x)$ and every $n \in S(x, [y', y])$ with $n > M$, there is a unique triple (u, v, p) such that $n = uv$ and such that v is the smallest initial factor in the factorisation above that exceeds M . More precisely,

- (1) $y' \leq p \leq y$ and $p \mid v$,
- (2) $u \in S(x/v, [y', p])$,
- (3) $M < v \leq Mp$ and $v \in S([p, y])$.

Using this factorisation, the exponential sum can be decomposed as

$$\sum_{\substack{n \in S(x, [y', y]) \\ (n, r)=1}} e(n^k \theta) = \sum_{y' \leq p \leq y} \sum_{\substack{v \in \mathfrak{B}(M, p, y) \\ (v, r)=1}} \sum_{\substack{u \in S(x/v, [y', p]) \\ (u, r)=1}} e(u^k v^k \theta) + O(M),$$

where the error term $O(M)$ bounds the contribution from all $n \in S(x, [y', y])$ with $n \leq M$. As in [Vaughan 1989, (10.9)], we can use the orthogonality principle to remove the dependence on v in the restriction $u \in S(x/v, [y', p])$. Since $v > M$, the inner sum above is equal to

$$\begin{aligned} \int_{\mathbb{T}} \left(\sum_{\substack{u \in S(x/M, [y', p]) \\ (u, r)=1}} e(u^k v^k \theta + u\beta) \right) \left(\sum_{m \leq x/v} e(-m\beta) \right) d\beta \\ \ll \sup_{\beta \in \mathbb{T}} \left| \sum_{\substack{u \in S(x/M, [y', p]) \\ (u, r)=1}} e(u^k v^k \theta + u\beta) \right| \int_{\mathbb{T}} \min\{x/v, \|\beta\|^{-1}\} d\beta \\ \ll \log x \sup_{\beta \in \mathbb{T}} \left| \sum_{\substack{u \in S(x/M, [y', p]) \\ (u, r)=1}} e(u^k v^k \theta + u\beta) \right|. \end{aligned}$$

The lemma follows by combining this bound with the previous expression. □

The following lemma is a generalisation of the Weyl sum estimate for y -smooth numbers given by [Wooley 1995], and it relies on the observation that the conclusion of [loc. cit., Lemma 3.1] continues to hold when we restrict the y -smooth numbers to $[y', y]$ -smooth numbers.

Lemma 6.3 (Weyl sum, θ is highly irrational). *Let $k \in \mathbb{N}$ be a fixed positive integer and suppose that $\sigma \in (0, \frac{1}{2})$ and $\eta \in (0, \sigma/(4k)]$. Then there exists $c = c(k, \sigma) > 0$ such that the following statement is true. Let $1 \leq y' < y \leq x^\eta$ and $\theta \in (0, 1)$. Then, if $\theta \notin \mathfrak{M}(x^\sigma, x)$, we have*

$$E_k(x, [y', y]; \theta) \ll x^{1-c}.$$

Proof. We seek to apply a modified version of [Wooley 1995, Lemma 3.1] in a similar way as at the start of the proof of [loc. cit., Theorem 4.2]. Wooley’s lemma produces bounds on

$$f(\theta; x, y) := E_k(x, [1, y]; \theta)$$

in terms of M and q , where $M = x^\lambda > y$ is a parameter such that $\lambda \in (\frac{1}{2}, 1)$ and q is any natural number such that $\|q\theta\| = |q\theta - a| < q^{-1}$ for some $a \in \mathbb{Z}$ with $(a, q) = 1$. As we shall explain below, one can deduce from the proof of the lemma that the same bounds hold for $E_k(x, [y', y]; \theta)$ in place of $f(\theta; x, y) = E_k(x, [1, y]; \theta)$, i.e., for the exponential sum relevant to us. In order for those bounds to be useful, additional assumptions on M and q are necessary, and our first aim is to show that, under the assumption of our lemma, we can choose M and q in such a way that

$$M > y, \quad q \leq 2(yM)^k, \quad |q\theta - a| < (yM)^{-k}/2 \quad \text{and} \quad q > (x/M)^k. \quad (6-2)$$

These conditions correspond to those in place at the start of the proof of [Wooley 1995, Theorem 4.2]. When combined⁶ with the upper bound produced by [loc. cit., Lemma 3.1] it follows that

$$E_k(x, [y', y]; \theta) \ll x^{1-c(k,\sigma)}, \quad (6-3)$$

once we prove that the lemma can be extended to the case of $[y', y]$ -smooth numbers.

Concerning the conditions (6-2), let $0 < \sigma^* < \sigma/2$. Then $\sigma^* + 2k\eta < \sigma < \frac{1}{2}$. By the Dirichlet approximation theorem, there is a positive integer $q \leq x^{k-\sigma^*}$ and some $a \in \mathbb{Z}$ with $(a, q) = 1$ such that

$$\|q\theta\| = |q\theta - a| \leq x^{\sigma^*-k} < x^{\sigma-k}.$$

Since $\theta \notin \mathfrak{M}(x^\sigma, x)$, this implies that $q > x^\sigma$. Let M be defined by the equation

$$x^{k-\sigma^*} = 2(yM)^k.$$

Then, since $\sigma^* + 2\eta < \frac{1}{2}$, we have

$$y < x^\eta < x^{1-\sigma^*/k-\eta} 2^{-1/k} \leq M = 2^{-1/k} x^{1-\sigma^*/k} y^{-1} \leq x^{1-\sigma^*/k}$$

as soon as x is sufficiently large. Hence, $y < M$. If we write $M = x^\lambda$, then the above line of inequalities also shows that $\lambda \in (\frac{1}{2}, 1)$ since $\sigma^*/k + \eta < \frac{1}{2}$. Observe further that $\sigma^* + k\eta < \sigma$ implies that

$$q > x^\sigma > 2x^{\sigma^*+k\eta} \geq (x/M)^k$$

as soon as x is sufficiently large. Hence, all conditions from (6-2) are satisfied and (6-3) therefore follows under the assumptions of our lemma.

It remains to show that [Wooley 1995, Lemma 3.1] can be extended to the case of exponential sums over $[y', y]$ -smooth numbers. The proof strategy is to follow the original argument, applied to the exponential sum $E_k(x, [y', y]; \theta)$ instead of $E_k(x, [1, y]; \theta)$, and explain how all dependencies on y' that appear in any of the bounds on the way can be removed, so as to eventually arrive at an intermediate bound that agrees with the corresponding bound in the original proof. Following the original proof through to the end shows then that the upper bound that [loc. cit., Lemma 3.1] provides for $f(\theta; x, y) = E_k(x, [1, y]; \theta)$ is also an upper bound for $E_k(x, [y', y]; \theta)$.

⁶See the proof of [loc. cit., Theorem 4.2] for details.

Turning to the details, let $M = x^\lambda$ for any $\lambda \in (\frac{1}{2}, 1)$ and let $q \in \mathbb{N}$ be such that $|q\theta - a| < q^{-1}$ for some $a \in \mathbb{Z}$ with $(a, q) = 1$. The $r = 1$ case of Lemma 6.2 (which replaces [Wooley 1995, Lemma 2.3]) produces a prime $y' < p \leq y$ and some frequency $\gamma \in \mathbb{T}$ such that

$$E_k(x, [y', y]; \theta) \ll (\log x)y \sum_{v \in S(My, y)} \left| \sum_{u \in S(x/M, [y', p])} e(u^k v^k \theta + u\gamma) \right| + M.$$

Note carefully that the sum over v only involves y -smooth numbers. Following the argumentation of the start of the proof of [loc. cit., Lemma 3.1] (which involves expressing the absolute value of the inner sum as $\varepsilon(v, \theta) \sum_{u \in S(x/M, [y', p])} e(u^k v^k \theta + u\gamma)$ for some $\varepsilon(v, \theta) \in \mathbb{C}$ of unit modulus, reinterpreting certain exponential sums, as well as two applications of Hölder’s inequality) shows that, for all positive integers $t, w \in \mathbb{N}$,

$$\left(\sum_{v \in S(My, y)} \left| \sum_{u \in S(x/M, [y', p])} ye(u^k v^k \theta + u\gamma) \right| \right)^{2tw} \ll (My)^{2w(t-1)} \left(\sum_c n_c \right)^{2w-2} \left(\sum_c n_c^2 \right) J_w(\theta),$$

where $J_w(\theta)$ is as in [loc. cit., (3.4)] and where n_c denotes the number of solutions to the equation $u_1^k + \dots + u_t^k = c$ with $u_i \in S(x/M, [y', p])$. Since $p \leq y$, we have $S(x/M, [y', p]) \subset S(x/M, y)$, which implies that $\sum_c n_c^2$ is trivially bounded above by $\int_{\mathbb{T}} \left| \sum_{u \in S(x/M, y)} e(u^k \theta) \right|^{2t} d\theta$. Using the trivial bound $\sum_c n_c \ll (x/M)^t$ for the remaining sum over c completes our task of removing all dependencies on y' from the upper bound on the given exponential sum $E_k(x, [y', y]; \theta)$. To summarise, we obtain

$$\left(\sum_{v \in S(My, y)} \left| \sum_{u \in S(x/M, [y', p])} e(u^k v^k \theta + u\gamma) \right| \right)^{2tw} \ll (My)^{2w(t-1)} (x/M)^{t(2w-2)} J_w(\theta) \int_{\mathbb{T}} \left| \sum_{u \in S(x/M, y)} e(u^k \theta) \right|^{2t} d\theta,$$

corresponding to the bound in [loc. cit., (3.5)]. □

The following final lemma of this section is an easy generalisation of [Drappeau and Shao 2016, Proposition 5.7], which itself generalises the $k = 1$ case established in [Harper 2016, Theorem 1].

Lemma 6.4 (Weyl sum, θ is irrational). *Suppose that $\theta = a/q + \delta$ for some $q \in \mathbb{N}$, $a \in \mathbb{Z}$ and $\delta \in \mathbb{R}$ such that $(a, q) = 1$ and $|\delta| \leq 1/(2q)$. Further, let $x > 2$, $1 \leq y' \leq (\log x)^{K'}$ for some $K' \geq 1$ and let $y'^2 \leq y \leq x$. Write $\mathcal{Q} = q + x^k \|q\theta\|$, and assume that $4y^3 \mathcal{Q}^3 \leq x$. Then there exists some constant $c = c(k) > 0$ such that*

$$E_k(x, [y', y]; \theta) \ll \mathcal{Q}^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y]).$$

Proof. The proof of [Drappeau and Shao 2016, Proposition 5.7] carries over almost directly when replacing any application of [loc. cit., Lemma 3.2] by one of Lemma 3.4(i) combined with Lemma 3.3. In doing so, we however have to ensure that the stronger condition that $1 \leq d \leq x/y$ of Lemma 3.4(i) is satisfied. For this reason we make the stronger assumption that $4y^3 \mathcal{Q}^3 \leq x$ instead of $4y^2 \mathcal{Q}^3 \leq x$. As only fairly straightforward changes are required, we explain below how to deduce our lemma from the proof of [Drappeau and Shao 2016, Proposition 5.7], instead of repeating the entire proof.

Extracting, as before, the greatest common divisor of $n \in S(x, [y', y])$ and q^∞ , we obtain

$$E_k(x, [y', y]; \theta) = \sum_{\substack{d|q^\infty \\ d \leq x}} \mathbf{1}_{S([y', y])}(d) \sum_{\substack{n \leq x/d \\ (n, q)=1}} \mathbf{1}_{S([y', y])}(n) e(n^k d^k \theta).$$

We claim that the contribution from those terms with $d > \mathcal{Q}^{1/2}$ and those terms with $n \leq x/\mathcal{Q}$ is negligible, in the sense that

$$E_k(x, [y', y]; \theta) = \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}^{1/2}}} \mathbf{1}_{S([y', y])}(d) \sum_{\substack{x/\mathcal{Q} \leq n \leq x/d \\ n \in S([y', y]) \\ (n, q)=1}} e(n^k d^k \theta) + O_\varepsilon\left(\frac{\Psi(x, [y', y])}{\mathcal{Q}^{\alpha/2-\varepsilon}}\right). \tag{6-4}$$

We start with the contribution from $\max(x/y, \mathcal{Q}^{1/2}) < d \leq x$. In this case $x/d < y < x^{1/3}$. Note that $q \leq \mathcal{Q} \leq x^{1/3}$ has $\omega(q) \leq \log x$ prime factors. Hence, the contribution to $E_k(x, [y', y]; \theta)$ is bounded by

$$\sum_{\substack{d|q^\infty \\ x/y < d \leq x}} y \leq x^{1/3} \Psi(x, \log x) \leq x^{1/3+o(1)}$$

by Erdős's bound stated in [Montgomery and Vaughan 2007, Section 7.1.1, Exercise 12]. Since $\alpha > \frac{1}{2}$ and $\mathcal{Q} < x^{1/3}$, it follows that

$$x^{1/3+o(1)} \ll \mathcal{Q}^{-\alpha/2} \Psi(x, [y', y]).$$

If $\mathcal{Q}^{1/2} < d < x/y$, then Lemma 3.4(i) applies and we obtain

$$\begin{aligned} & \sum_{\substack{d|q^\infty \\ \mathcal{Q}^{1/2} < d \leq x/y}} \Psi(x/d, [y', y]) \mathbf{1}_{S([y', y])}(d) \\ & \ll \Psi(x, [y', y]) \sum_{\substack{\mathcal{Q}^{1/2} < d \leq x/y \\ d|q^\infty}} d^{-\alpha} \mathbf{1}_{S([y', y])}(d) \\ & \ll \Psi(x, [y', y]) \mathcal{Q}^{-\alpha/2} \sum_{d|q^\infty} d^{-\alpha/2} \mathbf{1}_{S([y', y])}(d) \ll \Psi(x, [y', y]) \mathcal{Q}^{-\alpha/2} \exp\left(\sum_{\substack{p > y' \\ p|q}} p^{-\alpha/2} + O(1)\right) \\ & \ll \Psi(x, [y', y]) \mathcal{Q}^{-\alpha/2} \exp\left(O\left(y'^{-\alpha/2} \frac{\log \mathcal{Q}}{\log y'}\right)\right) \ll_\varepsilon \Psi(x, [y', y]) \mathcal{Q}^{-\alpha/2+\varepsilon}, \end{aligned}$$

since $\alpha > \frac{1}{2}$.

To remove the contribution from $n \leq x/\mathcal{Q}$ we may now assume that $d < \mathcal{Q}^{1/2}$. This contribution is trivially bounded above by

$$\sum_{d < \mathcal{Q}^{1/2}} \sum_{\substack{n \leq x/\mathcal{Q} \\ (n, d)=1}} \mathbf{1}_{S([y', y])}(dn) \leq \Psi(x/\mathcal{Q}^{1/2}; [y', y]) \ll \mathcal{Q}^{\alpha/2} \Psi(x; [y', y]),$$

where we used Lemma 3.4(i) together with the bound $\mathcal{Q}^{1/2} < x/y$ which follows from the assumption that $4y^3 \mathcal{Q}^3 \leq x$. This shows that (6-4) holds.

Following [Drappeau and Shao 2016], set $L = 4y\mathcal{Q}$ and decompose any element of $[x/\mathcal{Q}, x/d] \cap S([y', y])$ as the unique product mn for which $m \in [L, yL]$ and $P^+(m) \leq P^-(n)$, which is possible since

$y < Ly < x/\mathcal{Q}$ by the assumption that $4y^3\mathcal{Q}^3 \leq x$. Decomposing the range of m dyadically and extracting the largest prime factor of m implies that

$$E_k(x, [y', y]; \theta) \ll \log x \sup_{L \leq M \leq yL} \mathcal{E}(M) + \frac{\Psi(x, [y', y])}{\mathcal{Q}^{\alpha/4}},$$

where

$$\mathcal{E}(M) = \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}^{1/2} \\ d \in S([y', y])}} \sum_{y' \leq p \leq y} \sum_{\substack{m \in S([y', p]) \\ M/p \leq m \leq \min\{2M, yL\}/p}} \sum_{\substack{n \in S([p, y]) \\ (pmn, q)=1 \\ x/(pm\mathcal{Q}) \leq n \leq x/(pmd)}} e((pmnd)^k \theta).$$

The dependence of the summation condition of the inner sum on m can be removed with the help of the same trick as in the proof of Lemma 6.2, which leads to

$$\mathcal{E}(M) \ll (\log x) \sup_{\beta \in [0, 1)} \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}^{1/2} \\ d \in S([y', y])}} \sum_{y' \leq p \leq y} \sum_{\substack{m \in S([y', p]): \\ M/p \leq m \leq \min\{2M, yL\}/p}} \left| \sum_{\substack{n \in S([p, y]) \\ (pmn, q)=1 \\ x/(2M\mathcal{Q}) \leq n \leq x/(Md)}} e((pmnd)^k \theta + \beta n) \right|.$$

An application of Cauchy–Schwarz, combined with the bound

$$\sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}^{1/2} \\ d \in S([y', y])}} 1 \leq \mathcal{Q}^{\varepsilon/2} \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}^{1/2} \\ d \in S([y', y])}} d^{-\varepsilon} \leq \mathcal{Q}^{\varepsilon/2} \exp\left(O(1) \sum_{p|q} p^{-\varepsilon}\right) \ll \mathcal{Q}^{\varepsilon/2} \mathcal{Q}^{o(1)} \ll_\varepsilon \mathcal{Q}^\varepsilon,$$

which follows from $\omega(d) \leq (\log \mathcal{Q})/\log y'$, we obtain

$$\mathcal{E}(M) \ll_\varepsilon (\log x) \mathcal{Q}^\varepsilon M^{1/2} \mathcal{S}_1(M)^{1/2},$$

where $\mathcal{S}_1(M)$ is defined as in [Drappeau and Shao 2016, display just below (5.3)] except that the sum over primes runs over $y' \leq p \leq y$:

$$\mathcal{S}_1(M) = \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}}} \sum_{y' \leq p \leq y} \sum_{M/p < m \leq 2M/p} \left| \sum_{\substack{n \in S([p, y]) \\ (pmn, q)=1 \\ x/(2M\mathcal{Q}) \leq n \leq x/(Md)}} e((pmnd)^k \theta + \beta n) \right|^2.$$

Here we relaxed the summation conditions on d (to match the corresponding sum in [loc. cit.]) and, following [Harper 2016; Drappeau and Shao 2016], on m , which allows one to later use a standard exponential sum estimate where the summation variable runs over a full interval. Expanding out the square, swapping the order of summation, applying the triangle inequality and relaxing some of the summation conditions in the outer sum leads to

$$\mathcal{S}_1(M) \ll \sum_{\substack{d|q^\infty \\ d \leq \mathcal{Q}}} \sum_{y' \leq p \leq y} \sum_{\substack{n_1, n_2 \in S([y', y]) \\ (n_1 n_2, q)=1 \\ x/(2M\mathcal{Q}) < n_1 \leq n_2 \leq x/(Md)}} \left| \sum_{M/p < m \leq 2M/p} e((pmd)^k \theta (n_1^k - n_2^k)) \right|.$$

Note carefully that instead of relaxing the condition $n_1, n_2 \in S([p, y])$ to $n_1, n_2 \in S(y)$, as in [Drappeau and Shao 2016], we replaced it by $n_1, n_2 \in S([y', y])$. We now follow the analysis of $\mathcal{S}_1(M)$ from [loc. cit.], with the only change that in the definitions of \mathcal{S}_j , $j \in \{2, \dots, 5\}$, the variables n_1 and n_2 continue to be restricted to $S([y', y])$. Instead of bounding the quantity $\mathcal{S}_5(M; r', d; n_1, b; T)$ with the help of [loc. cit., Lemma 3.3], we use Lemma 5.7, which produces the analogous bound with $\Psi(x/(Md), y)$ replaced by $\Psi(x/(Md), [y', y])$. In the definition of \mathcal{S}'_3 , we include the restriction to $n_1 \in S([y', y])$. Adapting the analysis of \mathcal{S}_3 requires the lower bound

$$\begin{aligned} \Psi(x/(Md), y) &\asymp \left(\frac{x}{Md}\right)^{\alpha'} (\log x \log y)^{-1/2} \prod_{y' \leq p \leq y} (1 - p^{-\alpha'})^{-1} \\ &\gg \left(\frac{x}{Md}\right)^{\alpha'} (\log x \log y)^{-1/2} \gg \left(\frac{x}{Md}\right)^{\alpha' + o(1)} \gg \left(\frac{x}{Md}\right)^{\alpha}, \end{aligned}$$

where $\alpha' = \alpha(x/(Md), y)$, which follows from Lemma 3.3, the asymptotic expansion (3-3) as well as the bound $\alpha' = \alpha(x/(Md), y) > \alpha(x, y) = \alpha$ implied by the lower bound $M \leq L = 4y\mathcal{Q} > \log x$ and the estimate (3-2).

The remaining analysis in the proof of [loc. cit., Proposition 5.7] involves two applications of [loc. cit., Lemma 3.2], which may be replaced by applications of Lemma 3.4(i) combined with Lemma 3.3 since for $d \leq \mathcal{Q}$ and $M \leq yL$ we have

$$Md \leq yL\mathcal{Q} \leq 4y^2\mathcal{Q}^2 \leq x/(y\mathcal{Q}) \leq x/y,$$

thanks to our stronger assumption that $4y^3\mathcal{Q}^3 \leq x$. This ensures that $d \leq x/(yM)$ in the first application and that $M < x/y$ in the second application, i.e., that the conditions of Lemma 3.4(i) are satisfied. Our lemma thus follows from the proof of [loc. cit., Proposition 5.7]. \square

Proof of Theorem 6.1. The first part of the result follows directly from Lemma 6.3 applied with $\sigma = \frac{1}{12}$. Concerning the second part, suppose that $\theta \in \mathfrak{M}(x^{1/12}, x)$ and recall that by definition (6-1) there exists a positive integer $q \leq x^{0.1}$ such that $\|q\theta\| \leq x^{1/12}x^{-k}$. For any such value of q , we have $\mathcal{Q} = q + \|q\theta\|x^k \leq 2x^{1/12}$ and $4y^3\mathcal{Q}^3 \leq x$ since $y < x^\eta$ and $\eta < \frac{1}{4}$. It then follows from Lemma 6.4 that

$$E_k(x, [y', y]; \theta) \ll \mathcal{Q}^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y]). \quad \square$$

7. Strongly recurrent polynomial sequences over smooth numbers

Our aim in this section is to show that if a sequence $\{\|\beta n^k\|\}_{n \in S(x, [y', y])}$ is strongly recurrent, then β is very close to a rational with small denominator q . More precisely:

Theorem 7.1. *Let $N > 2$ be a parameter, let $\theta \in \mathbb{R}$, and let $k \geq 1$ be a fixed integer. Let $K' > 0$, $K > \max(2, 2K')$, $1 \leq y' \leq (\log N)^{K'}$, and suppose that $(\log N)^K \leq y \leq N^\eta$ for some small constant $\eta = \eta(k) \in (0, 1)$. Let $\delta : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ be a function of N that satisfies*

$$\log_5 N \ll \delta(N)^{-1} < \log_2 N$$

for all sufficiently large N . Finally suppose that $0 < \varepsilon \leq \delta/2$ and that $\varepsilon < \delta(N)^{-O(1)}(\log N)^{-C_1}$ for some fixed constant $C_1 \geq 1$. Then the following assertion holds provided that K and C_1 are sufficiently large depending on k .

If the bound $\|n^k\theta\| \leq \varepsilon$ holds for at least $\delta\Psi(N, [y', y])$ elements $n \in S(N, [y', y])$, then there exists a positive integer $0 < q \ll \delta^{-O(1)}$ such that

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}/N^k.$$

The corresponding problem for strongly recurrent *unrestricted* polynomial sequences $\{\|\beta n^k\|\}_{n \leq x}$ has been treated in [Green and Tao 2012a, Section 3] where the polynomial case was reduced to the linear case via an application of bounds in Waring's problem. While a suitable Waring-type result was proved in [Drappeau and Shao 2016, Theorem 2.4] for the set $S(x, y)$, no such result is currently available for the sparse subset $S(x, [y', y])$, although we expect such an analogue to hold.

For this reason, our approach proceeds instead via first reducing the problem via Fourier analysis to bounds on Weyl sums over $S(x, [y', y])$, which can then be combined with the results of Section 6. The Fourier analysis reduction is standard (see [Green and Tao 2012b, Proposition 3.1 and Lemma 3.2]). The bounds on the relevant Weyl sums obtained in Section 6 by themselves are, however, not strong enough in order to deduce the theorem above; they only provide part (1) of Lemma 7.4 below. In particular, those bounds on $\|q\theta\|$ are not strong enough in order to later analyse the correlation of $g_{[y', y]}(n)$ with nilsequences for very small values of y , that is, when $\log y \ll \log \log x$. To work around this problem, we will employ a bootstrapping argument in order to improve the bounds on $\|q\theta\|$. This argument is based on the combination of the following lemma due to [Drappeau and Shao 2016, Lemma 3.7] (which is a higher-dimensional version of the bootstrapping argument used in the proof of [Green and Tao 2012b, Lemma 3.2]) and the bounds we obtained in Section 5.2 on equidistribution of $S([y', y])$ in short progressions.

Lemma 7.2 [Drappeau and Shao 2016, bootstrapping lemma]. *Let $k > 0$ be a fixed integer and let $\varepsilon', \delta \in (0, 1)$. Let $1 \leq L \leq x$ be parameters and let $\mathcal{A} \subset [x, 2x]$ be a nonempty subset with the property that*

$$|\mathcal{A} \cap P| \leq \Delta \frac{|\mathcal{A}| |P|}{x}$$

for some $\Delta \geq 1$ and for any arithmetic progression $P \subseteq [x, 2x]$ of length at least L and common difference $q = 1$. Suppose that, for some $\vartheta \in \mathbb{R}$ with $\|\vartheta\| \leq \varepsilon'/(Lx^{k-1})$, there are at least $\delta|\mathcal{A}|$ elements $m \in \mathcal{A}$ satisfying $\|m^k\vartheta\| \leq \varepsilon'$. Then either $\varepsilon' \gg \delta/\Delta$ or

$$\vartheta \ll \Delta\delta^{-1}\varepsilon'/x^k.$$

Remark 7.3. The original statement of [Drappeau and Shao 2016, Lemma 3.7] does not restrict the progressions P to have common difference $q = 1$. An inspection of the proof reveals, however, that the corresponding assumption is only required for progressions with common difference $q = 1$, i.e., for discrete intervals.

Theorem 7.1 is an immediate consequence of the following lemma, which will help us structure our proof.

Lemma 7.4. *Let $\theta \in \mathbb{R}$, $k \geq 1$ a fixed integer, and let $K' > 0$ and $K > \max(2K', 2)$ be constants. Let $1 \leq y' \leq (\log x)^{K'}$ and suppose that $(\log x)^K \leq y \leq x^\eta$ for some small constant $\eta \in (0, 1)$. Let $\delta = \delta(x)$ be such that $\delta(x)^{-B} \ll_B x$ for all $B > 0$ and suppose that $0 < \varepsilon \leq \delta/2$. Suppose that there are at least $\delta\Psi(x, [y', y])$ elements $n \in S(x, [y', y])$ for which $\|n^k\theta\| \leq \varepsilon$. Then:*

(1) *There is some integer $0 < q \ll \delta^{-O(1)}$ and a constant $c > 0$ depending on k such that*

$$\|q\theta\| \ll \delta^{-O(1)}(\log x)^{10/c}/x^k$$

provided K is sufficiently large in terms of c .

(2) *If, furthermore, $\delta(x) > (\log_2 x)^{-1}$ and $\varepsilon < \delta(x)^{-O(1)}(\log x)^{-C_1}$ for some fixed constant $C_1 \geq 5 + 20/c$, then the integer $0 < q \ll \delta^{-O(1)}$ from part (1) satisfies*

$$\|q\theta\| \ll \varepsilon\delta^{-O(1)}(\log x)^{C_1}/x^k,$$

provided that K is sufficiently large in terms of c .

(3) *Under the assumptions of part (2) the integer $0 < q \ll \delta^{-O(1)}$ from part (1) satisfies the bound*

$$\|q\theta\| \ll \varepsilon\delta^{-O_{c_1}(1)}(\log_3 x)/x^k,$$

provided that K is sufficiently large in terms of c .

(4) *If, in addition to all previous assumptions, $\delta^{-1} \gg \log_5 x$, then the integer $0 < q \ll \delta^{-O(1)}$ from part (1) is such that*

$$\|q\theta\| \ll \varepsilon\delta^{-O_{c_1}(1)}/x^k,$$

provided that K is sufficiently large in terms of c .

Proof of Lemma 7.4. Suppose that $I \subseteq [0, 1]$ is an interval of length $|I| = \varepsilon$, and that there are at least $\delta\Psi(x, [y', y])$ elements $n \in S(x, [y', y])$ such that $n^k\theta \pmod{\mathbb{Z}} \in I$. By approximating the characteristic function of the interval I by a suitable smooth Lipschitz function F , we obtain

$$\left| \sum_{n \in S(x, [y', y])} F(n^k\theta \pmod{\mathbb{Z}}) \right| \geq \delta'\Psi(x, [y', y]) \tag{7-1}$$

for some $\delta' \in [\delta/2, \delta]$. We may assume that F is supported on a set of measure δ . Hence, on rescaling F and redefining δ' in the bound above as $\delta' = \delta^C$ for some $C \geq 1$, we may assume that $\|F\|_{\text{Lip}} = 1$, which implies that, for every $k \in \mathbb{Z}$,

$$|\hat{F}(k)| = \int_{\mathbb{R}/\mathbb{Z}} F(\theta)e(-\theta k) \, d\theta \leq \|F\|_\infty \leq \|F\|_{\text{Lip}} = 1. \tag{7-2}$$

Following the proof of [Green and Tao 2012b, Lemma 3.1], our next step is to approximate F by a function whose Fourier transform is finitely supported, with a support defined in terms of δ' . For this purpose, consider the Fejér kernel $K(\theta) = \chi_Q * \chi_Q(\theta)$, where $\chi_Q(\theta) = (\delta'/16)\mathbf{1}_Q(\theta)$ is the normalised

characteristic function of the short interval $Q = [-\delta'/16, \delta'/16]$. Since $\int_{\mathbb{R}/\mathbb{Z}} K = 1$ and by (7-2), the convolution $F_1 = F * K$ satisfies

$$\|F - F_1\|_\infty \leq \frac{\delta'}{4} \quad \text{and} \quad |\hat{F}_1(k)| \leq 1 \quad \text{for all } k \in \mathbb{Z}.$$

The rapid decay of the Fourier coefficients of K allows one to approximate F_1 by a finite truncation of its Fourier series as follows. If

$$F_2(\theta) := \sum_{0 < |q| \ll \delta'^{-3}} \hat{F}_1(q)e(q\theta), \quad \text{then} \quad \|F_2 - F_1\|_\infty \leq \frac{\delta'}{4}.$$

When applying both of these approximations to (7-1) and then swapping the order of summation, it follows from the triangle inequality that

$$\frac{\delta'}{2} \Psi(x, [y', y]) \leq \sum_{0 < |q| \ll \delta'^{-3}} |\hat{F}_1(q)| \left| \sum_{n \in S(x, [y', y])} e(qn^k\theta) \right| \leq \sum_{0 < |q| \ll \delta'^{-3}} \left| \sum_{n \in S(x, [y', y])} e(qn^k\theta) \right|.$$

By the pigeonhole principle there therefore is some integer $0 < |q| \ll \delta'^{-3}$ such that

$$\delta'^4 \Psi(x, [y', y]) \ll \left| \sum_{n \in S(x, [y', y])} e(qn^k\theta) \right| = |E_k(x, [y', y]; q\theta)|. \tag{7-3}$$

Following the above reduction via Fourier analysis, we are now in the position to invoke the bounds on Weyl sums over smooth numbers from Section 6. To start with, the first part of Theorem 6.1 shows that if $q\theta \pmod{\mathbb{Z}}$ does not belong to $\mathfrak{M}(x^{1/12}, x)$, then the right-hand side of (7-3) is bounded by

$$E_k(x, [y', y]; q\theta) \ll x^{1-c} \ll x^\alpha x^{1-\alpha-c} \ll x^{-c'} \Psi(x, [y', y]),$$

where $c' := c - (1 - \alpha)$ and where the lower bound $\Psi(x, [y', y]) \gg x^{\alpha+o(1)}$ follows from (3-3) and Lemma 3.2. Observe that $c' > 0$ provided that K is sufficiently large in terms of c since $1 - \alpha(x, y) \leq 1/K + o(1)$ by (3-2). Hence, $\delta'^4 \ll x^{-c'}$ and, thus, $x \ll \delta'^{-4/c'} = \delta^{-4C/c'}$, contradicting our assumptions on δ , which implies $\delta^{-4C/c'} \ll_B x^{1/B}$ for all $B > 0$.

It follows that $q\theta \pmod{\mathbb{Z}} \in \mathfrak{M}(x^{1/12}, x)$. In this case, the second part of Theorem 6.1, applied with θ replaced by $q\theta$ (and applied in the special case where, in the statement of Theorem 6.1, $q = a = 1$), shows that

$$\sum_{n \in S(x, [y', y])} e(qn^k\theta) \ll (1 + x^k \|q\theta\|)^{-c+2(1-\alpha)} (\log x)^5 \Psi(x, [y', y]).$$

Suppose that K is sufficiently large to ensure that $c - 2(1 - \alpha) > c/2$. Then it follows from the Fourier analysis bound (7-3) that

$$1 + x^k \|q\theta\| \ll \delta^{-8C/c} (\log x)^{10/c},$$

and hence

$$\|q\theta\| \ll \delta^{-8C/c} (\log x)^{10/c} x^{-k},$$

which proves part (1). For later use in the proofs of all remaining parts, we record the following consequence of part (1) and our assumptions.

Recurrence condition. Let $0 < q \ll \delta^{-O(1)}$ denote the integer produced by part (1). Then we have

$$\|n^k q \theta\| \leq q \|n^k \theta\| < \varepsilon' \quad (7-4)$$

for some $\varepsilon' \asymp q \varepsilon \ll \delta^{-O(1)} \varepsilon$ and for at least $\delta \Psi(x, [y', y])$ elements n of $S(x, [y', y])$.

To establish part (2), observe that in view of Lemma 5.7, the conditions of Lemma 7.2 are satisfied for the set $\mathcal{A} = S(2x, [y', y])$ and any lower bound $L \geq 1$ with a correction factor of the form $\Delta = (x/L)^{1-\alpha} \log x$. Moreover, we have

$$\|q \theta\| \ll \delta^{-8C/c} (\log x)^{10/c} x^{-k} = \varepsilon' / (L x^{k-1})$$

if we set $L = \varepsilon' \delta^{8C/c} (\log x)^{-10/c} x$ and with ε' as in (7-4). In this case,

$$\begin{aligned} \varepsilon' \Delta &\leq \varepsilon' (\log x) (\varepsilon' \delta^{8C/c} (\log x)^{-10/c})^{\alpha-1} \ll (q \varepsilon)^\alpha \delta^{-O(1)} (\log x)^{1+5/c} \\ &\ll (\log x)^{-\alpha C_1} \delta^{-O(1)} (\log x)^{1+5/c} = o(\delta(x)) \end{aligned}$$

provided that C_1 is sufficiently large depending on c . Here, we used that $\alpha > \frac{1}{2}$ if $K > 2$. To find a simple bound for Δ , note that

$$x/L \ll (\varepsilon q \delta^{8C/c} (\log x)^{-10/c})^{-1} \ll (\log x)^{C_1} \delta^{-O(1)} (\log x)^{10/c} \ll (\log x)^{3C_1/2},$$

provided that C_1 is sufficiently large (e.g., $C_1 \geq 5 + 20/c$), which implies that

$$\Delta = (x/L)^{1-\alpha} \log x \ll (\log x)^{1+3C_1(1-\alpha)/2} \ll (\log x)^{3C_1/4+1} \leq (\log x)^{C_1}.$$

Thus, in view of the recurrence condition above, the conclusion of part (2) follows from Lemma 7.2, applied with the given value of ε' and $\vartheta := q \theta$, provided that C_1 is sufficiently large in terms of c .

To prove part (3), we use the information from part (2), i.e., that the positive integer $q \ll \delta^{-O(1)}$ produced by part (1) satisfies

$$\|q \theta\| \ll \varepsilon \delta^{-O(1)} (\log x)^{C_1} / x^k. \quad (7-5)$$

We shall now apply the short intervals case of Corollary 5.6(1). This result and the recurrence condition imply that the conditions of the bootstrapping lemma are satisfied for the set $\mathcal{A} = S(x, [y', y])$, the lower bound $L = x / (\log x)^\ell$ and

$$\Delta = \log_3 x + (\log x)^{\ell-1/24}$$

for any constant $\ell > 0$. Observe that, since $\varepsilon < \delta^{-O(1)} (\log x)^{-C_1}$ and $\delta(x)^{-1} < \log_2 x$, a bound of the form $\ell \leq C_1$ implies

$$\begin{aligned} \varepsilon' \delta(x)^{-O(1)} \Delta &< \delta(x)^{-O(1)} (\log_3 x + (\log x)^{\ell-1/24}) (\log x)^{-C_1} \\ &< (\log x)^{-C_1+o(1)} + (\log x)^{-1/24+o(1)} = o(\delta(x)). \end{aligned}$$

This ensures that the bootstrapping lemma produces bounds on $\|q \theta\|$ in all applications below.

Let $0 \leq j < 24C_1$ be an integer, and suppose inductively that

$$\|q \theta\| \ll \varepsilon \delta^{-O_{C_1}(1)} (\log x)^{C_1-j/24} / x^k,$$

the case $j = 0$ being (7-5). Let $\ell := C_1 - j/24$, and $L = x/(\log x)^\ell$. Then the bootstrapping lemma, applied with ε' replaced by $\varepsilon\delta^{-O_{C_1}(1)}$, implies that

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log x)^{C_1-(j+1)/24}/x^k$$

if $j + 1 < 24C_1$. Recalling the shape of Δ , we pick up a $(\log_3 x)$ -factor when $j + 1 \geq 24C_1$ and thus $\ell - 1/24 \leq 0$, and obtain

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log_3 x)/x^k$$

as claimed.

It remains to establish part (4). In view of part (3), we have

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}(\log_3 x)/x^k$$

and there are $\delta\Psi(x, [y', y])$ elements $n \in S(x, [y', y])$ for which $\|n^k q\theta\| < \varepsilon'$ by (7-4). We shall now appeal to Corollary 5.6(2) in order to remove the $(\log_3 x)$ -factor from the bound. For this purpose let $\varpi(x) = \log_4 x$ and recall the recurrence condition. Then Corollary 5.6(2) shows that the conditions of Lemma 7.2 are satisfied for the set $\mathcal{A} = S(x, [y', y])$, the lower bound $L = x/\log_3 x$ and $\Delta = \log \varpi(x) \ll \log_5 x \ll \delta^{-1}$. By the assumptions on δ and $\varpi(x)$ it follows easily that

$$\Delta\varepsilon'\delta^{-O_{C_1}(1)} \ll \delta^{-O_{C_1}(1)}(\log x)^{-C_1} = o(\delta).$$

Hence, Lemma 7.2 applied with ε' replaced by $\varepsilon\delta^{-O_{C_1}(1)}$ finally leads to the bound

$$\|q\theta\| \ll \varepsilon\delta^{-O_{C_1}(1)}/x^k. \quad \square$$

8. Noncorrelation with nilsequences: a reduction and general lemmas

The aim of this section is to first establish an initial reduction of the noncorrelation estimate stated in Theorem 1.1 to the case where the nilsequence is equidistributed. In preparation for the proof of the reduced version we then show that most sequences in certain sparse families of subsequences of an equidistributed polynomial sequence are equidistributed. We start by recalling some notation around nilsequences.

Definition 8.1 (filtered nilmanifold). Let $d, m_G \geq 0$ be integers and let $M > 0$. We define a *filtered nilmanifold* G/Γ of degree d , dimension m_G and complexity at most M to be an s -step nilmanifold G/Γ , for some $1 \leq s \leq d$, of dimension m_G in the sense of [Green and Tao 2012b, Definition 1.1] such that G is equipped with a filtration G_\bullet of degree $d \geq s$ in the sense of that definition and such that the Lie algebra $\mathfrak{g} = \log G$ is equipped with an M -rational Malcev basis adapted to G_\bullet in the sense of [loc. cit., Definition 2.1].

A Malcev basis \mathcal{X} gives rise to a metric $d_{\mathcal{X}}$ on G/Γ (see [loc. cit., Definition 2.2]). With respect to this metric, Lipschitz functions can be defined. More precisely, if $F : G/\Gamma \rightarrow \mathbb{C}$, we define (see [loc. cit., Definition 1.2]) the Lipschitz norm

$$\|F\|_{\text{Lip}} := \|F\|_\infty + \sup_{x, y \in G/\Gamma, x \neq y} \frac{|F(x) - F(y)|}{d_{\mathcal{X}}(x, y)}$$

and call F a Lipschitz function if $\|F\|_{\text{Lip}} < \infty$.

Definition 8.2 (polynomial sequence and nilsequence). Given a nilpotent Lie group G and a filtration

$$G_\bullet : G = G_0 = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_d \supseteq G_{d+1} = \{\text{id}_G\},$$

the set $\text{poly}(\mathbb{Z}, G_\bullet)$ of *polynomial sequences* is defined as the set of all maps $g : \mathbb{Z} \rightarrow G$ such that, if $\partial_h g(n) := g(n+h)g(n)^{-1}$, the j -th discrete derivative $\partial_{h_j} \dots \partial_{h_1} g$ takes values in G_j for all $j \in \{1, \dots, d+1\}$ and all $h_1 \dots h_j \in \mathbb{Z}$. If $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ and $F : G/\Gamma \rightarrow \mathbb{C}$ is a Lipschitz function, then the sequence $\mathbb{Z} \rightarrow F(g(\cdot)\Gamma)$ is called a *nilsequence*.

With this notation in place, we restate the main result of our paper, which shows that $n \mapsto (g_{[y',y]}^{(W,A)}(n) - 1)$ is orthogonal to nilsequences.

Theorem 8.3 (noncorrelation with nilsequences). *Let N be a large positive parameter and let $K' \geq 1$, $K > 2K'$ and $d \geq 0$ be integers. Let $\frac{1}{2} \log_3 N \leq y' \leq (\log N)^{K'}$ and suppose that $(\log N)^K < y_0 < y < N^\eta$ for some sufficiently small $\eta \in (0, 1)$ depending the value of d . Let $(G/\Gamma, G_\bullet)$ be a filtered nilmanifold of complexity Q_0 and degree d . Finally, let $w(N) = \frac{1}{2} \log_3 N + o(1)$, $W = P(w(N))$ and define $\delta(N) = \exp(-\sqrt{\log_4 N})$.*

If K is sufficiently large depending on the degree d of G_\bullet , then the estimate

$$\left| \frac{\tilde{W}}{N} \sum_{n \leq (N-A)/\tilde{W}} (g_{[y',y]}^{(\tilde{W},A)}(n) - 1) F(g(n)\Gamma) \right| \ll_{d,C} (1 + \|F\|_{\text{Lip}}) \delta(N) Q_0 + \frac{1}{\log w(N)} \quad (8-1)$$

holds uniformly for all $\tilde{W} = Wq$, where $q \leq (\log y_0)^C$ satisfies $p \mid q \Rightarrow p < w(N)$ and $C \geq 1$ is a fixed constant, for all $1 \leq A \leq \tilde{W}$ with $\text{gcd}(A, W) = 1$, all polynomial sequences $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ and all 1-bounded Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$.

By decomposing the summation range $1 \leq n \leq (N - A)/\tilde{W}$ into dyadic intervals $n \sim (N' - A)/\tilde{W}$ for $N' \in (N/\log N, N]$ and the initial segment $1 \leq n \ll N/(\tilde{W} \log N)$, and approximating $g_{[y',y]}$ by $h_{[y',y]}$ on each dyadic interval, it follows from Lemma 4.4 that the task of proving Theorem 8.3 may be reduced to proving that the estimate

$$\left| \frac{\tilde{W}}{N} \sum_{n \sim (N-A)/\tilde{W}} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1) F(g(n)\Gamma) \right| \ll_{d,C} (1 + \|F\|_{\text{Lip}}) \delta(N) Q_0 + \frac{1}{\log w(N)} \quad (8-2)$$

holds under the assumptions of Theorem 8.3 and for the function $h_{[y',y]}^{(\tilde{W},A)}$ that is defined on the interval $n \sim (N - A)/\tilde{W}$.

8.1. Reduction to noncorrelation with equidistributed nilsequences. Observe that the bound in the statement of Theorem 8.3 holds trivially unless

$$Q_0 \leq \delta(N)^{-1}.$$

This information can be used in combination with Green and Tao’s factorisation theorem for polynomial sequences (which states that every polynomial sequence is the product of a slowly varying sequence, a highly equidistributed sequence and a periodic polynomial sequence) together with our results on the distribution of $n \mapsto h_{[y',y]}$ in short arithmetic progressions in order to reduce the statement to one in which

the sequence g can be assumed to be equidistributed. Before stating the reduced version and proving the reduction, we recall the relevant definitions around equidistribution as well as the factorisation theorem.

Definition 8.4 (δ -equidistributed and totally δ -equidistributed sequence [Green and Tao 2012b, Definition 1.2]). Let G/Γ be a nilmanifold.

(1) Given a length $N > 0$ and an error tolerance $\delta > 0$, a finite sequence $(g(n)\Gamma)_{n \in [N]}$ is said to be δ -equidistributed if we have

$$\left| \mathbb{E}_{n \in [N]} F(g(n)\Gamma) - \int_{G/\Gamma} F \right| \leq \delta \|F\|_{\text{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$.

(2) A finite sequence $(g(n)\Gamma)_{n \in [N]}$ is said to be *totally δ -equidistributed* if we have

$$\left| \mathbb{E}_{n \in P} F(g(n)\Gamma) - \int_{G/\Gamma} F \right| \leq \delta \|F\|_{\text{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$ and all arithmetic progressions $P \subset [N]$ of length at least δN .

Definition 8.5 (rational sequence [Green and Tao 2012b, Definition 1.17]). Let G/Γ be a nilmanifold and let $Q > 0$ be a parameter. We say that $\gamma \in G$ is Q -rational if $\gamma^r \in \Gamma$ for some integer r , $0 < r \leq Q$. A Q -rational point is any point in G/Γ of the form $\gamma\Gamma$ for some Q -rational group element γ . A sequence $(\gamma(n))_{n \in \mathbb{Z}}$ is Q -rational if every element $\gamma(n)\Gamma$ in the sequence is a Q -rational point.

Definition 8.6 (smooth sequences [Green and Tao 2012b, Definition 1.18]). Let G/Γ be a nilmanifold with a Malcev basis \mathcal{X} . Let $(\varepsilon(n))_{n \in \mathbb{Z}}$ be a sequence in G , and let $M, N \geq 1$. We say that $(\varepsilon(n))_{n \in \mathbb{Z}}$ is (M, N) -smooth if we have $d(\varepsilon(n), \text{id}_G) \leq M$ and $d(\varepsilon(n), \varepsilon(n-1)) \leq M/N$ for all $n \in [N]$, where id_G denotes the identity element of G .

The following proposition is Green and Tao’s factorisation theorem [2012b, Theorem 1.19] for polynomial sequences, which asserts that any polynomial sequence can be decomposed into a product of a smooth, a highly equidistributed and a rational polynomial sequence.

Proposition 8.7 (Green–Tao factorisation theorem [2012b]; see also [Tao and Teräväinen 2023]). *Let $m, d \geq 1$, and let $M_0, N \geq 1$ and $A > 0$ be real numbers. Suppose that G/Γ is an m -dimensional nilmanifold together with a filtration G_\bullet of degree d . Suppose that \mathcal{X} is an M_0 -rational Malcev basis adapted to G_\bullet and that $g \in \text{poly}(\mathbb{Z}, G_\bullet)$. Then there is an integer M with $M_0 \leq M \ll M_0^{O_{A,m,d}(1)}$, a rational subgroup $G' \subseteq G$, a Malcev basis \mathcal{X}' for G'/Γ' in which each element is an M -rational combination of the elements of \mathcal{X} , and a decomposition $g = \varepsilon g' \gamma$ into polynomial sequences $\varepsilon, g', \gamma \in \text{poly}(\mathbb{Z}, G_\bullet)$ with the following properties:*

- (1) $\varepsilon : \mathbb{Z} \rightarrow G$ is (M, N) -smooth.
- (2) $g' : \mathbb{Z} \rightarrow G'$ takes values in G' , and the finite sequence $(g'(n)\Gamma')_{n \in [N]}$ is totally $1/M^A$ -equidistributed in G'/Γ' , using the metric $d_{\mathcal{X}'}$ on G'/Γ' .
- (3) $\gamma : \mathbb{Z} \rightarrow G$ is M -rational, and $(\gamma(n)\Gamma)_{n \in \mathbb{Z}}$ is periodic with period at most M .

With the above notation in place, we may now state the equidistributed version of Theorem 8.3 and deduce this theorem from it.

Proposition 8.8 (orthogonality to equidistributed nilsequences). *Let N be a large positive parameter, let $K' \geq 1$, $K > 2K'$ and $d \geq 0$ be integers. Let $\frac{1}{2} \log_3 N \leq y' \leq (\log N)^{K'}$ and suppose that $(\log N)^K < y_0 < y \leq N^\eta$ for some $\eta \in (0, 1)$ that is sufficiently small depending on d . Let $w(N) = \frac{1}{2} \log_3 N$, $W = P(w(N))$ and $\delta(N) = \exp(-C_0 \sqrt{\log_4 N})$ with $C_0 \in [1, (\log N)^{1/4}]$.*

Let G/Γ be a nilmanifold together with a filtration G_\bullet of degree d , and suppose that \mathcal{X} is a $\delta(N)^{-1}$ -rational Malcev basis adapted to G_\bullet . Let $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ be any polynomial sequence such that the finite sequence $(g(n)\Gamma)_{n \leq 2N/W}$ is totally $\delta(N)^{E_0}$ -equidistributed for some $E_0 > 1$. Let $F : G/\Gamma \rightarrow \mathbb{C}$ be any 1-bounded Lipschitz function such that $\int_{G/\Gamma} F = 0$.

If $1 \leq q \leq (\log y_0)^C$ is $(w(N)-1)$ -smooth, where $1 \leq C \ll 1$, if $0 \leq a < q$ and $0 < A < W$ are integers such that $(W, A) = 1$ (and thus $(Wq + A, Wq) = 1$), and if $0 < N_1 \leq N$, then we have

$$\left| \frac{Wq}{N} \sum_{\substack{m \in \mathbb{N}: \\ N < Wqm < N + N_1}} h_{[y', y]}^{(W, A)}(qm + a) F(g(m)\Gamma) \right| \ll_{d, \dim G, \|F\|_{\text{Lip}}, E_1} \delta(N)^{E_1} \tag{8-3}$$

for any given $E_1 \geq 1$, provided that E_0 is sufficiently large with respect to d , $\dim G$ and E_1 , provided that K is sufficiently large depending on the degree d of G_\bullet , and provided that N is sufficiently large in terms of $\dim G$, d and E_0 .

Proposition 8.8 implies Theorem 8.3. We shall prove that Proposition 8.8 implies (8-2), from which Theorem 8.3 follows. We may assume that $Q_0 \leq \delta(N)^{-1}$ as (8-2) is trivially true otherwise. Let $B > 1$ be a parameter. Then, by Proposition 8.7, applied with N replaced by $2N/\tilde{W}$, there exists $Q_0 \leq Q \ll Q_0^{O_{B, \dim G, d}(1)}$ and a factorisation of the polynomial sequences g as $\varepsilon g' \gamma$ that satisfies properties (1)–(3) of that proposition. In particular, the polynomial sequence $\gamma : \mathbb{Z} \rightarrow G$ gives rise to a \tilde{q} -periodic function $\gamma(\cdot)\Gamma : \mathbb{Z} \rightarrow G/\Gamma$ for some period $1 \leq \tilde{q} \leq Q$, and the sequence $\varepsilon : \mathbb{Z} \rightarrow G$ is $(Q, 2N/\tilde{W})$ -smooth. The sequence $g' : \mathbb{Z} \rightarrow G'$ takes values in a Q -rational subgroup G' of G , it is a polynomial sequence with respect to the filtration $G'_\bullet := G_\bullet \cap G'$ and the finite sequence $(g'(n)\Gamma')_{n \leq 2N/\tilde{W}}$ is Q^{-B} -equidistributed in G'/Γ' , where $\Gamma' = \Gamma \cap G'$ and where equidistribution is defined with respect to the metric $d_{\mathcal{X}'}$ arising from a Malcev basis \mathcal{X}' adapted to G'_\bullet . The existence of the Malcev basis \mathcal{X}' is guaranteed by [Green and Tao 2012b, Lemma A.10], which also allows us to assume that each of its basis elements is a Q -rational combination of basis elements from \mathcal{X} .

In order to reduce the noncorrelation estimate to the case where the polynomial sequence is highly equidistributed, we seek to decompose the summation range of n in (8-2) into (short) subprogressions on which γ and ε are almost constant. Splitting the interval $(N/\tilde{W}, 2N/\tilde{W}]$ into arithmetic progressions with common difference \tilde{q} , we obtain

$$\sum_{n \sim N/\tilde{W}} (h_{[y', y]}^{(\tilde{W}, A)}(n) - 1) F(g(n)\Gamma) = \sum_{0 \leq a < \tilde{q}} \sum_{\substack{n \sim N/\tilde{W} \\ n \equiv a \pmod{\tilde{q}}}} (h_{[y', y]}^{(\tilde{W}, A)}(n) - 1) F(\varepsilon(n)g'(n)\gamma_a\Gamma),$$

where $\gamma_a \in G$ is such that $\gamma(n)\Gamma = \gamma_a\Gamma$ whenever $n \equiv a \pmod{\tilde{q}}$.

Since F is a Lipschitz function and since $d_{\mathcal{X}}$ is right-invariant (see [Green and Tao 2012b, Appendix A]), we deduce that, for any $n_0, n \in \mathbb{Z}$,

$$\begin{aligned} |F(\varepsilon(n_0)g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)| &\leq \|F\|_{\text{Lip}} d_{\mathcal{X}}(\varepsilon(n_0)g'(n)\gamma_a, \varepsilon(n)g'(n)\gamma_a) \\ &= \|F\|_{\text{Lip}} d_{\mathcal{X}}(\varepsilon(n_0), \varepsilon(n)). \end{aligned}$$

It then follows from the assumption ε is $(Q, 2N/\tilde{W})$ -smooth that

$$d_{\mathcal{X}}(\varepsilon(n_0), \varepsilon(n)) \leq \frac{Q|n_0 - n|}{N/\tilde{W}}$$

whenever $n, n_0 \leq 2N/\tilde{W}$, and therefore

$$|F(\varepsilon(n_0)g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)| \ll_{\|F\|_{\text{Lip}}} \log^{-1} Q \tag{8-4}$$

if, in addition, $|n_0 - n| \ll N/(\tilde{W}Q \log Q)$. With this in mind, we refine the partition of our summation range and consider a subpartition

$$(N/\tilde{W}, 2N/\tilde{W}] = \bigcup_j P_j,$$

where each P_j is a ‘‘short’’ progression of common difference \tilde{q} , as before, but with diameter bounded by $O(N/(Q\tilde{W} \log Q))$. The bound on the diameter ensures that ε is almost constant on each P_j . Note that the total number of short progressions P_j is $O(\tilde{q}Q \log Q)$. By fixing an element $\varepsilon_j \in \varepsilon(P_j)$ in the image of P_j under ε for each progression P_j , we obtain

$$\begin{aligned} &\sum_{n \sim N/\tilde{W}} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1)F(g(n)\Gamma) \\ &= \sum_j \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1)F(\varepsilon_j g'(n)\gamma_a\Gamma) \\ &\quad + O\left\{ \sup_{P_j} \sup_{n \in P_j} |F(\varepsilon_j g'(n)\gamma_a\Gamma) - F(\varepsilon(n)g'(n)\gamma_a\Gamma)| \left(\sum_{n \sim N/\tilde{W}} h_{[y',y]}^{(\tilde{W},A)}(n) + \frac{N}{\tilde{W}} \right) \right\}. \end{aligned} \tag{8-5}$$

By construction or, more precisely, by Lemma 5.5, we have

$$\sum_{n \sim N/\tilde{W}} h_{[y',y]}^{(\tilde{W},A)}(n) \ll N/\tilde{W}.$$

Bounding the error in (8-5) with the help of this bound and (8-4), we obtain

$$\sum_{n \sim N/\tilde{W}} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1)F(g(n)\Gamma) = \sum_j \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1)F(\varepsilon_j g'(n)\gamma_a\Gamma) + O\left(\frac{N}{\tilde{W} \log Q}\right).$$

Observe that in the argument of F , apart from two constant factors, only the sequence g' occurs, which has the property that $(g'(n)\Gamma)_{n \leq 2N/\tilde{W}}$ is Q^{-B} -equidistributed. We aim to use this equidistribution property in combination with Proposition 8.8 in order to bound the correlations on the right-hand side above. For this purpose, we shall now first show that the sequence $n \mapsto \varepsilon_j g'(n)\gamma_a$ can be reinterpreted as a polynomial sequence $n \mapsto g^*(n)$ that is equidistributed on some filtered nilmanifold H/Λ . At the same

time, we show that $F(\varepsilon_j g'(n) \gamma_a \Gamma) = \tilde{F}(g^*(n) \Lambda)$ for a Lipschitz function $\tilde{F} : H/\Lambda \rightarrow \mathbb{C}$. As a second step, we carry out a reduction that allows us to assume that $\int_{H/\Lambda} \tilde{F} = 0$. And, thirdly and finally, we will apply Proposition 8.8 and complete the proof. We denote these steps (1)–(3) below.

(1) Define $g^*(n) := \gamma_a^{-1} g'(n) \gamma_a$. For an application of Proposition 8.8, it is necessary to verify that g^* is a polynomial sequence and that it inherits the equidistribution properties of g' . These questions have been addressed by Green and Tao [2012a, Section 2] and we follow their argument here. Let $H = \gamma_a^{-1} G' \gamma_a$ and define $H_\bullet = \gamma_a^{-1}(G')_\bullet \gamma_a$. Let $\Lambda = \Gamma \cap H$ and define $\tilde{F} = \tilde{F}_{a,j} : H/\Lambda \rightarrow \mathbb{R}$ via

$$\tilde{F}(x\Lambda) = F(\varepsilon_j \gamma_a x \Gamma).$$

Then $g^* \in \text{poly}(\mathbb{Z}, H_\bullet)$, we have $\tilde{F}(g^*(n) \Lambda) = F(\varepsilon_j g'(n) \gamma_a \Gamma)$, and the correlation that we seek to bound takes the form

$$\sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1) \tilde{F}(g^*(n) \Lambda). \quad (8-6)$$

The ‘‘Claim’’ from the end of [Green and Tao 2012a, Section 2] guarantees the existence of a Malcev basis \mathcal{Y} for H/Λ adapted to H_\bullet such that each basis element Y_i is a $Q^{O(1)}$ -rational combination of basis elements X_i from \mathcal{X} . Thus, there is $C' = O(1)$ such that \mathcal{Y} is $Q^{C'}$ -rational. Furthermore, the ‘‘Claim’’ implies that there is $c' > 0$, depending only on $\dim G$ and on the degree d of G_\bullet , such that whenever B is sufficiently large the sequence

$$(g^*(n) \Lambda)_{n \leq 2N/\tilde{w}} \quad (8-7)$$

is totally $Q^{-c'B+O(1)}$ -equidistributed in H/Λ , equipped with the metric $d_{\mathcal{Y}}$ induced by \mathcal{Y} . Taking B sufficiently large, we may assume that the sequence (8-7) is totally $M^{-c'B/2}$ -equidistributed. Finally, the ‘‘Claim’’ also provides the bound $\|\tilde{F}\|_{\text{Lip}} \leq Q^{C''} \|F\|_{\text{Lip}}$ for some $C'' = O(1)$. This shows that all conditions of Proposition 8.8 are satisfied except for $\int_{H/\Lambda} \tilde{F} = 0$.

(2) Let $\mu(\tilde{F}) = \int_{H/\Gamma} \tilde{F}$ denote the mean value of \tilde{F} and observe that $\mu(\tilde{F}) \ll 1$ since \tilde{F} is 1-bounded. Then $\int_{H/\Gamma} \bar{F} = 0$ if $\bar{F} := \tilde{F} - \mu(\tilde{F}) : H/\Gamma \rightarrow \mathbb{C}$ and

$$\begin{aligned} \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1) \tilde{F}(g^*(n) \Lambda) &= \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1) \bar{F}(g^*(n) \Lambda) + O\left(\left|\sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1)\right|\right) \\ &= \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1) \bar{F}(g^*(n) \Lambda) + O\left(\frac{|P_j|}{\log w(N)}\right) \end{aligned}$$

by Lemma 5.5. We may thus assume that $\int_{H/\Gamma} \bar{F} = 0$.

(3) By the previous two steps it remains to bound

$$\sum_j \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W}, A)}(n) - 1) \tilde{F}(g^*(n) \Lambda),$$

where we may assume that $\int_{H/\Gamma} \tilde{F} = 0$ and that $(g^*(n) \Lambda)_{n \leq 2N/\tilde{w}}$ is $Q^{-c'B/2}$ -equidistributed. Since P_j has common difference $\tilde{q} < Q$, the bound on the diameter of P_j implies that $|P_j| \gg N/(\tilde{q} Q \tilde{W} \log Q)$.

Note further that $Q \ll Q_0^{O_{B,d,\dim G}(1)} \ll \delta(N)^{-C}$ for some $C = O_{B,d,\dim G}(1)$ since $Q_0 \leq \delta(N)^{-1}$. We may suppose that N is sufficiently large for $C \leq (\log N)^{1/4}$ to hold. Define $\delta'(N) = \delta(N)^{-C}$.

We may thus apply Proposition 8.8 with Wq replaced by $\tilde{W}\tilde{q} = W(q\tilde{q})$, with δ replaced by δ' , with $E_0 = c'B/2$, $g = g^*$, $G/\Gamma = H/\Lambda$, $\mathcal{X} = \mathcal{Y}$, and with $N_1 \gg N/(Q \log Q) > N\delta'(N)^{E_1/2}$ (assuming that $E_1 > 2$, which we may) to deduce that

$$\begin{aligned} \sum_j \sum_{n \in P_j} (h_{[y',y]}^{(\tilde{W},A)}(n) - 1) \tilde{F}(g^*(n)\Lambda) &\ll_d \sum_j (1 + \|\tilde{F}\|_{\text{Lip}}) \delta'(N)^{E_1} \frac{N}{\tilde{W}\tilde{q}} \\ &\ll_d (1 + Q^{C''} \|F\|_{\text{Lip}}) \delta'(N)^{E_1} \frac{\tilde{q} Q N \log Q}{\tilde{W}\tilde{q}} \\ &\ll_d (1 + \|F\|_{\text{Lip}}) Q^{O(1)} \delta'(N)^{E_1} \frac{N}{\tilde{W}} \\ &\ll_d (1 + \|F\|_{\text{Lip}}) \delta(N) Q_0 \frac{N}{\tilde{W}}, \end{aligned}$$

provided that B , and hence E_1 , is sufficiently large to imply the final bound. This completes the proof of the deduction of (8-2) and, hence, Theorem 8.3. □

8.2. Sparse families of linear subsequences of equidistributed nilsequences. In Section 9.2 we will relate our task of bounding the one-parameter correlation (8-3) to that of bounding a bilinear sum. This reduction naturally leads to the problem of understanding equidistribution properties in families of linear subsequences of polynomial sequences that arise as follows. Let $(g(n)\Gamma)_{n \leq N}$ be a polynomial sequence and consider the family of sequences

$$\{n \mapsto (g(mn)\Gamma)_{n \leq N/m} \}_{m \in S([y',y]) \cap [M,2M)},$$

where the parameter $m \in [M, 2M)$ is further restricted to the sparse set $S([y', y])$, yielding a sparse family of subsequences. Our aim in this subsection is to show that if $(g(n)\Gamma)_{n \leq N}$ is δ -equidistributed for a suitable choice of δ , then almost all sequences in this family are $\delta^{1/C}$ -equidistributed. Equidistribution properties in unrestricted families of linear subsequences have been studied by the first author in [Matthiesen 2018, Section 7]. In this section we show that, thanks to the sparse recurrence result from Section 7, the unrestricted result [loc. cit., Proposition 7.4] can indeed be extended to the sparse situation where $m \in S([y', y])$. Part of this section follows [loc. cit., Section 7] very closely.

The proof of Proposition 8.11 below uses the notion of a *horizontal character* [Green and Tao 2012b, Definition 1.5] on a nilmanifold G/Γ , which is defined to be a continuous additive homomorphism $\eta : G \rightarrow \mathbb{R}/\mathbb{Z}$ which annihilates Γ . The set of horizontal characters may be equipped with a height function $|\eta|$ as defined in [loc. cit., Definition 2.6]. This specific height function is called the *modulus* of η . All that is relevant to us in the present paper is that there are at most $Q^{O(1)}$ horizontal characters $\eta : G \rightarrow \mathbb{R}/\mathbb{Z}$ of modulus $|\eta| \leq Q$.

If $\eta : G \rightarrow \mathbb{R}/\mathbb{Z}$ is a horizontal character and $g \in \text{poly}(\mathbb{Z}, G_\bullet)$, where G_\bullet is a filtration of degree d , then $\eta \circ g : \mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ is a polynomial of degree at most d . For an arbitrary polynomial $P : \mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ of

degree at most d , we may define two sets of coefficients, $\alpha_0, \dots, \alpha_d$ and β_0, \dots, β_d , in \mathbb{R}/\mathbb{Z} via

$$P(n) = \alpha_0 + \alpha_1 \binom{n}{1} + \dots + \alpha_d \binom{n}{d} = \beta_d n^d + \dots + \beta_1 n + \beta_0.$$

The *smoothness norm* [Green and Tao 2012b, Definition 1.5] of P with respect to N is then defined to be

$$\|P\|_{C^\infty[N]} = \sup_{1 \leq j \leq d} N^j \|\alpha_j\|_{\mathbb{R}/\mathbb{Z}}$$

and we have

$$\|P\|_{C^\infty[N]} \ll_d \sup_{1 \leq j \leq d} N^j \|\beta_j\|_{\mathbb{R}/\mathbb{Z}} \quad \text{and} \quad \sup_{1 \leq j \leq d} N^j \|q\beta_j\|_{\mathbb{R}/\mathbb{Z}} \ll \|P\|_{C^\infty[N]}$$

for some positive integer $q \ll_d 1$ by [loc. cit., Lemma 3.2].

Smoothness norms and horizontal characters allow one to characterise δ -equidistributed polynomial sequences in the following sense:

Lemma 8.9 [Green and Tao 2012b, Theorem 2.9]. *Let m_G and d be nonnegative integers, let $0 < \delta < \frac{1}{2}$ and let $N \geq 1$. Suppose that G/Γ is an m_G -dimensional nilmanifold together with a filtration G_\bullet of degree d and that \mathcal{X} is a δ^{-1} -rational Malcev basis adapted to G_\bullet . Suppose $g \in \text{poly}(\mathbb{Z}, G_\bullet)$. If $(g(n)\Gamma)_{n \leq N}$ is not δ -equidistributed, then there exists a nontrivial horizontal character η with $0 < |\eta| \ll \delta^{-O_{d,m_G}(1)}$ such that*

$$\|\eta \circ g\|_{C^\infty[N]} \ll \delta^{-O_{d,m_G}(1)},$$

where C_d is a sufficiently large constant depending only on d .

In order to pass between the notions of equidistribution and total equidistribution for polynomial sequences, we shall use the following lemma, which is [Matthiesen 2018, Lemma 7.2].

Lemma 8.10. *Let N and A be positive integers and let $\delta : \mathbb{N} \rightarrow [0, 1]$ be a function that satisfies $\delta(x)^{-t} \ll_t x$ for all $t > 0$. Suppose that G has a $\delta(N)^{-1}$ -rational Malcev basis adapted to the filtration G_\bullet . Then there is $1 \leq B \ll_{d, \dim G} 1$ such that the following holds provided N is sufficiently large. If $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ is a polynomial sequence such that $(g(n)\Gamma)_{n \leq N}$ is $\delta(N)^A$ -equidistributed for some $A > B$, then $(g(n)\Gamma)_{n \leq N}$ is totally $\delta(N)^{A/B}$ -equidistributed.*

With these preparations in place, we now turn towards the main result of this section.

Proposition 8.11 (equidistribution in sparse families of linear subsequences). *Let N be a large positive parameter, let $d \geq 0$, and let $K' > 0$ and $K > \max(2K', 2)$. Suppose that $1 \leq y' \leq (\log N)^{K'}$ and $(\log N)^K < y \leq N^\mu$, where $\mu = \mu(d) \in (0, 1)$ is sufficiently small depending on d . Let $\delta : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ be a function of N that satisfies $\log_5 N \ll \delta(N)^{-1}$ and $\delta(N)^{-B} \ll_B \log_2 N$ for all $B > 0$. Suppose that $(G/\Gamma, G_\bullet)$ is a nilmanifold together with a filtration G_\bullet of degree d and a $\delta(N)^{-1}$ -rational Malcev basis adapted to it. Let $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ be a polynomial sequence and suppose that the finite sequence $(g(n)\Gamma)_{n \leq N}$ is totally $\delta(N)^{E_1}$ -equidistributed in G/Γ for some $E_1 \geq 1$. Then there is some $c_1 \in (0, 1)$ depending on d and $\dim G$ such that the following assertion holds for all integers*

$$M \in [N^{1/2}, N/y^{1/2}] \tag{8-8}$$

provided that K is sufficiently large depending on d , and that $c_1 E_1 \geq 1$.

Given any sequence $(A_m)_{m \in \mathbb{N}}$ of integers satisfying $|A_m| \leq m$, write $g_m(n) := g(mn + A_m)$ and let \mathcal{B}_M denote the set of integers $m \in [M, 2M) \cap S([y', y])$ for which

$$(g_m(n)\Gamma)_{n \leq N/m}$$

fails to be totally $\delta(N)^{c_1 E_1}$ -equidistributed. Then

$$\#\mathcal{B}_M \ll \Psi(2M, [y', y])\delta(N)^{c_1 E_1}.$$

Proof. Let M be a fixed integer in the range (8-8) and let $c_1 > 0$ to be determined in the course of the proof. Suppose that $E_1 > 1/c_1$. It follows from Lemma 8.10 that for every $m \in \mathcal{B}_M$, the sequence $(g_m(n)\Gamma)_{n \leq N/m}$ fails to be $\delta(N)^{c_1 E_1 B}$ -equidistributed on G/Γ for some $1 \leq B \ll_{d, \dim G} 1$. By Lemma 8.9, there is a nontrivial horizontal character $\eta_m : G \rightarrow \mathbb{R}/\mathbb{Z}$ of magnitude $|\eta_m| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)}$ such that

$$\|\eta_m \circ g_m\|_{C^\infty[N/M]} \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)}. \tag{8-9}$$

For each nontrivial horizontal character $\eta : G \rightarrow \mathbb{R}/\mathbb{Z}$ we define the set

$$\mathcal{M}_\eta = \{m \in \mathcal{B}_M : \eta_m = \eta\}.$$

Note that this set is empty unless $|\eta| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)}$. Suppose that

$$\#\mathcal{B}_M \geq \Psi(2M, [y', y])\delta(N)^{c_1 E_1}.$$

Since there are only $M^{O(1)}$ horizontal characters of modulus bounded by M , it follows from the pigeonhole principle that there is some η of modulus $|\eta| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)}$ such that

$$\#\mathcal{M}_\eta \geq \Psi(2M, [y', y])\delta(N)^{c_1 E_1 C}$$

for some $C \asymp_{d, \dim G} 1$. Suppose

$$\eta \circ g(n) = \beta_d n^d + \dots + \beta_1 n + \beta_0.$$

Then

$$\eta \circ g_m(n) = \eta \circ g(mn + A_m) = \alpha_d^{(m)} n^d + \dots + \alpha_1^{(m)} n + \alpha_0^{(m)},$$

where

$$\alpha_j^{(m)} = m^j \sum_{i=j}^d \binom{i}{j} A_m^{i-j} \beta_i \quad (0 \leq j \leq d). \tag{8-10}$$

The bound (8-9) on the smoothness norm asserts that

$$\sup_{1 \leq j \leq d} \frac{N^j}{M^j} \|\alpha_j^{(m)}\| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)},$$

which by downwards induction combined with (8-10) implies

$$\sup_{1 \leq j \leq d} \frac{N^j}{M^j} \|\beta_j m^j\| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)}.$$

Hence,

$$\|\beta_j m^j\| \ll \delta(N)^{-O_{d, \dim G}(c_1 E_1)} (M/N)^j \quad (1 \leq j \leq d)$$

for every $m \in \mathcal{M}_\eta$.

In view of the lower bound on $\#\mathcal{M}_\eta$, we seek to apply Theorem 7.1 with $k = j \leq d$ and with the cut-off parameter N in the theorem replaced by M . Observe that all assumptions in Theorem 7.1 about the relation between the parameters N , y , y' , as well as those on the function δ and the size of ε are sufficiently flexible to allow for N to be replaced by any M with $\log M \asymp \log N$. Note that $N/M > y^{1/2} \geq (\log N)^{K/2} \gg (\log M)^{K/2}$, which allows us to choose

$$\varepsilon = \delta(N)^{-O_{d,\dim G}(c_1 E_1)} (M/N)^j < \delta(N)^{-O_{d,\dim G}(c_1 E_1)} y^{-j/2} < \delta(N)^{-O_{d,\dim G}(c_1 E_1)} (\log M)^{-jK/2}$$

and $C_1 = jK/2$ in the application of Theorem 7.1. Our assumptions on $\delta(N)$ imply that $\varepsilon < \delta^*/2$ for any δ^* of the form $\delta(N)^{c_1 E_1 C}$ with $C \asymp_{d,\dim G} 1$ as soon as N is sufficiently large. Hence, it follows from Theorem 7.1 that there exists an integer $1 \leq q_j \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$ such that

$$\|q_j \beta_j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} (M/N)^{-j} M^{-j} = \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}.$$

Thus,

$$\beta_j = \frac{a_j}{\kappa_j} + \tilde{\beta}_j, \tag{8-11}$$

where $\kappa_j \mid q_j$, $\gcd(a_j, \kappa_j) = 1$ and

$$0 \leq \tilde{\beta}_j \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}.$$

Hence,

$$\|\kappa_j \beta_j\| \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} N^{-j}. \tag{8-12}$$

Let $\kappa = \text{lcm}(\kappa_1, \dots, \kappa_d)$ and set $\tilde{\eta} = \kappa \eta$. We proceed in a similar fashion as in [Green and Tao 2012a, Section 3]: The above implies that

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)} n/N \quad (n \leq N),$$

which is small provided n is not too large. Indeed, if $N' = \delta(N)^{c_1 E_1 C''} N$ for some sufficiently large constant $C'' \geq 1$ depending only on d and $\dim G$, and if $n \in \{1, \dots, N'\}$, then

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \leq \frac{1}{10}.$$

Let $\chi : \mathbb{R}/\mathbb{Z} \rightarrow [-1, 1]$ be a function of bounded Lipschitz norm that equals 1 on $[-\frac{1}{10}, \frac{1}{10}]$ and satisfies $\int_{\mathbb{R}/\mathbb{Z}} \chi(t) dt = 0$. Then, by setting $F := \chi \circ \tilde{\eta}$, we obtain a Lipschitz function $F : G/\Gamma \rightarrow [-1, 1]$ that satisfies $\int_{G/\Gamma} F = 0$ and $\|F\|_{\text{Lip}} \ll \delta(N)^{-O_{d,\dim G}(c_1 E_1)}$. By choosing c_1 sufficiently small depending on d and $\dim G$ we may ensure that

$$\|F\|_{\text{Lip}} < \delta(N)^{-E_1}$$

and, moreover, that

$$N' > \delta(N)^{E_1} N.$$

This choice of N' , F and c_1 implies that, for all sufficiently large values of N , we have

$$\left| \frac{1}{N'} \sum_{1 \leq n \leq N'} F(g(n)\Gamma) \right| = 1 > \delta(N)^{E_1} \|F\|_{\text{Lip}},$$

which contradicts our assumption that $(g(n)\Gamma)_{n \leq N}$ is totally $\delta(N)^{E_1}$ -equidistributed. \square

9. Noncorrelation with equidistributed nilsequences

In this section we prove Proposition 8.8, that is to say we show that the function $m \mapsto \mathbf{1}_{S([y',y])}(\tilde{W}m + A')$ and its weighted version $m \mapsto h_{[y',y]}^{(W,A)}(qm + a)$ are orthogonal to highly equidistributed nilsequences for $w(N)$ -smooth values of q and $0 \leq a < q$.

Both of these functions are W -tricked versions of *sparse multiplicative functions*. Building on a Montgomery–Vaughan-type decomposition, which allows one to replace twisted sums of multiplicative functions by bilinear expressions, it was proved in [Matthiesen 2018] that W -tricked and centralised versions of “dense” *multiplicative functions* are orthogonal to nilsequences. We shall show that this approach can be extended to the sparse setting we are looking at here in order to establish Proposition 8.8.

9.1. Removing the weight. We will prove Proposition 8.8 by a sequence of reductions. Recall that

$$h_{[y',y]}^{(W,A)}(m) = \frac{\phi(W)}{W} h_{[y',y]}(Wm + A) = \frac{\phi(W)}{W} \frac{N^\alpha (Wm + A)^{1-\alpha}}{\alpha \Psi(N, [y', y])} \mathbf{1}_{S([y',y])}(Wm + A),$$

where $\gcd(A, W) = 1$. Our first step is to remove the weight and reduce the correlation estimate for $h_{[y',y]}^{(W,A)}(qm + a)$ to one that only involves the characteristic function $\mathbf{1}_{S([y',y])}(\tilde{W}m + A')$, where $\tilde{W} = Wq$ and $A' = Wa + A$.

Lemma 9.1 (removing the weight). *Let $\tilde{W} = Wq$ and $A' = Wa + A$ and define for any $x \in \mathbb{N}$ the quantity*

$$T(x) := \sum_{N/\tilde{W} < m \leq x} \mathbf{1}_{S([y',y])}(\tilde{W}m + A') F(g(m)\Gamma).$$

Then the conclusion (8-3) of Proposition 8.8 holds provided that

$$T(x) \ll_{d, \dim G, \|F\|_{\text{Lip}}, E_1} \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \tag{9-1}$$

for all $x \in (N/\tilde{W}, 2N/\tilde{W}]$.

Proof. To start with, recall that

$$\phi(\tilde{W})^{-1} = (\phi(W)q)^{-1} \prod_{p \mid q, p > w(N)} (1 - p^{-1})^{-1} \asymp (\phi(W)q)^{-1}$$

by (5-12) and note that

$$\frac{N^\alpha (\tilde{W}m + A')^{1-\alpha}}{\alpha N} \asymp 1$$

for $N/\tilde{W} < m < (N + N_1)/\tilde{W}$. With this in mind, partial summation shows that

$$\begin{aligned} & \frac{\tilde{W}}{N} \frac{\phi(W)}{W} \sum_{N/\tilde{W} < m < (N+N_1)/\tilde{W}} h_{[y',y]}(\tilde{W}m + A') F(g(m)\Gamma) \\ &= \frac{1}{\Psi(N, [y', y])} \frac{\tilde{W}\phi(W)}{W} \sum_{N/\tilde{W} < m < (N+N_1)/\tilde{W}} \frac{N^\alpha (\tilde{W}m + A')^{1-\alpha}}{\alpha N} \mathbf{1}_{S([y',y])}(\tilde{W}m + A') F(g(m)\Gamma) \end{aligned}$$

$$\begin{aligned}
 &\ll \frac{|T((N + N_1)/\tilde{W})|\phi(\tilde{W})}{\Psi(N, [y', y])} \\
 &\quad + \frac{\phi(\tilde{W})}{\Psi(N, [y', y])} \sum_{N/\tilde{W} < m < (N+N_1)/\tilde{W}} \frac{|T(m)|}{\alpha} \left| \left(\frac{\tilde{W}(m+1) + A'}{N} \right)^{1-\alpha} - \left(\frac{\tilde{W}m + A'}{N} \right)^{1-\alpha} \right| \\
 &\ll \frac{\phi(\tilde{W})}{\Psi(N, [y', y])} \left\{ \left| T\left(\frac{N + N_1}{\tilde{W}} \right) \right| + \sum_{N/\tilde{W} < m < (N+N_1)/\tilde{W}} \frac{|T(m)|(1-\alpha)|}{\alpha N} \left| \int_{\tilde{W}(m+1)+A'}^{\tilde{W}m+A'} (N/t)^\alpha dt \right| \right\} \\
 &\ll \frac{\phi(\tilde{W})}{\Psi(N, [y', y])} \left\{ \left| T\left(\frac{N + N_1}{\tilde{W}} \right) \right| + \frac{\tilde{W}}{N} \sum_{N/\tilde{W} < m < (N+N_1)/\tilde{W}} |T(m)| \right\} \\
 &\ll_{d, \dim G, \|F\|_{\text{Lip}}, E_1} \delta(N)^{E_1},
 \end{aligned}$$

provided that the stated bounds on $T(x)$ hold. □

9.2. Montgomery–Vaughan-type reduction to a Type II estimate. Our proof strategy for establishing (9-1) is to employ a Montgomery–Vaughan-type decomposition which replaces the given one-parameter correlation by a bilinear sum and eventually allows us to reduce (9-1) to a noncorrelation estimate of the von Mangoldt function with nilsequences. Since the parameter y can be very small and since the correlation involving the von Mangoldt function will have length y^c in the end, this approach requires a careful choice of the cut-off parameter $w(N)$ in the W -trick and the parameter $\delta(N)$ that controls the level of equidistribution in the nilsequence to ensure that uniform noncorrelation estimates for the von Mangoldt function, valid over the whole range produced by the Montgomery–Vaughan decomposition, can be deduced.

Proposition 9.2 (reduction to a bilinear correlation). *The bound (9-1) holds for all $x \in (N/\tilde{W}, 2N/\tilde{W}]$ provided that for any sequence $(A_n)_{n \in \mathbb{N}}$ with $|A_n| \leq n$, for any sequence $(A'_n)_{n \in \mathbb{N}}$ of integers $0 < A'_n < \tilde{W}$ coprime to \tilde{W} and for any sufficiently large $E_0 > 1$, we have*

$$\frac{1}{\log N} \left| \sum_{1 < n \leq N+N_1} \sum_{\substack{\tilde{W}m + A'_n \in [y', y] \\ N < n(\tilde{W}m + A'_n) \leq N+N_1}} \mathbf{1}_{S([y', y])}(n) \Lambda(\tilde{W}m + A'_n) F(g(mn + A_n)\Gamma) \right| \ll \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})q}$$

for all $N_1 \in (0, N]$ and all $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ such that $(g(n)\Gamma)_{n \leq N}$ is totally $\delta(N)^{E_0}$ -equidistributed. The implied constant is allowed to depend on $d, \dim G, \|F\|_{\text{Lip}}$ and E_1 .

Proposition 9.2 will be deduced from the following simple lemma, which is inspired by a bound from [Montgomery and Vaughan 1977], but does not involve a second moment.

Lemma 9.3. *Suppose that $f : \mathbb{N} \rightarrow \mathbb{C}$, let N be a positive parameter and let $\mathcal{S}(N) \subset [N/2, N] \cap \mathbb{N}$ denote a set. Then*

$$\sum_{n \in \mathcal{S}(N)} f(n) \leq \frac{1}{\log N} \sum_{n \in \mathcal{S}(N)} |f(n)| + \frac{1}{\log N} \left| \sum_{nm \in \mathcal{S}(N)} f(nm) \Lambda(m) \right|.$$

Proof. This is an immediate consequence of the bound

$$\sum_{N/2 \leq n \leq N} \log(N/n) f(n) \leq \log 2 \sum_{N/2 \leq n \leq N} |f(n)|. \quad \square$$

Note that we have

$$T(x) = \sum_{N/\tilde{W} < m \leq x} \mathbf{1}_{S([y', y])}(\tilde{W}m + A') F(g(m)\Gamma) = \sum_{\substack{N < n \leq x\tilde{W} \\ n \equiv A' \pmod{\tilde{W}}}} \mathbf{1}_{S([y', y])}(n) F\left(g\left(\frac{n - A'}{\tilde{W}}\right)\Gamma\right).$$

Writing $N_1 := x\tilde{W} - N \in (0, N]$ and applying Lemma 9.3 to this expression yields

$$T(x) \ll \frac{1}{\log N} \left| \sum_{\substack{N < mn \leq N+N_1 \\ mn \equiv A' \pmod{\tilde{W}}}} \mathbf{1}_{S([y', y])}(mn) \Lambda(m) F\left(g\left(\frac{nm - A'}{\tilde{W}}\right)\Gamma\right) \right| + O\left(\frac{\Psi(N, [y', y]; \tilde{W}, A')}{\log N}\right). \quad (9-2)$$

If $A'_n \in \{0, \tilde{W} - 1\}$ is such that $nA'_n \equiv A' \pmod{\tilde{W}}$, we may expand the congruence condition on m and replace m by $\tilde{W}m + A'_n$. With this choice of A'_n we have

$$\frac{n(\tilde{W}m + A'_n) - A'}{\tilde{W}} = nm + A_n$$

for some integer A_n with $|A_n| \leq n$. Hence the first term in the bound on $T(x)$ may be rewritten as

$$\mathcal{M}(x) := \frac{1}{\log N} \left| \sum_{1 \leq n \leq N+N_1} \sum_{N/n < \tilde{W}m + A'_n \leq (N+N_1)/n} \mathbf{1}_{S([y', y])}(n(\tilde{W}m + A'_n)) \Lambda(\tilde{W}m + A'_n) F(g(nm + A_n)\Gamma) \right|.$$

The error term in (9-2) is negligible in view of Theorem 5.1, Lemma 3.3 and (3-3), provided that $\log^{-1} N \ll_{E_1} \delta(N)^{E_1}$ which holds for the choice of $\delta(N)$ for Proposition 8.8. Thus (9-1) holds provided we can show that, under the assumptions of Proposition 8.8, the bound

$$\mathcal{M}(x) \ll_{d, \|F\|_{\text{Lip}}} \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \quad (9-3)$$

holds for all $N_1 \in (0, N]$. To complete the proof of Proposition 9.2, it remains to show that we can replace the function $\Lambda(m)\mathbf{1}_{S([y', y])}(m)$ by $\Lambda(m)\mathbf{1}_{[y', y]}(m)$ in the condition (9-3). This is the content of the following lemma.

Lemma 9.4 (removing large prime powers). *We have*

$$\frac{1}{\log N} \sum_{\substack{N < nm \leq N+N_1 \\ nm \equiv A' \pmod{\tilde{W}} \\ \Omega(m) \geq 2, m > y}} \mathbf{1}_{S([y', y])}(nm) \Lambda(m) \ll_B \delta(N)^B \frac{\Psi(N, [y', y])}{\phi(\tilde{W})}.$$

Proof. Theorem 5.1 and Lemma 3.4(i) show that the left-hand side above is bounded by

$$\begin{aligned} &\ll \frac{\log y}{\log N} \sum_{p \leq y} \sum_{k \geq \max(2, (\log y)/\log p)} \max_{(A'', \tilde{W})=1} \Psi\left(\frac{N}{p^k}, [y', y]; \tilde{W}, A''\right) \\ &\ll \frac{\log y}{\log N} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \sum_{p \leq y} \sum_{k \geq \max(2, (\log y)/\log p)} p^{-\alpha k} + O\left(\frac{\log y}{\log N} \frac{y^2}{\log y}\right), \end{aligned}$$

where the error term trivially bounds the contribution of those choices of p^k for which $N/y < p^k < (N + N_1)/y'$, i.e., to which Lemma 3.4(i) does not apply. This error term is acceptable since $y < N^\eta$ for some sufficiently small $\eta > 0$. The sum over primes in the main term satisfies

$$\begin{aligned} \sum_{p \leq y} \sum_{k \geq \max(2, (\log y)/\log p)} p^{-\alpha k} &\ll \sum_{p \leq y^{1/2}} y^{-\alpha} + \sum_{\sqrt{y} < p \leq y} p^{-2\alpha} \ll \frac{y^{1/2-\alpha}}{\log y} + (y^{1/2})^{-2\alpha+1} \\ &\ll y^{-1/2} u \log u \ll_\varepsilon \frac{\log N}{\log y} (\log N)^{-K/2+\varepsilon} \ll_B \delta(N)^B \frac{\log N}{\log y} \end{aligned}$$

for all $B > 0$, since $\alpha > \frac{1}{2}$ and $\delta(N) = \exp(-C_0\sqrt{\log_4 N})$. □

9.3. Explicit bounds for the correlation between Λ and nilsequences. In view of the inner sum in the bilinear expression in Proposition 9.2, we may reduce the problem of establishing that bound to a noncorrelation estimate for the von Mangoldt function with equidistributed nilsequences. For this purpose, we require explicit bounds for correlations of length ξ of the W -tricked von Mangoldt function with nilsequences. A specific requirement on these bounds is that they work with the same W -trick (determined by $w(N)$ and independent of ξ) and are uniform over a large range of ξ . Taking into account that y can be as small as $(\log N)^K$, we shall prove a result that is applicable to $\xi \in [\log N, N]$ if $w(N)$ is suitably chosen (see Remarks 9.6(ii) below).

Theorem 9.5. *Let $N_0 > 2$ be a large constant and $N > N_0$ a parameter. Let $x_0, w, \delta, \kappa : \mathbb{N}_{>N_0} \rightarrow \mathbb{R}_{>0}$ be functions that are defined for all sufficiently large integers and satisfy the relations $\kappa(N) \geq 1$, $1 \leq w(N) \leq \frac{1}{2} \log_2 x_0(N)$ and $\kappa(N)^2 w(N)^{-1/\kappa(N)} \ll_B \delta(N)^B$ for all $N > N_0$ and all $B \geq 1$. Suppose that $w(N) \rightarrow \infty$ as $N \rightarrow \infty$.*

Suppose that G/Γ is a filtered nilmanifold of dimension $\dim G \leq \kappa(N)$ and complexity at most $\delta(N)^{-1}$, let $d \geq 0$ denote its degree, and let $g \in \text{poly}(\mathbb{Z}, G_\bullet)$. Let $1 \leq q \leq (\log x_0(N))^E$, $1 \leq E \ll 1$, be an integer, let $W = P(w(N))$ and write $\tilde{W} = Wq$. If $0 < A' < \tilde{W}$ is an integer such that $(\tilde{W}, A') = 1$, then

$$\frac{\tilde{W}}{\xi} \sum_{n \leq \xi/\tilde{W}} \left(\frac{\phi(\tilde{W})}{\tilde{W}} \Delta(\tilde{W}n + A') - 1 \right) F(g(n)\Gamma) \ll_{d,B} (1 + \|F\|_{\text{Lip}}) \delta(N)^B$$

for all $\xi \in [x_0(N), N]$, all $B \geq 1$, and for all Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$.

Remarks 9.6. (i) Note that $\xi/\tilde{W} = \xi^{1+o(1)}$ by the definition of W and the conditions on $w(N)$, $\delta(N)$ and q . More precisely, we have $\xi \geq x_0(N)$, $\tilde{W} = Wq$, $W = \exp(w(N) + o(1)) = (\log x_0(N))^{1/2+o(1)}$, and $q \leq (\log x_0(N))^{O(1)}$.

(ii) The choices $x_0(N) := \log N$, $w(N) := \frac{1}{2} \log_3 N$, $\kappa(N) = (\log_5 N)^C$ with $1 \leq C = O(1)$ and $\delta(N) = \exp(-C_0\sqrt{\log_4 N})$ with $1 \leq C_0 \leq (\log_4 N)^{1/4}$ and $N > N_0$ for sufficiently large N_0 are permissible in the theorem above.

(iii) Observe that the parameter choices in (ii) are consistent with the assumptions of Proposition 8.11 as well as the assumptions of Theorem 7.1.

Proof. This proof is a small modification of the proof of [Matthiesen 2018, Lemma 9.5] given in the appendix to that paper. The starting point of that proof is the following decomposition of the von Mangoldt function as $\Lambda = \Lambda^b + \Lambda^\sharp$. Let $\text{id}_{\mathbb{R}}(x) = x$ denote the identity function and let $\chi^b + \chi^\sharp = \text{id}_{\mathbb{R}}$ be a smooth decomposition with the property that $\text{supp}(\chi^\sharp) \subset (-1, 1)$ and $\text{supp}(\chi^b) \subset \mathbb{R} \setminus [-\frac{1}{2}, \frac{1}{2}]$. Then, for any $\gamma \in (0, 1)$,

$$\frac{\phi(\tilde{W})}{\tilde{W}} \Lambda(\tilde{W}n + A') - 1 = \frac{\phi(\tilde{W})}{\tilde{W}} \Lambda^b(\tilde{W}n + A') + \left(\frac{\phi(\tilde{W})}{\tilde{W}} \Lambda^\sharp(\tilde{W}n + A') - 1 \right),$$

where, see [Green and Tao 2010, (12.2)],

$$\Lambda^\sharp(n) = -\log \xi^\gamma \sum_{d|n} \mu(d) \chi^\sharp \left(\frac{\log d}{\log \xi^\gamma} \right) \quad (|t| \geq 1 \Rightarrow \chi^\sharp(t) = 0)$$

is a truncated divisor sum, where

$$\Lambda^b(n) = -\log \xi^\gamma \sum_{d|n} \mu(d) \chi^b \left(\frac{\log d}{\log \xi^\gamma} \right) \quad (|t| \leq \frac{1}{2} \Rightarrow \chi^b(t) = 0)$$

is an average of $\mu(d)$ running over large divisors of n .

It follows as in [Green and Tao 2010, Section 12] and [Matthiesen 2018, Appendix A] from the orthogonality of the Möbius function with nilsequences that

$$\frac{\tilde{W}}{\xi} \sum_{n \leq \xi/\tilde{W}} \frac{\phi(\tilde{W})}{\tilde{W}} \Lambda^b(\tilde{W}n + A') F(g(n)\Gamma) \ll_{\|F\|_{\text{Lip}}, G/\Gamma, B} (\log \xi)^{-B} \quad (B > 0).$$

Here, it is important that $\tilde{W} \leq (\log x_0(N))^{O(1)} \leq (\log \xi)^{O(1)}$.

Concerning the contribution from Λ^\sharp , define $\lambda^\sharp : \mathbb{N} \rightarrow \mathbb{R}$,

$$\lambda^\sharp(n) := \frac{\phi(\tilde{W})}{\tilde{W}} \Lambda^\sharp(\tilde{W}n + A') - 1.$$

Since \mathcal{X} is $\delta(N)^{-1}$ -rational, [Matthiesen 2018, Lemmas A.2 and A.3] imply that the following bound holds with $m = 2^d \dim G$ and for every $\varepsilon > 0$

$$\begin{aligned} \frac{\tilde{W}}{\xi} \sum_{n \leq \xi/\tilde{W}} \lambda^\sharp(n) F(g(n)\Gamma) &\ll \varepsilon(1 + \|F\|_{\text{Lip}}) + \|\lambda^\sharp\|_{U^{d+1}[\xi/\tilde{W}]}(m/\varepsilon)^{2m} \delta(N)^{-O(m)} \\ &\ll \varepsilon(1 + \|F\|_{\text{Lip}}) + \frac{m^{2m} \|\lambda^\sharp\|_{U^{d+1}[\xi/\tilde{W}]}}{(\varepsilon^2 \delta(N)^{O(1)})^m}. \end{aligned} \tag{9-4}$$

We shall show below that $\|\lambda^\sharp\|_{U^{d+1}[\xi/\tilde{W}]} \ll_d w(N)^{-1/2^{d+1}}$. Choosing $\varepsilon = \delta(N)^B$ and recalling the assumption that $\kappa(N)^2 w(N)^{-1/\kappa(N)} \ll_B \delta(N)^B$ and that $m = 2^d \dim G \leq 2^d \kappa(N)$, it follows that the above is bounded by

$$\ll_{d,B} (1 + \|F\|_{\text{Lip}}) \delta(N)^B + \delta(N)^{O_d(B\kappa(N))} \ll_{d,B} (1 + \|F\|_{\text{Lip}}) \delta(N)^B,$$

as required.

The uniformity norm $\|\lambda^\sharp\|_{U^{d+1}[\xi/\tilde{W}]}$, can, as in [Matthiesen 2018, Appendix A], be estimated using [Green and Tao 2010, Theorem D.3] since $w(N) \leq \frac{1}{2} \log_2 x_0(N)$ is sufficiently small (see the ‘‘important convention’’ in [loc. cit., Section 5]). In our case, the system of forms takes the shape

$$\Psi_{\mathcal{B}}(n, \mathbf{h}) = (\tilde{W}(n + \boldsymbol{\omega} \cdot \mathbf{h}) + A')_{\boldsymbol{\omega} \in \mathcal{B}}, \quad (n, \mathbf{h}) \in \mathbb{Z} \times \mathbb{Z}^{d+1},$$

where $\mathcal{B} \subset \{0, 1\}^{d+1}$ is any nonempty subset. The corresponding set $\mathcal{P}_{\Psi_{\mathcal{B}}}$ of exceptional primes consists of those primes dividing $\tilde{W} = Wq$, i.e., $|\mathcal{P}_{\Psi_{\mathcal{B}}}| = \pi(w(N)) \ll \log_2 \xi / \log w(N)$. We further have

$$\prod_{p \notin \mathcal{P}_{\Psi_{\mathcal{B}}}} \beta_p^{(\mathcal{B})} = 1 + O_d\left(\frac{1}{w(N)}\right) \quad \text{and} \quad \prod_{p \in \mathcal{P}_{\Psi_{\mathcal{B}}}} \beta_p^{(\mathcal{B})} = \left(\frac{\tilde{W}}{\phi(\tilde{W})}\right)^{|\mathcal{B}|}.$$

Provided the constant γ is chosen sufficiently small, it follows from [Green and Tao 2010, Theorem D.3] applied with $K_z = \{(n, \mathbf{h}) : 0 < n + \boldsymbol{\omega} \cdot \mathbf{h} \leq z \text{ for all } \boldsymbol{\omega} \in \{0, 1\}^{d+1}\}$, where $z = \xi/\tilde{W}$, that

$$\begin{aligned} \|\lambda^\sharp\|_{U^{d+1}[z]}^{2^{d+1}} &= \frac{\text{vol}(K_z)}{z^{d+2}} \sum_{\mathcal{B} \subseteq \{0, 1\}^{d+1}} (-1)^{|\mathcal{B}|} \prod_{p \notin \mathcal{P}_{\Psi_{\mathcal{B}}}} \beta_p^{(\mathcal{B})} + O_d((\log \xi^\gamma)^{-1/20} \exp(O_d(|\mathcal{P}_{\Psi_{\mathcal{B}}|}))) \\ &\ll_d \frac{\text{vol}(K_z)}{z^{d+2}} \frac{1}{w(N)} + \exp\left(-\frac{\log_2 \xi}{20} + O_d\left(\frac{\log_2 \xi}{\log w(N)}\right)\right) \ll_d \frac{1}{w(N)}. \quad \square \end{aligned}$$

As it provides a different approach which could prove useful for future generalisations, we include a second, alternative, proof for the special case of Theorem 9.5 in which $q = 1$. This proof is based on one of the main results of [Tao and Teräväinen 2023]. For this proof to work, $w(n)$ needs to be redefined as $w(n) = c' \log_3 N$ for some sufficiently small constant $c' > 0$.

Alternative proof for special case of Theorem 9.5. Assume for simplicity the choice of parameters from Remarks 9.6(ii) except for $w(N) = \frac{1}{2} \log_3 N$, which we replace by $w(n) = c' \log_3 N$ for some small $c' > 0$. Further, let $q = 1$ so that $\tilde{W} = W$ and $A' = A$.

Define $\lambda_{W,A} = W^{-1} \phi(W) \Lambda(W \cdot + A) - 1$. By the proof of [loc. cit., Corollary 1.5] (see the end of Section 8 in that paper), we have

$$\|\lambda_{W,A}\|_{U^k[(\xi-A)/W]} \ll (\log_2 \xi)^{-c} \tag{9-5}$$

provided that the constant $c' > 0$ in the definition of $w(N)$ is sufficiently small to ensure that the estimate [loc. cit., (8.19)] holds with $W = P(w(N))$ and $N = x_0(N)$.

Proceeding as in the proof of Theorem 9.5 above, we use [Matthiesen 2018, Lemmas A.2 and A.3] which show that the following bound holds with $m = 2^d \dim G$ and for every $\varepsilon > 0$:

$$\frac{W}{\xi} \sum_{n \leq (\xi-A)/W} \lambda_{W,A}(n) F(g(n)\Gamma) \ll \varepsilon (1 + \|F\|_{\text{Lip}}) + \|\lambda_{W,A}\|_{U^{d+1}[(\xi-A)/W]} (m/\varepsilon)^{2m} \delta(N)^{-O(m)}.$$

Choosing $\varepsilon = \delta(N)^B$ and recalling that $\log_2 \xi \gg \log_2 x_0(N) \gg \log_3 N$, the above is seen to be bounded by

$$\ll_d (1 + \|F\|_{\text{Lip}}) \delta(N)^B + \frac{\kappa(N)^2 (\log_2 \xi)^{-c}}{\delta(N)^{O_d(B\kappa(N))}} \ll_d (1 + \|F\|_{\text{Lip}}) \delta(N)^B,$$

where the constants c and the final implied constant depend on the exponent in the bound (9-5) from [Tao and Teräväinen 2023, Section 8] as well as on the degree d of the filtration. \square

9.4. Conclusion of the proof of Proposition 8.8. In view of the reductions carried out in the previous subsections it remains to show that the condition of Proposition 9.2 is in fact valid in order to complete the proof of the equidistributed noncorrelation estimate stated in Proposition 8.8. Hence, the proof of Proposition 8.8 is complete once we have established the following lemma.

Lemma 9.7. *Let $E_1 \geq 1$. Under the assumptions of Proposition 8.8 and with the notation $\tilde{W} = Wq$, we have*

$$\frac{1}{\log N} \left| \sum_{n \leq N} \sum_{\substack{\tilde{W}m + A'_n \in \\ [y', \min(y, N/n)]}} \mathbf{1}_{S([y', y])}(n) \Lambda(\tilde{W}m + A'_n) F(g(mn + A_n)\Gamma) \right| \ll \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})}$$

for all sequences $(A_n)_{n \in \mathbb{N}}$, $(A'_n)_{n \in \mathbb{N}}$ of integers such that $|A_n| \leq n$ and $\gcd(A'_n, \tilde{W}) = 1$, and for all $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ such that the finite sequence $(g(n)\Gamma)_{n \leq N}$ is totally $\delta(N)^{E_0}$ -equidistributed for some sufficiently large $E_0 > 1$. The implied constant may depend on the degree d of G_\bullet , $\dim G$, $\|F\|_{\text{Lip}}$ and E_1 .

Proof. Splitting the summation range of the outer sum into three intervals and abbreviating

$$G(m, n) = \mathbf{1}_{S([y', y])}(n) \Lambda(\tilde{W}m + A'_n) F(g(mn + A_n)\Gamma),$$

we obtain

$$\begin{aligned} \sum_{n \leq N} \sum_{\tilde{W}m + A'_n \in [y', \min(y, N/n)]} G(m, n) &= \sum_{n < \delta^{2E_1} N/y} \sum_{y' \leq \tilde{W}m + A'_n \leq y} G(m, n) \\ &+ \sum_{\delta^{2E_1} N/y < n \leq N/y^{2/3}} \sum_{y' \leq \tilde{W}m + A'_n \leq \min(N/n, y)} G(m, n) \\ &+ \sum_{0 \leq k \leq K} \sum_{N/y^{2/3} < n \leq N/(y'2^k)} \sum_{\tilde{W}m + A'_n \in [y'2^k, \min(y'2^{k+1}, N/n)]} G(m, n), \end{aligned} \tag{9-6}$$

where $K = \lceil \log(y^{2/3}/y') / \log 2 \rceil$ and in particular $y'2^K \asymp y^{2/3}$.

The decomposition above has been chosen in such a way that the contributions from the initial and final segment are negligible while in the middle segment the summation range of the inner sum is guaranteed to be sufficiently long and $\min(y, N/n)$ is not too small compared to N/n . More precisely, by Lemma 3.4(i), the final term is trivially bounded above by

$$\begin{aligned} \sum_{0 \leq k \leq K} \frac{y'2^k}{\phi(\tilde{W})} \Psi\left(\frac{N}{y'2^k}, [y', y]\right) &\ll \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \sum_{0 \leq k \leq K} (y'2^k)^{(1-\alpha)} \ll \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} (y'2^K)^{(1-\alpha)} \\ &\ll y^{2(1-\alpha)/3} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \ll (\log N)^{2/3} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})}, \end{aligned}$$

which is acceptable in view of the $(\log N)^{-1}$ factor in the statement.

Similarly, the first term is trivially bounded above by

$$\frac{y}{\phi(\tilde{W})} \Psi\left(\delta(N)^{2E_1} \frac{N}{y}, [y', y]\right) \ll \delta(N)^{2E_1\alpha} y^{1-\alpha} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})} \ll (\log N) \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})},$$

since $\alpha > \frac{1}{2}$, which is also acceptable.

It remains to analyse the middle range and we start by observing that the summation range of the inner sum over $\tilde{W}m + A'_n$ is always sufficiently long, that is, of length $\gg \delta(N)^{E_1} m_1$ if m_1 denotes the upper endpoint of the range. In fact, $m_1 = \min(y, N/n) \geq y^{2/3} > y^{1/2} > y'$, which shows that $|\min(y, N/n) - y'| \sim \min(y, N/n)$ and moreover $y' = o(\delta(N)^B \min(y, N/n)/\phi(\tilde{W}))$ for any $B \asymp 1$.

In order to apply Proposition 8.11 on the equidistribution of linear subsequences of an equidistributed nilsequence, we dyadically decompose the sum over n in the middle range of (9-6) into $O(\log N)$ intervals $(M_j, 2M_j]$, where $\delta(N)^{2E_1} N/y \leq M_j < N/y^{2/3}$, and one potentially shorter interval. Note that $N^{1/2} < Ny^{-2} < N\delta(N)^{2E_1} y^{-1} \leq M_j$, assuming that $\eta < \frac{1}{4}$ (so that $y < N^{1/4}$) and that N is sufficiently large. For each M_j as above consider the sum

$$\sum_{\substack{n \sim M_j \\ n \leq N/y^{2/3}}} \mathbf{1}_{S([y', y])}(n) \sum_{\tilde{W}m + A'_n \in [y', \min(y, N/n)]} \Lambda(\tilde{W}m + A'_n) F(g(mn + A_n)\Gamma). \tag{9-7}$$

Let $E_2 > 1$ be a parameter to be chosen later. Then, by Proposition 8.11, all but $O(\delta(N)^{E_2} \Psi(M_j, [y', y]))$ of the finite sequences

$$(g_n(m)\Gamma)_{m \leq N/n}, \quad n \in (M_j, 2M_j] \cap S([y', y]),$$

where $g_n(m) := g(mn + A_n)$, are totally $\delta(N)^{E_2}$ -equidistributed in G/Γ provided that the parameter $E_0 > 1$ from our assumptions is sufficiently large. We bound the contribution to (9-7) from all exceptional values of $n \in (M_j, 2M_j] \cap S([y', y])$ trivially by

$$O(\delta(N)^{E_2} \Psi(M_j, [y', y])) \frac{\min(y, N/M_j)}{\phi(\tilde{W})}.$$

For all other values of $n \in (M_j, 2M_j] \in S([y', y])$, we seek to apply Theorem 9.5 with the choice of parameters given in Remarks 9.6(ii). The lower bound $M_j \geq \delta(N)^{2E_1} N/y$ implies that $y \geq \delta^{2E_1} N/M_j \geq \delta^{E_2/2} N/n$ for all $n \in (M_j, 2M_j]$ provided that $E_2 > 4E_1$. Thus $\min(N/n, y) \geq \delta^{E_2/2} N/n$ and the sequence

$$(g_n(m)\Gamma)_{m \leq \min(N/n, y)}$$

is totally $\delta^{E_2/2}$ -equidistributed in G/Γ whenever n is nonexceptional, and $\xi = \min(y, N/n)$ satisfies $\xi \in [\log N, N]$. Finally, the upper bound on y' implies that, for any $B \asymp 1$, $y' = o(\delta(N)^B \min(y, N/M_j)/\phi(q))$. For nonexceptional n as above, it thus follows from Theorem 9.5, applied with $B = E_2/4$, that

$$\begin{aligned} & \frac{\phi(\tilde{W})}{\min(y, N/M_j)} \sum_{\tilde{W}m + A'_n \in [y', \min(y, N/n)]} \Lambda(\tilde{W}m + A'_n) F(g(mn + A_n)\Gamma) \\ &= \sum_{\tilde{W}m + A'_n \in [y', \min(y, N/n)]} F(g(mn + A_n)\Gamma) + O_{d, E_1, \|F\|_{\text{Lip}}}(\delta(N)^{E_2/4}) \ll_{d, E_1, \|F\|_{\text{Lip}}} \delta(N)^{E_2/4}, \end{aligned}$$

where we used that $p < w(N)$ for every $p | q$.

Taking the contributions from both exceptional and nonexceptional $n \in (M_j, 2M_j] \in S([y', y])$ into account, the expression (9-7) is thus bounded by

$$\ll_{d, E_1, \|F\|_{\text{Lip}}} (\delta(N)^{E_2/4} + \delta(N)^{E_1}) \min(y, N/M_j) \frac{\Psi(M_j, [y', y])}{\phi(\tilde{W})}.$$

Summing over all j , the contribution of the middle segment to (9-6) is therefore bounded above by

$$\begin{aligned} &\ll_{d, E_1, \|F\|_{\text{Lip}}} (\delta(N)^{E_2/4} + \delta(N)^{E_1}) \sum_j \min(y, N/M_j) \frac{\Psi(M_j, [y', y])}{\phi(\tilde{W})} \\ &\ll_{d, E_1, \|F\|_{\text{Lip}}} \frac{\delta(N)^{E_1}}{\phi(\tilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y', y])}(n) \min(y, N/n). \end{aligned} \tag{9-8}$$

We shall now make use of the fact that the inner sum in the middle range is guaranteed to be sufficiently long in order to deduce that we make an $\delta(N)^{E_1}$ -saving on average on each inner sum provided E_0 and, hence, E_2 are sufficiently large. Reversing the steps in the Montgomery–Vaughan-type decomposition will then complete the proof.

Turning towards the details, recall that for any $n \sim M_j$ the interval $[y', \min(y, N/n)]$ has length $|\min(y, N/n) - y'| \sim \min(y, N/n)$, so that

$$\min(y, N/n) \ll |\min(y, N/n) - y'| \ll \sum_{y' \leq m \leq \min(y, N/n)} \Lambda(m) \quad (n \sim M_j),$$

by the prime number theorem. By combining this estimate with the bound (9-8), the contribution of the middle segment to (9-6) is seen to be bounded above by

$$\begin{aligned} &\ll \frac{\delta(N)^{E_1}}{\phi(\tilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y', y])}(n) \min(y, N/n) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\tilde{W})} \sum_j \sum_{n \sim M_j} \mathbf{1}_{S([y', y])}(n) \sum_{y' \leq m \leq \min(y, N/n)} \Lambda(m) \mathbf{1}_{S([y', y])}(m) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\tilde{W})} \sum_{\delta(N)^{2E_1} N/y \leq n < 2N/y^{2/3}} \sum_{y' \leq m \leq \min(y, N/n)} \Lambda(m) \mathbf{1}_{S([y', y])}(mn) \\ &\ll \frac{\delta(N)^{E_1}}{\phi(\tilde{W})} \sum_{n < 2N} \mathbf{1}_{S([y', y])}(n) \log n \ll (\log N) \delta(N)^{E_1} \frac{\Psi(N, [y', y])}{\phi(\tilde{W})}, \end{aligned}$$

provided that E_0 is sufficiently large (to ensure that $E_2 > 4E_1$), and where the implied constant may depend on $d, \dim G, E_1$ and $\|F\|_{\text{Lip}}$. The lemma follows when taking into account the $(\log N)^{-1}$ -factor in the statement. □

10. The proof of Theorem 1.3

We will use the transferred generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] combined with a simple majorising function in order deduce Theorem 1.3 from Theorem 1.1. There are

two choices of majorants for $h_{[y',y]}$ readily available in the case where $y \geq N^\varepsilon$ for any fixed $\varepsilon \in (0, 1)$. These can still be used for somewhat smaller values of y . We state both majorant constructions below.

10.1. GPY-type sieve majorant. Consider the following GPY-type sieve majorant for numbers free from prime factors $p < y'$:

$$\Lambda_{\chi,y'}(n) = \log y' \left(\sum_{\substack{d|n \\ d \leq y'}} \mu(d) \chi \left(\frac{\log d}{\log y'} \right) \right)^2,$$

where $\chi : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ is a smooth function with support in $[-1, 1]$ which takes the value $\chi(x) = 1$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$. In particular,

$$\Lambda_{\chi,y'}(n) = \log y' \left(\sum_{\substack{d|n \\ d \leq y'}} \mu(d) \chi \left(\frac{\log d}{\log y'} \right) \right)^2,$$

which implies that $\Lambda_{\chi,y'}(n) = \log y'$ for all integers n that are free from prime factors $p \leq y'$. Since $\alpha = 1 - O((\varepsilon \log N)^{-1} \log_2 N)$ for $y = N^\varepsilon$, Lemma 3.2 shows that $\zeta(\alpha, y) \asymp \log y \asymp \log N$ and $g_{P(y')} \asymp (\log y')^{-1}$. Thus, by (3-3) and Lemma 3.3, we have

$$h_{[y',N^\varepsilon]}(n) \leq \frac{(N/n)^{\alpha n}}{\Psi(N, [y', N^\varepsilon])} \ll \frac{N \log N}{N g_{P(y')}(\alpha) \zeta(\alpha, N^\varepsilon)} \ll \log y'.$$

This shows that

$$h_{[y',N^\varepsilon]}(n) \ll \Lambda_{\chi,y'}(n) \quad (N < n \leq 2N).$$

Hence, $\Lambda_{\chi,y'}(n)$ satisfies the majorisation property (1) from above. It follows from [Green and Tao 2010, Theorem D.3] that the W -tricked version of $\Lambda_{\chi,y'}(n)$ satisfies the *linear forms condition* as stated in [loc. cit., Definition 6.2]. We omit the proof here as it is essentially contained in [loc. cit., Appendix D] and only mention that it relies on the fact that $\Lambda_{\chi,y'}$ carries the structure of a truncate divisor sum.

10.2. The normalised characteristic function $\mathbf{1}_{S([y',2N])}$. An alternative majorant function in the dense setting is given by the function

$$v(n) = \frac{P(y')}{\phi(P(y'))} \mathbf{1}_{S([y',2N])}(n) \quad (n \leq 2N).$$

Since

$$\mathbf{1}_{S([y',2N])}(n) = \mathbf{1}_{(n, P(y'))=1}(n)$$

for $n \leq 2N$, this function corresponds to the *Cramér model* for the von Mangoldt function that was studied in [Tao and Teräväinen 2023]: we have

$$v(n) = \Lambda_{\text{Cramér},y'}(n) := \frac{P(y')}{\phi(P(y'))} \mathbf{1}_{(n, P(y'))=1}(n)$$

for all $n \leq 2N$.

Lemma 10.1. *Let $\varepsilon \in (0, 1)$. Then we have*

$$h_{[y',y]}(n) \ll \Lambda_{\text{Cramér},y'}(n)$$

uniformly for all $y > N^\varepsilon$, all $n \in (N, 2N]$ and all sufficiently large N .

Proof. Since $\mathbf{1}_{(n, P(y'))=1} = \mathbf{1}_{S([y', 2N])}(n)$ for $n \leq 2N$, it remains to note that

$$h_{[y', y]}(n) \leq \frac{n(N/n)^\alpha}{\Psi(N, [y', N^\varepsilon])} \ll \frac{N \log N}{Ng_{P(y')}(\alpha)\zeta(\alpha, N^\varepsilon)} \ll \log y' \asymp \frac{P(y')}{\phi(P(y'))},$$

when $n \sim N$ and $y > N^\varepsilon$. □

Let $2 < w(N) < y' = (\log N)^{K'}$, $W = P(w(N))$, $0 < A < W$ and $(A, W) = 1$. Then, by [Tao and Teräväinen 2023, Corollary 5.3], we have

$$\left\| \frac{\phi(W)}{W} \nu(W \cdot + A) - 1 \right\|_{U^k[\frac{N-A}{W}]} \ll_k w(N)^{-c} \tag{10-1}$$

for some constant $c > 0$. We observe that, using $\Lambda_{\chi, y'}(n)$ as a majorant for ν , it follows from this uniformity norm estimate and the generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] that the W -tricked version of ν satisfies the linear forms condition from [Green and Tao 2008, Definition 6.2].

10.3. Proof of Theorem 1.3. Let $y_0 \leq y \leq N^\eta$, where $\eta \in (0, 1)$ is sufficiently small depending on d and y_0 will be determined in the course of the proof, and let $w(N) \leq y' \leq (\log N)^{K'}$ for some fixed $K' \geq 1$. The transferred version of the quantitative inverse theorem for U^k -norms of [Manners 2018] proved in [Tao and Teräväinen 2023, Theorem 8.3] allows us to deduce explicit U^k -norm estimates from Theorem 1.3, which we will later combine with the transferred generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1]. More precisely, Theorem 1.1 implies that if $\Psi(N, [y', y])P(y')/(N\phi(P(y'))) \geq (\log_4 N)^{-1/2}$, then

$$\begin{aligned} \left| \frac{W}{N} \sum_{N/W < n \leq 2N/W} (g_{[y', y]}^{(W, A)}(n) - 1)F(g(n)\Gamma) \right| &\ll (1 + \|F\|_{\text{Lip}})(\log_4 N)^{-1} \\ &\ll \frac{1 + \|F\|_{\text{Lip}}}{\sqrt{\log_4 N}} \frac{N\phi(P(y'))}{\Psi(N, [y', y])P(y')} \end{aligned}$$

for all filtered nilmanifolds of dimension $O((\log_5 N)^{O(1)})$ and complexity Q_0 that is bounded by

$$Q_0 < \exp\left(\frac{1}{2}\sqrt{\log_4 N}\right),$$

and all nilsequences attached to it. Observe that

$$(\log_4 N)^{-1/2} = o(\exp(-\exp(C_1/\tilde{\delta}^{C_2})))$$

and

$$\exp \exp(C_1/\tilde{\delta}^{C_2}) = o(\exp(\frac{1}{2}\sqrt{\log_4 N}))$$

for all constants $C_1, C_2 > 1$ if $\tilde{\delta} = 1/\log_7 N$. Note further that the function from Section 10.2 may be used as a trivial majorising function for the following rescaled version of $g_{[y', y]}^{(W, A)}(n)$, which essentially corresponds to $\mathbf{1}_{S([y', y])}(n)$. We have

$$\frac{P(y')}{\phi(P(y'))} \frac{\Psi(N, [y', y])}{N} g_{[y', y]}^{(W, A)}(n) \ll \nu(n) \quad (n \leq (N - A)/W).$$

By (10-1), we have

$$\|\nu - 1\|_{U^k[(N-A)/W]} \ll_k (\log_3 N)^{-c} \ll_B \tilde{\delta}^B$$

for all $B \geq 1$. Hence the transferred quantitative U^k inverse theorem of [Manners 2018], as stated in [Tao and Teräväinen 2023, Theorem 8.3], implies that uniformly for all $0 < A < W$ with $(W, A) = 1$ we have

$$\|g_{[y',y]}^{(W,A)} - 1\|_{U^k((N-A)/W)} \ll \frac{\tilde{\delta} N \phi(P(y'))}{\Psi(N, [y', y]) P(y')},$$

which is small compared to the mean value 1 of $g_{[y',y]}^{(W,A)}$ provided that

$$\Psi(N, [y', y]) P(y') / (N \phi(P(y'))) = o(\log_7 N).$$

In order to apply the generalised von Neumann theorem [Green and Tao 2010, Proposition 7.1] in our situation, observe that the $o(1)$ error term in [loc. cit., Proposition 7.1'] corresponds to the error term in the (D, D, D) -linear forms condition for ν . It follows from [Tao and Teräväinen 2023, Proposition 5.2] that any function of the form

$$\tilde{\nu}(n) = \frac{1}{t} \sum_{i=1}^t \frac{\phi(W)}{W} \nu^{(N)}(Wn + A_i),$$

with $0 < A_i < W$, $\gcd(A_i, W) = 1$ for all $1 \leq i \leq t$, satisfies the (D, D, D) -linear forms condition from [Green and Tao 2010, Definition 6.2] for any given $D = O_{r,s,L}(1)$ with an error term of the form

$$w(N)^{-c} = (\log_3 N)^{-c}.$$

Applying [loc. cit., Proposition 7.1] with

$$f_i(n) = \frac{P(y')}{\phi(P(y'))} \frac{\Psi(N, [y', y])}{N} g_{[y',y]}^{(W,A)}(n)$$

and with $\tilde{\nu}$ as above, we obtain

$$\sum_{\mathbf{n} \in \mathbb{Z}^s \cap (N\mathfrak{K})/W} \prod_{j=1}^r g_{[y',y]}(W\psi_j(\mathbf{n}) + A_j) = \left(\frac{W}{\phi(W)}\right)^r \left\{ \frac{N^s \text{vol } \mathfrak{K}}{W^s} + o_{s,r,\|\Psi\|} \left(\left(\frac{N}{W}\right)^s \right) \right\}$$

provided that

$$\|g_{[y',y]}^{(W,A)} - 1\|_{U^k((N-A)/W)} = o\left(\left(\frac{\Psi(N, [y', y]) P(y')}{N \phi(P(y'))}\right)^{r-1}\right)$$

for $k = O_{r,s}(1)$, as well as

$$w(N)^{-c} = (\text{linear forms condition error term}) = o\left(\left(\frac{\Psi(N, [y', y]) P(y')}{N \phi(P(y'))}\right)^r\right).$$

These conditions are certainly satisfied if

$$\frac{\Psi(N, [y', y]) P(y')}{N \phi(P(y'))} \gg (\log_8 N)^{-1}.$$

Since

$$\frac{\Psi(N, [y', y]) P(y')}{N \phi(P(y'))} \asymp \frac{\Psi(N, y)}{N} \prod_{p \leq y'} \frac{1 - p^{-\alpha(N,y)}}{1 - p^{-1}} \leq \frac{\Psi(N, y)}{N} = u^{-u+o(u)}$$

by Lemma 3.3, it follows that $N^{-1}\Psi(N, y) \gg (\log_8 N)^{-1}$. Let y_0 be such that

$$\frac{\Psi(N, y_0)}{N} = (\log_8 N)^{-1}$$

and write $u_0 = (\log N)/\log y_0$. Then

$$\frac{\log N}{\log y_0} = u_0 < u_0^{u_0+o(u_0)} = \log_8 N$$

and, if $y > y_0$, then

$$1 - \alpha(N, y) \ll \frac{\log(u_0 \log(u_0 + 2))}{\log y_0} \ll \frac{(\log_8 N)^2}{\log N}.$$

Hence, the first part of Lemma 3.2 applies to $g_{P(y')}(\alpha(N, y))$ and yields

$$g_{P(y')}(\alpha(N, y)) \asymp \frac{1}{\log y'} \asymp \frac{\phi(P(y'))}{P(y')}$$

for $y \geq y_0$ and $y' = (\log N)^{K'}$. Thus

$$\frac{\Psi(N, [y', y])P(y')}{N\phi(P(y'))} \asymp \frac{\Psi(N, y)}{N} \gg (\log_8 N)^{-1}$$

for $y \geq y_0$. From

$$\exp(u_0^2) > u_0^{u_0+o(u_0)} = \log_8 N,$$

which holds for sufficiently large N , we deduce that $y_0 < N^{1/\sqrt{\log_8 N}}$.

Suppose now that $y_0 \leq y \leq N^\eta$. Then we obtain

$$\begin{aligned} & \sum_{\mathbf{n} \in \mathbb{Z}^s \cap N\mathfrak{K}} \prod_{j=1}^r g_{[y', y]}(\psi_j(\mathbf{n}) + a_j) \\ &= \sum_{\mathbf{A} \in \{0, \dots, W-1\}^s} \sum_{\substack{W\mathbf{n} + \mathbf{A} \\ \in \mathbb{Z}^s \cap N\mathfrak{K}}} \prod_{j=1}^r g_{[y', y]}(\psi_j(W\mathbf{n} + \mathbf{A}) + a_j) \\ &= \sum_{\mathbf{A} \in \{0, \dots, W-1\}^s} \sum_{\mathbf{n} \in \mathbb{Z}^s \cap (N\mathfrak{K} - \mathbf{A})/W} \prod_{j=1}^r g_{[y', y]}(W\psi_j(\mathbf{n}) + \psi_j(\mathbf{A}) + a_j) \\ &= \left(\frac{W}{\phi(W)}\right)^r \left\{ \frac{N^s \text{vol } \mathfrak{K}}{W^s} + o_{s,r,\|\Psi\|} \left(\left(\frac{N}{W}\right)^s \right) \right\} \sum_{\mathbf{A} \in \{0, \dots, W-1\}^s} \prod_{j=1}^r \mathbf{1}_{\gcd(W, \psi_j(\mathbf{A}) + a_j) = 1} \\ &= \{N^s \text{vol } \mathfrak{K} + o_{s,r,\|\Psi\|}(N^s)\} \prod_{p < w(N)} \beta_p, \end{aligned}$$

where

$$\beta_p = \frac{1}{p^s} \sum_{\mathbf{u} \in (\mathbb{Z}/p\mathbb{Z})^s} \prod_{j=1}^r \frac{p}{p-1} \mathbf{1}_{\psi_j(\mathbf{u}) + a_j \not\equiv 0 \pmod{p}}.$$

Theorem 1.3 now follows on recalling⁷ from [Green and Tao 2010, Lemma 1.3] that the condition that the forms ψ_i be pairwise linearly independent over \mathbb{Q} implies $\beta_p = 1 + O_{s,r,L}(p^{-2})$ for all primes $p > w(N)$.

⁷Observe that our local factors are identical to those defined in [Green and Tao 2010, (1.6)]

Acknowledgements

We would like to thank Adam Harper for very helpful comments on a previous version of this paper and the anonymous referee for their astute questions and suggestions, leading us, in particular, to rework our arguments and remove the previous restriction to large values of y' ($\asymp (\log N)^C$). While working on this paper, the authors were supported by the Swedish Research Council (VR) grant 2021-04526 and the Göran Gustafsson Foundation.

References

- [Balog et al. 2012] A. Balog, V. Blomer, C. Dartyge, and G. Tenenbaum, “Friable values of binary forms”, *Comment. Math. Helv.* **87**:3 (2012), 639–667. MR
- [de la Bretèche and Tenenbaum 2005] R. de la Bretèche and G. Tenenbaum, “Propriétés statistiques des entiers friables”, *Ramanujan J.* **9**:1-2 (2005), 139–202. MR
- [Drappeau and Shao 2016] S. Drappeau and X. Shao, “Weyl sums, mean value estimates, and Waring’s problem with friable numbers”, *Acta Arith.* **176**:3 (2016), 249–299. MR
- [Fouvry 2010] É. Fouvry, “Friabilité des valeurs d’un polynôme”, *Arch. Math. (Basel)* **95**:5 (2010), 411–421. MR
- [Fouvry and Tenenbaum 1996] É. Fouvry and G. Tenenbaum, “Répartition statistique des entiers sans grand facteur premier dans les progressions arithmétiques”, *Proc. Lond. Math. Soc. (3)* **72**:3 (1996), 481–514. MR
- [Granville 2008] A. Granville, “Smooth numbers: computational number theory and beyond”, pp. 267–323 in *Algorithmic number theory: lattices, number fields, curves and cryptography*, edited by J. P. Buhler and P. Stevenhagen, Math. Sci. Res. Inst. Publ. **44**, Cambridge Univ. Press, 2008. MR
- [Green and Tao 2008] B. Green and T. Tao, “The primes contain arbitrarily long arithmetic progressions”, *Ann. of Math. (2)* **167**:2 (2008), 481–547. MR
- [Green and Tao 2010] B. Green and T. Tao, “Linear equations in primes”, *Ann. of Math. (2)* **171**:3 (2010), 1753–1850. MR
- [Green and Tao 2012a] B. Green and T. Tao, “The Möbius function is strongly orthogonal to nilsequences”, *Ann. of Math. (2)* **175**:2 (2012), 541–566. MR
- [Green and Tao 2012b] B. Green and T. Tao, “The quantitative behaviour of polynomial orbits on nilmanifolds”, *Ann. of Math. (2)* **175**:2 (2012), 465–540. MR
- [Green et al. 2012] B. Green, T. Tao, and T. Ziegler, “An inverse theorem for the Gowers $U^{s+1}[N]$ -norm”, *Ann. of Math. (2)* **176**:2 (2012), 1231–1372. MR
- [Harper 2012a] A. J. Harper, “Bombieri–Vinogradov and Barban–Davenport–Halberstam type theorems for smooth numbers”, preprint, 2012. arXiv 1208.5992
- [Harper 2012b] A. J. Harper, “On a paper of K. Soundararajan on smooth numbers in arithmetic progressions”, *J. Number Theory* **132**:1 (2012), 182–199. MR
- [Harper 2016] A. J. Harper, “Minor arcs, mean values, and restriction theory for exponential sums over smooth numbers”, *Compos. Math.* **152**:6 (2016), 1121–1158. MR
- [Hildebrand 1986a] A. Hildebrand, “On the local behavior of $\Psi(x, y)$ ”, *Trans. Amer. Math. Soc.* **297**:2 (1986), 729–751. MR
- [Hildebrand 1986b] A. Hildebrand, “On the number of positive integers $\leq x$ and free of prime factors $> y$ ”, *J. Number Theory* **22**:3 (1986), 289–307. MR
- [Hildebrand and Tenenbaum 1986] A. Hildebrand and G. Tenenbaum, “On integers free of large prime factors”, *Trans. Amer. Math. Soc.* **296**:1 (1986), 265–290. MR
- [Lachand 2017] A. Lachand, “On the representation of friable integers by linear forms”, *Acta Arith.* **181**:2 (2017), 97–109. MR
- [Lagarias and Soundararajan 2011] J. C. Lagarias and K. Soundararajan, “Smooth solutions to the abc equation: the xyz conjecture”, *J. Théor. Nombres Bordeaux* **23**:1 (2011), 209–234. MR

- [Lagarias and Soundararajan 2012] J. C. Lagarias and K. Soundararajan, “Counting smooth solutions to the equation $A + B = C$ ”, *Proc. Lond. Math. Soc.* (3) **104**:4 (2012), 770–798. MR
- [Leng et al. 2024] J. Leng, A. Sah, and M. Sawhney, “Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$ -norm”, preprint, 2024. arXiv 2402.17994
- [Manners 2018] F. Manners, “Quantitative bounds in the inverse theorem for the Gowers U^{s+1} -norms over cyclic groups”, preprint, 2018. arXiv 1811.00718
- [Matthiesen 2012] L. Matthiesen, “Linear correlations amongst numbers represented by positive definite binary quadratic forms”, *Acta Arith.* **154**:3 (2012), 235–306. MR
- [Matthiesen 2018] L. Matthiesen, “Generalized Fourier coefficients of multiplicative functions”, *Algebra Number Theory* **12**:6 (2018), 1311–1400. MR
- [Matthiesen 2020] L. Matthiesen, “Linear correlations of multiplicative functions”, *Proc. Lond. Math. Soc.* (3) **121**:2 (2020), 372–425. MR
- [Montgomery and Vaughan 1977] H. L. Montgomery and R. C. Vaughan, “Exponential sums with multiplicative coefficients”, *Invent. Math.* **43**:1 (1977), 69–82. MR
- [Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Stud. Adv. Math. **97**, Cambridge Univ. Press, 2007. MR
- [Soundararajan 2008] K. Soundararajan, “The distribution of smooth numbers in arithmetic progressions”, pp. 115–128 in *Anatomy of integers* (Montreal, 2006), edited by J.-M. De Koninck et al., CRM Proc. Lecture Notes **46**, Amer. Math. Soc., Providence, RI, 2008. MR
- [Tao and Teräväinen 2023] T. Tao and J. Teräväinen, “Quantitative bounds for Gowers uniformity of the Möbius and von Mangoldt functions”, *J. Eur. Math. Soc.* (online publication December 2023).
- [Vaughan 1989] R. C. Vaughan, “A new iterative method in Waring’s problem”, *Acta Math.* **162**:1-2 (1989), 1–71. MR
- [Wooley 1992] T. D. Wooley, “Large improvements in Waring’s problem”, *Ann. of Math.* (2) **135**:1 (1992), 131–164. MR
- [Wooley 1995] T. D. Wooley, “New estimates for smooth Weyl sums”, *J. Lond. Math. Soc.* (2) **51**:1 (1995), 1–13. MR

Communicated by Andrew Granville

Received 2022-12-10 Revised 2024-10-17 Accepted 2024-12-05

lilian.matthiesen@mathematik.uni-goettingen.de

Mathematisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany

mengdi.wang@epfl.ch

Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR
Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR
David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada	Philippe Michel	École Polytechnique Fédérale de Lausanne
Bhargav Bhatt	University of Michigan, USA	Martin Olsson	University of California, Berkeley, USA
Frank Calegari	University of Chicago, USA	Irena Peeva	Cornell University, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France	Jonathan Pila	University of Oxford, UK
Brian D. Conrad	Stanford University, USA	Anand Pillay	University of Notre Dame, USA
Samit Dasgupta	Duke University, USA	Bjorn Poonen	Massachusetts Institute of Technology, USA
Hélène Esnault	Freie Universität Berlin, Germany	Victor Reiner	University of Minnesota, USA
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Vasudevan Srinivas	SUNY Buffalo, USA
Wee Teck Gan	National University of Singapore	Shunsuke Takagi	University of Tokyo, Japan
Andrew Granville	Université de Montréal, Canada	Pham Huu Tiep	Rutgers University, USA
Ben J. Green	University of Oxford, UK	Ravi Vakil	Stanford University, USA
Christopher Hacon	University of Utah, USA	Akshay Venkatesh	Institute for Advanced Study, USA
Roger Heath-Brown	Oxford University, UK	Melanie Matchett Wood	Harvard University, USA
János Kollár	Princeton University, USA	Shou-Wu Zhang	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA		

PRODUCTION

production@msp.org
Silvio Levy, Scientific Editor


See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2025 is US \$565/year for the electronic version, and \$820/year (+\$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW[®] from MSP.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2025 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 19 No. 10 2025

Smooth numbers are orthogonal to nilsequences LILIAN MATTHIESEN and MENGDI WANG	1881
Stacked pseudo-convergent sequences and polynomial Dedekind domains GIULIO PERUGINELLI	1947
Affine Deligne–Lusztig varieties via the double Bruhat graph, I: Semi-infinite orbits FELIX SCHREMMER	1973
Affine Deligne–Lusztig varieties via the double Bruhat graph, II: Iwahori–Hecke algebra FELIX SCHREMMER	2015
Paucity of rational points on fibrations with multiple fibres TIM BROWNING, JULIAN LYCZAK and ARNE SMEETS	2049