

Quantitative asymptotics for polynomial patterns in the primes

Lilian Matthiesen¹ | Joni Teräväinen² | Mengdi Wang³

¹Mathematisches Institut, Georg-August Universität Göttingen, Göttingen, Germany

²Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK

³École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Correspondence

Joni Teräväinen, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK.

Email: joni.p.teravainen@gmail.com

Funding information

European Union's Horizon Europe research and innovation programme; Marie Skłodowska-Curie, Grant/Award Number: 101058904; ERC, Grant/Award Number: 101162746; Academy of Finland, Grant/Award Number: 333707; Swedish Research Council, Grant/Award Numbers: 2021-04526, 2021-06594

Abstract

We prove quantitative estimates for averages of the von Mangoldt and Möbius functions along polynomial progressions $n + P_1(m), \dots, n + P_k(m)$ for a large class of polynomials P_i . The error terms obtained save an arbitrary power of logarithm, matching the classical Siegel–Walfisz error term. These results give the first quantitative bounds for the Tao–Ziegler polynomial patterns in the primes result. The proofs are based on a quantitative generalised von Neumann theorem of Peluse, a recent result of Leng on strong bounds for the Gowers uniformity of the primes, and analysis of a ‘Siegel model’ for the von Mangoldt function along polynomial progressions.

MSC 2020

11B30, 11N32

1 | INTRODUCTION

1.1 | Möbius and von Mangoldt averages along polynomial progressions

Given $k \in \mathbb{N}$ and polynomials $P_1, \dots, P_k \in \mathbb{Z}[y]$, a *polynomial progression* is a sequence of the form $n + P_1(m), \dots, n + P_k(m)$ with $m, n \in \mathbb{Z}$. In a seminal work, Tao and Ziegler [19] proved that the

primes contain infinitely many non-constant polynomial progressions for any distinct polynomials P_i satisfying $P_i(0) = 0$ for all $1 \leq i \leq k$. This generalised the Green–Tao theorem [6] that corresponded to the polynomials P_i being linear.

In a subsequent work, Tao and Ziegler [20] went further by showing that, if $d = \max_{1 \leq i \leq k} \deg(P_i)$ and Λ denotes the von Mangoldt function,[†] then there is an asymptotic formula for the count

$$\frac{1}{N^{d+1}} \sum_{n \leq N^d} \sum_{m \leq N} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m)) \tag{1.1}$$

of polynomial patterns in the primes, provided that $P_i - P_j$ has degree d for all $1 \leq i < j \leq k$. This generalised the work of Green, Tao and Ziegler [7–9] that handled the case of linear polynomials P_i .

The results in [19] are qualitative in nature, as they rely on the Gowers uniformity result for the von Mangoldt function from [7] (which is based on the qualitative Green–Tao–Ziegler [9] inverse theorem for the Gowers norms). Subsequently, a quantitative bound for the Gowers norms of the primes was established in [18] (using a quantitative inverse theorem for the Gowers norms due to Manners [14]). Very recently, Leng [12] made a substantial improvement to this result, obtaining arbitrary power of logarithm savings (using the quasipolynomial inverse theorem of Leng, Sah and Sawhney [13]). Leng applied his result in [12] to give an asymptotic formula, with arbitrary power of logarithm savings, for the count (1.1) in the case of linear polynomials.

In the case where the P_i are not all linear polynomials, we are only aware of quantitative results in the case $k \leq 2$ where classical Fourier analysis applies. Zhou [22] handled the case of the polynomials $\{0, y^d\}$ for any $d \in \mathbb{N}$ with savings of an arbitrary power of logarithm. See also [1, 2] for results of this type.

When it comes to a qualitative asymptotic formula for (1.1), it is only known in the aforementioned case when $P_i - P_j$ has degree d for all $1 \leq i < j \leq k$ and a few additional cases with $k \leq 4$ (given in [20, Theorem 5 and Remark 2]).

The purpose of this paper is to prove, building on Leng’s strong bounds for the Gowers uniformity of the primes [12] and a slightly extended version of a quantitative generalised von Neumann theorem of Peluse [16], a quantitative asymptotic formula for the count (1.1) for a large class of polynomials. We obtain arbitrary power of logarithm savings, as in [12], which is the best one can hope for in the absence of an improvement to the Siegel–Walfisz theorem.[‡] The class of polynomials we can handle is the same for which qualitative results were obtained in [20].

Theorem 1.1 (Quantitative polynomial patterns in the primes). *Let $k, d \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed. Let $\max_{1 \leq i \leq k} \deg(P_i) = d$ and suppose that $\deg(P_i - P_j) = d$ for all $1 \leq i < j \leq k$. Then, for any $N \geq 3$ and $A > 0$, we have*

[†] For convenience, we define Λ on all of \mathbb{Z} as an even function; thus $\Lambda(n) = \log p$ if $|n| = p^j$ for some prime p and some $j \in \mathbb{N}$, and $\Lambda(n) = 0$ otherwise.

[‡] For example, in the case of the polynomials $0, y, 2y$, it should be possible (e.g. following the approach in [18]) to extract the contribution of a Siegel zero to the main term, with the conclusion that any improvement to the error term beyond an arbitrary power of logarithm would lead to an improvement to Siegel’s theorem and hence to the Siegel–Walfisz theorem. We will not pursue the details of this, however.

$$\frac{1}{N^{d+1}} \sum_{n \leq N^d} \sum_{m \leq N} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m)) = \prod_p \beta_p + O_A((\log N)^{-A}), \tag{1.2}$$

where

$$\beta_p := \mathbb{E}_{m, n \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_p(n + P_j(m)),$$

and the local von Mangoldt function Λ_p is given by

$$\Lambda_p(n) := \frac{p}{p-1} 1_{n \not\equiv 0 \pmod{p}}. \tag{1.3}$$

Remarks.

- By [19, Lemma 9.5], we have

$$\beta_p = 1 + O(1/p^2), \tag{1.4}$$

since all large enough primes p are ‘good’ in the sense of [19, Definition 9.4]. This implies that the product $\prod_p \beta_p$ is convergent. Thus, $\prod_p \beta_p > 0$ unless there exists a prime p such that for all $m, n \in \mathbb{Z}$ at least one of $n + P_1(m), \dots, n + P_k(m)$ is divisible by p .

- The result of Tao and Ziegler [20] also works for multivariate polynomials. If Proposition 3.1 could be extended to multivariate polynomials, then it seems likely that Theorem 1.1 could also be extended to them.

We also prove a similar quantitative bound for averages of the Möbius function[†] μ along a larger class of polynomial progressions.

Theorem 1.2 (Quantitative polynomial patterns with Möbius weight). *Let $k, d \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed. Let $\max_{1 \leq i \leq k} \deg(P_i) = d$. Suppose that there exists $1 \leq \ell \leq k$ such that $\deg(P_\ell - P_i) = d$ for all $i \neq \ell$. Then, for any $N \geq 3$ and $A > 0$, we have*

$$\left| \frac{1}{N^{d+1}} \sum_{n \leq N^d} \sum_{m \leq N} \mu(n + P_1(m)) \cdots \mu(n + P_k(m)) \right| \ll_A (\log N)^{-A}.$$

Remarks.

- Theorem 1.2 holds equally well with the Liouville function λ in place of the Möbius function μ . This follows from the same proof along with Remark 4.2.
- Note that in Theorem 1.2 we can allow some of the polynomials to differ only by a constant (or even to be equal). Thus, for example, if $P_1, \dots, P_{k-1} \in \mathbb{Z}[y]$ are any polynomials of degree at most $d - 1$, then $\{P_1, \dots, P_{k-1}, y^d\}$ is a valid collection for this theorem.[‡]

[†] For convenience, we define μ on all of \mathbb{Z} as an even function; thus, $\mu(n) = (-1)^k$ if $|n|$ is the product of k distinct primes and $\mu(n) = 0$ if $|n|$ is not squarefree.

[‡] In Theorem 1.1, we needed a stricter condition on the polynomials, ultimately because in the proof of that theorem one needs to split each copy of the von Mangoldt function into a ‘structured’ and ‘random’ part, and a result of the type of Theorem 1.2 only applies to the random parts.

As far as we are aware, Theorem 1.2 gives the first quantitative bounds for the averages of the Möbius function along polynomial progressions with non-linear polynomials (and more than two polynomials). Tao and Ziegler stated that their method in [20] works for giving qualitative cancellation for averages of μ along polynomial progressions in the same cases where they handled the von Mangoldt weight. Qualitative cancellation for averages of μ over general polynomial progressions (with no two polynomials differing by a constant) was subsequently proved in [15, Theorem 1.10].

1.2 | One-dimensional averages

It turns out that we can give quantitative estimates not only for the double averages (1.1), but also for the single averages

$$\sum_{n \leq N^d} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m)) \quad \text{and} \quad \sum_{m \leq N} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m))$$

for almost all $m \in [N]$ (respectively, for almost all $n \in [N^d]$). This follows from the following theorem, from which Theorems 1.1 and 1.2 will be deduced.

Theorem 1.3. *Let $k, d \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed. Let $\max_{1 \leq i \leq k} \deg(P_i) = d$, $N \geq 3$ and $A > 0$.*

(1) *Suppose that there exists $1 \leq \ell \leq k$ such that $\deg(P_\ell - P_i) = d$ for all $i \neq \ell$. Then, we have*

$$\frac{1}{N^{d+1}} \sum_{m \leq N} \left| \sum_{n \leq N^d} \mu(n + P_1(m)) \cdots \mu(n + P_k(m)) \right| \ll_A (\log N)^{-A},$$

and

$$\frac{1}{N^{d+1}} \sum_{n \leq N^d} \left| \sum_{m \leq N} \mu(n + P_1(m)) \cdots \mu(n + P_k(m)) \right| \ll_A (\log N)^{-A}.$$

(2) *Suppose that $\deg(P_i - P_j) = d$ for all $1 \leq i < j \leq k$. Then we have*

$$\frac{1}{N^{d+1}} \sum_{m \leq N} \left| \sum_{n \leq N^d} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m)) - \prod_p \beta_p(m) \right| \ll_A (\log N)^{-A},$$

and if additionally all the P_i are non-constant, we have

$$\frac{1}{N^{d+1}} \sum_{n \leq N^d} \left| \sum_{m \leq N} \Lambda(n + P_1(m)) \cdots \Lambda(n + P_k(m)) - \prod_p \beta'_p(n) \right| \ll_A (\log N)^{-A},$$

where

$$\beta_p(m) := \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_p(n + P_j(m)), \quad \beta'_p(n) := \mathbb{E}_{m \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_p(n + P_j(m)). \quad (1.5)$$

Theorem 1.3(1) implies Theorem 1.2 directly by the triangle inequality. The implication from Theorem 1.3(2) to Theorem 1.1 is shown in Section 7. Also in Section 7, we prove Theorem 1.3(2) after giving an overview of the proof and the role of the Siegel model in the proof.

As an immediate corollary of Theorem 1.3 and Markov’s inequality, we have the following almost-all result.

Corollary 1.4 (An averaged multi-dimensional Bateman–Horn result). *Let $k, d \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed. Let $\max_{1 \leq i \leq k} \deg(P_i) = d$. Suppose that $\deg(P_i - P_j) = d$ for all $1 \leq i < j \leq k$. Let $N \geq 3$ and $A > 0$. Then, we have*

$$\left| \sum_{m \leq N} \Lambda(P_1(m) + n) \cdots \Lambda(P_k(m) + n) - \prod_p \beta'_p(n) \right| \ll_A N(\log N)^{-A}$$

for all but $\ll_A N^d / (\log N)^A$ integers $n \in [N^d]$.

The Bateman–Horn conjecture [3] predicts that the same asymptotic holds for all $n \in [N^d]$; thus, Corollary 1.4 can be seen as an averaged version of it. Specialising to $k = 1$, Corollary 1.4 extends an earlier result of Zhou [22], which handled the case where $P_1(y) = y^\ell$ for some $\ell \in \mathbb{N}$.

1.3 | Structure of the paper

After reviewing notation and providing some preliminary results in Section 2, we prove in Section 3 a slightly extended version of the quantitative generalised von Neumann theorem of Peluse [16, Theorem 6.1]. More precisely, the bounds in this extended version are given in terms of U^k norms instead of box norms and our version allows us to work with an additional weight factor. Section 4 collects together information about the Siegel model for the von Mangoldt function, which was introduced in [18], and proves a slight extension of Leng’s result [12, Theorem 6] on the Gowers uniformity of the primes which is applicable with a flexible W -trick. In Section 5, we use the version with the additional weight factor on Peluse’s result to prove Theorem 1.3(1) about the Möbius function. In preparation for the proof of Theorem 1.3(2), we establish in Section 6 a technical result (a local–global principle) that is used to handle the correlations of the Cramér model later. Finally, we prove Theorem 1.3(2) in Section 7 using the results from Sections 3, 4 and 6. Our approach, which we sketch at the start of that section, requires some flexibility in the W -trick and hence the version of Leng’s result provided in Section 4.

2 | NOTATION AND PRELIMINARIES

2.1 | Asymptotic and averaging notation

We use the usual Landau and Vinogradov asymptotic notations $O(\cdot)$, $o(\cdot)$, \ll , \gg . Thus, we write $X \ll Y$, $X = O(Y)$ or $Y \gg X$ if we have $|X| \leq CY$ for some constant C . We write $X \asymp Y$ if $X \ll Y$ and $Y \ll X$. We write $X = o(Y)$ as $N \rightarrow \infty$ if $|X| \leq c(N)Y$ for some function $c(N) \rightarrow 0$ as $N \rightarrow \infty$. If we attach subscripts to these notations, then the implied constants may depend on these subscripts. Thus, for example, $X \ll_A Y$ signifies that $|X| \leq C_A Y$ for some $C_A > 0$ depending on A .

For a set S , we write $1_S(x)$ for its indicator function, that is, a function that equals to 1 for $x \in S$ and equals to 0 otherwise. Similarly, if P is a proposition, the quantity 1_P equals to 1 if P is true and 0 if P is false.

For a non-empty finite set A and a function $f : A \rightarrow \mathbb{C}$, we use the averaging notation

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

For a real number $N > 0$, we denote $[N] := [1, N] \cap \mathbb{N}$. For $q \in \mathbb{Z}$, we also denote $q[N] := \{qn : n \in [1, N] \cap \mathbb{N}\}$.

We write $\|x\|$ for the distance from the real number x to the nearest integer. We use the standard notation $e(x) := e^{2\pi i x}$.

Given integers m, n , we write (m, n) for their greatest common divisor and $[m, n]$ for their least common multiple. We use $\tau(n)$ to denote the divisor function and $\Omega(n)$, $\omega(n)$ to denote the number of prime divisors of n , with and without multiplicities, respectively.

2.2 | Gowers norms

For $f : \mathbb{Z} \rightarrow \mathbb{C}$ with finite support and $s \in \mathbb{N}$, we define the unnormalised U^s Gowers norm of f as

$$\|f\|_{\tilde{U}^s(\mathbb{Z})} := \left(\sum_{x, h_1, \dots, h_s \in \mathbb{Z}} \prod_{\omega \in \{0, 1\}^s} C^{|\omega|} f(x + \omega \cdot (h_1, \dots, h_s)) \right)^{1/2^s},$$

where $C(z) = \bar{z}$ is the complex conjugation operator and for a vector $(\omega_1, \dots, \omega_s)$ we write $|\omega| := |\omega_1| + \dots + |\omega_s|$. For an interval $I \subset \mathbb{R}$ with $|I| \geq 1$, we further define the $U^s(I)$ Gowers norm of a function $f : \mathbb{Z} \rightarrow \mathbb{C}$ as

$$\|f\|_{U^s(I)} := \frac{\|f 1_I\|_{\tilde{U}^s(\mathbb{Z})}}{\|1_I\|_{\tilde{U}^s(\mathbb{Z})}}.$$

It is well known (see, e.g. [7, Appendix B]) that for $s \geq 2$ the $U^s(I)$ norm is indeed a norm and for $s = 1$ it is a semi-norm. When $I = [N_1, N_2]$, we abbreviate $\|f\|_{U^s(I)} = \|f\|_{U^s[N_1, N_2]}$ and if $N_1 = 1$ we further abbreviate this as $\|f\|_{U^s[N_2]}$.

The $\tilde{U}^s(\mathbb{Z})$ norms obey the Gowers–Cauchy–Schwarz inequality (see, e.g. [18, (4.2)]), which states that, for any functions $(f_\omega)_{\omega \in \{0,1\}^s}$ from \mathbb{Z} to \mathbb{C} with finite support, we have

$$\left| \sum_{x, h_1, \dots, h_s \in \mathbb{Z}} \prod_{\omega \in \{0,1\}^s} f_\omega(x + \omega \cdot (h_1, \dots, h_s)) \right| \leq \prod_{\omega \in \{0,1\}^s} \|f_\omega\|_{\tilde{U}^s(\mathbb{Z})}. \tag{2.1}$$

For $H \geq 1$, we define the weight

$$\mu_H(h) := \frac{|\{(h_1, h_2) \in [H]^2 : h_1 - h_2 = h\}|}{[H]^2}, \tag{2.2}$$

where $[x]$ denotes the floor of x . This weight will appear naturally when applying the Cauchy–Schwarz inequality to polynomial averages.

2.3 | Vinogradov’s Fourier expansion

We will make use of a Fourier approximation for the indicator function of an interval that is due to Vinogradov.

Lemma 2.1. *For any $-1/2 < \alpha < \beta < 1/2$ and $\eta \in (0, \min\{1/2 - \|\alpha\|, 1/2 - \|\beta\|, \|\alpha - \beta\|/2\})$, there exists a 1-periodic function $g : \mathbb{R} \rightarrow [0, 1]$ such that the following hold:*

- (1) $g(x) = 1$ for $x \in [\alpha + \eta, \beta - \eta]$ and $g(x) = 0$ for $x \in [-1/2, 1/2] \setminus [\alpha - \eta, \beta + \eta]$.
- (2) For some $|c_j| \leq 10\eta$, we have for all $x \in \mathbb{R}$ the Fourier expansion

$$g(x) = \beta - \alpha - \eta + \sum_{|j| > 0} c_j e(jx).$$

- (3) For any $K \geq 1$, we have

$$\sum_{|j| > K} |c_j| \leq \frac{10\eta^{-1}}{K}.$$

Proof. See [21, Chapter 1, Lemma 12] (with $r = 1, \Delta = 2\eta, (\alpha, \beta) \rightarrow (\alpha + \eta, \beta - \eta)$ there). □

3 | GOWERS NORM CONTROL OF POLYNOMIAL AVERAGES

For the proofs of the main theorems, we need to know that weighted polynomial averages (with maximal degree d) are controlled in terms of the $U^s[CN^d]$ norm of one of the weights for some constant C , assuming that the corresponding polynomials P_i satisfy the assumption of Theorem 1.3. This follows from a result that slightly extends the work of Peluse [16].

Proposition 3.1. *Let $k, d \in \mathbb{N}, C \geq 2$ and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be non-constant polynomials satisfying $\deg P_1 \leq \dots \leq \deg P_k = d$. Suppose that $\deg(P_k - P_i) = d$ for all $1 \leq i \leq k - 1$. Let c_i be the*

leading coefficient of P_i , and suppose that all of the coefficients of P_1, \dots, P_k are at most $C|c_k|$ in modulus. Suppose also that $|c_i| \geq |c_k|/C$ whenever $\deg(P_i) = d$ and that $|c_k| \leq N^\eta$ with $\eta > 0$ small enough in terms of d .

Let $N \geq 1$, and let $f_0, \dots, f_k : \mathbb{Z} \rightarrow \mathbb{C}$ be functions supported on $[-CN^d, CN^d]$ with $|f_i| \leq 1$ for all $0 \leq i \leq k$. Also let $\theta : [N] \rightarrow \mathbb{C}$ satisfy $|\theta| \leq 1$. Then, for some natural number $s \ll_{d,k} 1$ and some $1 \leq K \ll_{d,k} 1$, we have

$$\left| \frac{1}{N^{d+1}} \sum_{n \in \mathbb{Z}, m \in [N]} \theta(m) f_0(n) \prod_{j=1}^k f_j(n + P_j(m)) \right| \ll_{C,d,k} |c_k|^K (N^{-1} + \|f_k\|_{U^s[-CN^d, CN^d]})^{1/K}. \tag{3.1}$$

Remark 3.2. Compared with [16, Theorem 6.1], we have an additional weight θ on the m variable and we bound the averages in terms of the Gowers uniformity norms rather than Gowers box norms.

Proof. Reduction to the case $\theta \equiv 1$. We first reduce matters to proving the case $\theta \equiv 1$. Suppose that this case has been proved, and consider the general case. Let S denote the left-hand side of (3.1). Then, by the Cauchy–Schwarz inequality, we have

$$|S|^2 \leq \frac{1}{N^{2d+1}} \sum_{m \in [N]} \left| \sum_{n \in \mathbb{Z}} f_0(n) \prod_{j=1}^k f_j(n + P_j(m)) \right|^2.$$

Expanding out the square, the right-hand side becomes

$$\frac{1}{N^{2d+1}} \sum_{|h| \leq 2CN^d} \sum_{m \in [N]} \sum_{n \in \mathbb{Z}} \Delta_h f_0(n) \prod_{j=1}^k \Delta_h f_j(n + P_j(m)),$$

where $\Delta_h f(x) := \overline{f(x+h)}f(x)$. Applying the case $\theta \equiv 1$ to the inner sums here, we see that for some $1 \leq s, K \ll_{d,k} 1$ we have

$$|S|^2 \ll_{C,d,k} \frac{|c_k|^K}{N^d} \sum_{|h| \leq 2CN^d} (N^{-1} + \|\Delta_h f_k\|_{U^s[-CN^d, CN^d]})^{1/K}.$$

Using $(a + b)^{1/K} \leq a^{1/K} + b^{1/K}$ for $a, b \geq 0$ and Hölder’s inequality, we see that

$$|S|^2 \ll_{C,d,k} |c_k|^K \left(N^{-1/K} + \left(\frac{1}{N^d} \sum_{|h| \leq 2CN^d} \|\Delta_h f_k\|_{U^s[-CN^d, CN^d]}^{2^s} \right)^{1/(2^s K)} \right).$$

Expanding out the definition of the $U^s[-CN^d, CN^d]$ norm, we conclude that

$$|S|^2 \ll_{C,d,k} |c_k|^K \left(N^{-1/K} + \|f_k\|_{U^{s+1}[-CN^d, CN^d]}^{2/K} \right).$$

Now the claim (3.1) follows (with s replaced by $s + 1$ and K by $2K$) from the power mean inequality $(a^{1/K} + b^{1/K})/2 \leq ((a + b)/2)^{1/K}$ for $a, b \geq 0$.

Proof of the case $\theta \equiv 1$. We may also assume that C is large enough in terms of d, k .

We are going to apply [16, Theorem 6.1]. Let $s \ll_{d,k} 1$ be as in that theorem (we may assume that $s \geq 2$). Also let $B = B(d, k)$ be large enough. Let

$$\delta := \frac{1}{C^2 |c_k| N^{d+1}} \sum_{n \in \mathbb{Z}, m \in [N]} f_0(n) \prod_{j=1}^k f_j(n + P_j(m)). \tag{3.2}$$

We may suppose that

$$N \geq B \left(\frac{|c_k|}{\delta^B} \right)^B,$$

since otherwise solving this inequality for δ the claim of the proposition follows immediately. Since $|f_j| \leq 1$, we also have $\delta \leq 1/C^2$.

Now we apply [16, Theorem 6.1][†] with $(M, N, C) \rightarrow (N, C^2 |c_k| N^d, C^3)$ there (noting that the assumptions of that theorem are satisfied; in particular $|c_i| N^d / (C^2 |c_k| N^d) \in [C^{-3}, C^3]$ for all i). Recalling the definition of the Gowers box norms from [16, Section 2], we conclude that for some $t = t(d, k)$ (which may be assumed to be much smaller than B) and for $\delta' = \delta^t$ we have

$$\|f_k\|_{\square_{Q_1, \dots, Q_s}^s([C^2 |c_k| N^d])} \gg_{C,d,k} \delta', \tag{3.3}$$

where each Q_i equals to $r_i[\delta' N^d] \subset r_i \mathbb{Z}$ for some integer[‡] $0 < |r_i| \leq 2d! C |c_k|$. To simplify notation, we assume that all r_i are positive. Unwrapping the definition of the box norm, (3.3) yields

$$\begin{aligned} & \left| \frac{1}{C^2 |c_k| N^{d(2s+1)}} \sum_{n, h_1^{(0)}, h_1^{(1)}, \dots, h_s^{(0)}, h_s^{(1)} \in \mathbb{Z}} \prod_{j=1}^s 1_{h_j^{(0)}, h_j^{(1)} \in [r_j \delta' N^d] \cap (r_j \mathbb{Z})} \right. \\ & \quad \left. \cdot \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_k(n + h_1^{(\omega_1)} + \dots + h_s^{(\omega_s)}) \right| \\ & \gg_{C,d,k} (\delta')^{2^s}. \end{aligned}$$

Making the changes of variables $n' = n - (h_1^{(0)} + \dots + h_s^{(0)})$, $h_j = h_j^{(1)} - h_j^{(0)}$ and using the notation (2.2) and the trivial bound $1 \geq 1/(C^2 |c_k|)$, we see that

$$\left| \frac{1}{N^{d(s+1)}} \sum_{\substack{n' \in [-CN^d, CN^d] \\ h_1, \dots, h_s \in [-CN^d, CN^d]}} \prod_{j=1}^s 1_{h_j \equiv 0 \pmod{r_j}} \delta' N^d \mu_{\delta' N^d} \left(\frac{h_j}{r_j} \right) \right|$$

[†] That theorem is stated for functions supported on intervals of the form $[X]$, but the same proof works for functions supported on an interval of the form $[-X, X]$.

[‡] The numbers r_i are of the form c_k or $c_k - c_i$ for some i with $\deg P_i = d$; crucially, these numbers cannot be zero by our assumption on P_i .

$$\begin{aligned} & \left| \cdot \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_k(n' + \omega \cdot (h_1, \dots, h_s)) \right| \\ & \gg_{C,d,k} (\delta')^{2^s}. \end{aligned} \tag{3.4}$$

By the orthogonality of characters, we have

$$1_{h_j \equiv 0 \pmod{r_j}} = \frac{1}{r_j} \sum_{1 \leq a \leq r_j} e\left(\frac{ah_j}{r_j}\right). \tag{3.5}$$

Moreover, for any $H \geq 1$ and $x \in [-H, H]$ we have the Fourier expansion

$$H\mu_H(x) = 1 - \left| \frac{x}{H} \right| = \frac{1}{2} - \frac{2}{\pi^2} \sum_{m \equiv 1 \pmod{2}} \frac{1}{m^2} e\left(\frac{x}{2H}m\right). \tag{3.6}$$

Inserting (3.5), (3.6) into (3.4) and applying the triangle inequality, for some $\alpha_1, \dots, \alpha_s \in \mathbb{R}$ we obtain

$$\begin{aligned} & \left| \frac{1}{N^{d(s+1)}} \sum_{n, h_1, \dots, h_s \in [-2CN^d, 2CN^d]} \prod_{j=1}^s 1_{[-r_j\delta'N^d, r_j\delta'N^d]}(h_j) e(\alpha_j h_j) \right. \\ & \left. \cdot \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_k 1_{[-CN^d, CN^d]}(n + \omega \cdot \mathbf{h}) \right| \\ & \gg_{C,d,k} (\delta')^{2^s}. \end{aligned} \tag{3.7}$$

Observe that since $\delta \leq 1/C^2$ and C is large enough in terms of d, k , we have $r_j\delta' \leq 1/2$ for all $1 \leq j \leq s$. Note then that for any $c \in (0, 1)$ and $h \in [-2CN^d, 2CN^d]$ we have

$$1_{[-cN^d, cN^d]}(h) = 1_{\|h/(4CN^d)\| \leq c/(4C)}.$$

Hence, by Vinogradov’s Fourier expansion (Lemma 2.1), for any $1 \leq j \leq s$ there exists a trigonometric polynomial $T_j(x) = \sum_{0 \leq |m| \leq J} \gamma_{m,j} e(mx/(4CN^d))$ with $J = O_C((\delta')^{-10 \cdot 2^s})$ such that we have $|\gamma_{m,j}| \ll_C \delta'$ for all m and

$$\sum_{h \in [-2CN^d, 2CN^d]} |1_{[-r_j\delta'N^d, r_j\delta'N^d]}(h) - T_j(h)| \leq (\delta')^{5 \cdot 2^s}. \tag{3.8}$$

Replacing $1_{[-r_j\delta'N^d, r_j\delta'N^d]}(h_j)$ with $T_j(h_j)$ in (3.7), and crudely using the triangle inequality, the 1-boundedness of f_k and (3.8) to estimate the error from this replacement, we see that

$$\begin{aligned} & \left| \frac{1}{N^{d(s+1)}} \sum_{n, h_1, \dots, h_s \in [-2CN^d, 2CN^d]} \prod_{j=1}^s T_j(h_j) e(\alpha_j h_j) \cdot \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_k 1_{[-CN^d, CN^d]}(n + \omega \cdot \mathbf{h}) \right| \\ & \gg_{C,d,k} (\delta')^{2^s}. \end{aligned}$$

Writing out T_j as a trigonometric polynomial and applying the pigeonhole principle, we see that for some real numbers $\alpha'_1, \dots, \alpha'_s$ we have

$$\left| \frac{1}{N^{d(s+1)}} \sum_{n, h_1, \dots, h_s \in [-2CN^d, 2CN^d]} \prod_{j=1}^s e(\alpha'_j h_j) \cdot \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_k 1_{[-CN^d, CN^d]}(n + \omega \cdot \mathbf{h}) \right| \gg_{C,d,k} (\delta')^{100s2^s}.$$

Using $e(\alpha'_j h_j) = e(\alpha'_j(n + h_j))e(-\alpha'_j n)$, this can be written as

$$\left| \frac{1}{N^{d(s+1)}} \sum_{n, h_1, \dots, h_s \in [-2CN^d, 2CN^d]} \prod_{\omega \in \{0,1\}^s} C^{|\omega|} f_{k,\omega} 1_{[-CN^d, CN^d]}(n + \omega \cdot \mathbf{h}) \right| \gg_{C,d,k} (\delta')^{100s2^s},$$

where each $f_{k,\omega}$ is of the form $f_{k,\omega}(n) = f_k(n)e(\alpha_\omega n)$ for some real number α_ω .

Now, by the Gowers–Cauchy–Schwarz inequality (2.1) and the simple fact that

$$\|f_k e(\alpha \cdot)\|_{U^s([-CN^d, CN^d])} = \|f_k\|_{U^s([-CN^d, CN^d])}$$

for any real number α and any $s \geq 2$, we conclude that

$$\|f_k\|_{U^s([-CN^d, CN^d])}^{2^s} \gg_{C,d,k} (\delta')^{100s2^s}.$$

Recalling that $\delta' = \delta^t$ and taking $K = 100st$, this yields

$$\delta \ll_{C,d,k} \|f_k\|_{U^s([-CN^d, CN^d])}^{1/K}.$$

Taking into account the definition of δ in (3.2), the proof is now complete. □

4 | GOWERS NORMS OF THE MÖBIUS AND VON MANGOLDT FUNCTIONS

A key number-theoretic input needed for the proofs of our main theorems is a bound for the Gowers norms of the Möbius and von Mangoldt functions saving an arbitrary power of logarithm, due to Leng [12].

Proposition 4.1 (Quantitative U^k -uniformity of μ). *Let $k \in \mathbb{N}$, $A > 0$ and $N \geq 3$. Then, we have*

$$\|\mu\|_{U^k[N]} \ll_A (\log N)^{-A}.$$

Remark 4.2. Proposition 4.1 holds also with the Liouville function λ in place of the Möbius function μ ; this is seen from the identity

$$\lambda(n) = \sum_{d \geq 1} \mu(n/d) 1_{d^2|n}.$$

Proof of Proposition 4.1. This follows by combining [12, Theorem 6] with the triangle inequality for the Gowers norms and [18, Theorem 2.5] (where we can use Siegel’s bound to lower bound the size of a Siegel modulus instead of a weaker effective bound). \square

For the von Mangoldt function, we end up needing a stronger Gowers uniformity result, saving more than an arbitrary power of logarithm (as in [12]). This necessitates taking into account the contribution of a possible Siegel zero. We introduce some definitions for this.

Definition 4.3. Let $c_0 = 1/1000$ and $Q \geq 2$. A real number $\tilde{\beta} \in (0, 1)$ is called a *Siegel zero of level Q* if there exists a primitive Dirichlet character $\tilde{\chi}$ of conductor $\leq Q$ such that $L(\tilde{\beta}, \tilde{\chi}) = 0$ and $\tilde{\beta} \geq 1 - c_0/\log Q$. The character $\tilde{\chi}$ is then called a *Siegel character of level Q* .

By the Landau–Page theorem (see, e.g. [17]), we know that if a Siegel zero of level Q exists, it is unique and simple, and the corresponding Siegel character $\tilde{\chi}$ is a unique non-principal real character.

Definition 4.4 (The Siegel model). Let $w \geq 2$ and $W = \prod_{p \leq w} p$. Let $\tilde{\beta}$ be the Siegel zero of level w and $\tilde{\chi}$ be the Siegel character of level w , if they exist; otherwise let $\tilde{\beta} = 1$ and $\tilde{\chi} = 0$. Define for $n \in \mathbb{Z}$ the *Siegel model*

$$\tilde{\Lambda}_W(n) := \Lambda_W(n) \left(1 - \tilde{\chi}(|n|) |n|^{\tilde{\beta}-1} \right),$$

where

$$\Lambda_W(n) := \prod_{p|W} \Lambda_p(n), \tag{4.1}$$

with Λ_p as in (1.3).

With this notation, we have the following version of the prime number theorem in arithmetic progressions with a Siegel correction.

Proposition 4.5. Let $N \geq 2$. Let $2 \leq w \leq \exp((\log N)^{1/2})$, and let $\tilde{\Lambda}_W$ be the Siegel model as in Definition 4.4. Then

$$\max_{q \leq w} \max_{a \pmod{q}} \left| \sum_{\substack{m \leq N \\ m \equiv a \pmod{q}}} (\Lambda(m) - \tilde{\Lambda}_W(m)) \right| \leq N \exp(-c(\log N)^{1/2}) \tag{4.2}$$

for some absolute constant $c > 0$.

Proof. Let $1 \leq a \leq q \leq w$ be chosen so that the maximum in (4.2) is attained. Denote the sum on the left-hand side by $S(N; q, a)$. If a prime p divides (a, q) , then $p \leq w$, so $\tilde{\Lambda}_W(m) = 0$ for all $m \equiv 0 \pmod{p}$ and $\Lambda(m) = 0$ for all $m \equiv 0 \pmod{p}$ that are not of the form p^j with $j \in \mathbb{N}$. Hence, we have $|S(N; q, a)| \ll (\log N)^2$, which is much stronger than the claimed bound. This means that we can assume from now on that $(a, q) = 1$.

Expanding the indicator $1_{m \equiv a \pmod{q}}$ in terms of Dirichlet characters, we find

$$\sum_{\substack{m \leq N \\ m \equiv a \pmod{q}}} \tilde{\Lambda}_W(m) = \frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \bar{\chi}(a) \sum_{m \leq N} \tilde{\Lambda}_W(m) \chi(m). \tag{4.3}$$

We claim that for any character $\chi \pmod{q}$ we have

$$\sum_{m \leq N} \Lambda_W(m) \chi(m) = \begin{cases} N + O(N \exp(-(\log N)^{1/2}/100)), & \text{if } \chi \text{ is principal.} \\ O(N \exp(-(\log N)^{1/2}/100)), & \text{if } \chi \text{ is non-principal.} \end{cases} \tag{4.4}$$

Indeed, if $\chi \pmod{q}$ is principal, the fundamental lemma of sieve theory [5, Theorem 6.9] gives

$$\sum_{m \leq N} \Lambda_W(m) \chi(m) = \sum_{m \leq N} \Lambda_W(m) = N + O(N(\log w)e^{-s}),$$

where s is such that $w^s = N^{1/10}$, so in particular $s \geq (\log N)^{1/2}/10$. This gives the first part of the claim (4.4). If $\chi \pmod{q}$ is non-principal, in turn, the fundamental lemma of sieve theory again gives

$$\begin{aligned} \sum_{m \leq N} \Lambda_W(m) \chi(m) &= \sum_{\substack{b \pmod{q} \\ (b,q)=1}} \chi(b) \sum_{\substack{m \leq N \\ m \equiv b \pmod{q}}} \Lambda_W(m) \\ &= \frac{N}{\varphi(q)} \sum_{\substack{b \pmod{q} \\ (b,q)=1}} \chi(b) + O(N(\log w)e^{-s}) \\ &= O(N(\log w)e^{-s}), \end{aligned}$$

where s is as above, so in particular $s \geq (\log N)^{1/2}/10$. This completes the proof of (4.4).

Using (4.4), we obtain

$$\sum_{m \leq N} \tilde{\Lambda}_W(m) \chi(m) = N 1_{\chi \text{ principal}} - \frac{N^{\tilde{\beta}}}{\tilde{\beta}} 1_{\chi \tilde{\chi} \text{ principal}} + O(N \exp(-(\log N)^{1/2}/100)).$$

Inserting this into (4.3), we see that

$$\sum_{\substack{m \leq N \\ m \equiv a \pmod{q}}} \tilde{\Lambda}_W(m) = \frac{N}{\varphi(q)} \left(1 - 1_{\tilde{q}|q} \tilde{\chi}(a) \frac{N^{\tilde{\beta}-1}}{\tilde{\beta}} \right) + O(N \exp(-(\log N)^{1/2}/100)).$$

Now the claim (4.2) follows from [10, Theorem 5.27] (which is a strengthening of the prime number theorem in arithmetic progressions taking the possible Siegel zero into account). \square

We are now ready to prove the key Gowers uniformity property of the von Mangoldt function used in this paper.

Proposition 4.6 (Quantitative U^k -uniformity of Λ). *For each $k \in \mathbb{N}$, there exists $c_k > 0$ such that the following holds. Let $N \geq 3$ and $\exp((\log \log N)^{1/c_k}) \leq w \leq \exp((\log N)^{1/10})$. Also let $W = \prod_{p \leq w} p$. Then we have*

$$\|\Lambda - \tilde{\Lambda}_W\|_{U^k[N]} \ll_k \exp(-(\log w)^{c_k}). \tag{4.5}$$

Proof. For $w = \exp((\log N)^{1/10})$, this is Leng’s [12, Theorem 6]. The general case follows from the same argument; for the sake of completeness we give some details. We defer any notation related to nilmanifolds to [12].

Let C_k be a large enough constant, and let $c_k > 0$ be sufficiently small in terms of C_k . Let

$$\delta = \exp(-(\log w)^{c_k}).$$

We have $\delta \leq 1/\log N$ by assumption, and clearly either (4.5) holds or

$$\|\delta(\Lambda - \tilde{\Lambda}_W)\|_{U^k[N]} \geq \delta^2. \tag{4.6}$$

The function $\delta(\Lambda - \tilde{\Lambda}_W)$ is bounded pointwise by $O(1)$. Hence, we can apply the quasipolynomial inverse theorem for the Gowers norms [13]. Recalling that C_k is large, the quasipolynomial inverse theorem and (4.6) imply that there exists a nilmanifold G/Γ of degree $\leq k - 1$, step $\leq k - 1$, dimension $D \leq (\log(1/\delta))^{C_k}$ and complexity $M \leq \exp((\log(1/\delta))^{C_k})$, and a function $F : G/\Gamma \rightarrow \mathbb{C}$ which has Lipschitz norm at most 1 (and hence is 1-bounded) such that

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \tilde{\Lambda}_W(n))F(g(n)\Gamma) \right| \gg \exp\left(-\left(\log \frac{1}{\delta}\right)^{C_k}\right). \tag{4.7}$$

Combining [12, Lemma A.6] (which is a Fourier expansion on the vertical torus) and [12, Lemma 2.2], we may assume that F is a vertical character with frequency ξ satisfying $|\xi| \leq \delta^{-(2D)^{C_k}}$ and that G/Γ has a one-dimensional vertical component, at the cost of weakening (4.7) to

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \tilde{\Lambda}_W(n))F(g(n)\Gamma) \right| \gg \exp\left(-\left(\log \frac{1}{\delta}\right)^{C_k^3}\right). \tag{4.8}$$

From [18, Proposition 7.1], we have the estimate[†]

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda_W(n) - \tilde{\Lambda}_W(n))F(g(n)\Gamma) \right| \ll M^{(2D)^{C_k}} (\tilde{q})^{-1/(2D)^{C_k}}. \tag{4.9}$$

Hence, if (4.7) holds, we may assume that either we have

$$\tilde{q} \leq \left(M / \exp\left(\left(\log \frac{1}{\delta}\right)^{C_k^3}\right) \right)^{(2D)^{C_k^2}} \tag{4.10}$$

[†] There, w was fixed to be $\exp((\log N)^{1/10})$, but this makes no difference in the argument.

or

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \Lambda_W(n)) F(g(n)\Gamma) \right| \gg \exp\left(-\left(\log \frac{1}{\delta}\right)^{C_k^3}\right). \tag{4.11}$$

Now, applying [12, Lemma 5.1] (which is a version of Vaughan’s identity), we may split $\Lambda - \tilde{\Lambda}_W$ and $\Lambda - \Lambda_W$ into type I, twisted type I and type II sums, where we may assume that (4.10) holds if twisted type I sums appear. Applying now [12, Propositions 5.1–5.4] (which are estimates for type I, twisted type I and type II sums involving nilsequences $F(g(n)\Gamma)$ with a vertical character F and where G/Γ has a one-dimensional vertical component, with the twisted type I sum estimate requiring (4.10)), the claim follows unless we have a factorisation

$$g(n) = \epsilon(n)g_1(n)\gamma(n)$$

for all $n \in \mathbb{Z}$, where g_1 takes values in a $\exp((\log(1/\delta))^{C_k^{10}})$ -rational subgroup of G whose step is $\leq s - 1$, ϵ is $(\exp((\log(1/\delta))^{C_k^{10}}), N)$ -smooth and γ is $\exp((\log(1/\delta))^{C_k^{10}})$ -rational.

Let Q be the period of γ . Then by the pigeonhole principle there exists an arithmetic progression $P \subset [N]$ of common difference Q and length $\asymp \exp(-(\log(1/\delta))^{C_k^{10}})N$ such that

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \tilde{\Lambda}_W(n)) F(g(n)\Gamma) \right| \leq \left| \frac{1}{|P|} \sum_{n \in P} (\Lambda(n) - \tilde{\Lambda}_W(n)) F(\epsilon_0 g_1(n) \gamma_0 \Gamma) \right| + O(\exp(-(\log(1/\delta))^{C_k^{10}}))$$

for some $\epsilon_0, \gamma_0 \in G$. Factorise $\gamma_0 = \{\gamma_0\}[\gamma_0]$ where $[\gamma_0] \in \Gamma$ and $|\psi(\{\gamma_0\})| \leq 1/2$, with ψ the Mal’cev coordinate map. Writing $g'(n) = \{\gamma_0\}^{-1} g_1(n) \{\gamma_0\}$ and $F_1 = F(\epsilon_0 \{\gamma_0\} \cdot) 1_P$, we conclude from (4.7) that

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \tilde{\Lambda}_W(n)) F_1(g'(n)\Gamma) \right| \gg \exp\left(-\left(\log \frac{1}{\delta}\right)^{C_k^{10}}\right), \tag{4.12}$$

where F_1 is defined on a nilmanifold of step at most $s - 1$. Iterating the arguments from (4.7) to (4.12) $\leq s$ times (and adjusting the value of δ each time), we conclude that

$$\left| \frac{1}{N} \sum_{n \leq N} (\Lambda(n) - \tilde{\Lambda}_W(n)) 1_{P'}(n) \right| \gg \exp\left(-\left(\log \frac{1}{\delta}\right)^{C_k^{10s}}\right)$$

for some arithmetic progression P' . But this contradicts Proposition 4.5. Hence the claim (4.5) follows. □

5 | MÖBIUS CORRELATIONS ALONG POLYNOMIAL PATTERNS

We are now ready to prove Theorem 1.3(1), which immediately implies Theorem 1.2 by the triangle inequality.

Proof of Theorem 1.3(1). Let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be polynomials satisfying the assumption of Theorem 1.3(1).

For some unimodular $\theta_j : \mathbb{Z} \rightarrow \mathbb{C}$, we can write

$$\mathbb{E}_{m \leq N} \left| \mathbb{E}_{n \leq Nd} \prod_{j=1}^k \mu(n + P_j(m)) \right| = \mathbb{E}_{m \leq N} \mathbb{E}_{n \leq Nd} \theta_1(m) \prod_{j=1}^k \mu(n + P_j(m))$$

and

$$\mathbb{E}_{n \leq Nd} \left| \mathbb{E}_{m \leq N} \prod_{j=1}^k \mu(n + P_j(m)) \right| = \mathbb{E}_{n \leq Nd} \mathbb{E}_{m \leq N} \theta_2(n) \prod_{j=1}^k \mu(n + P_j(m)).$$

Let $s \ll_{d,k} 1$ be as in Proposition 3.1. Now, by Proposition 3.1 (with $f_i = \mu$) and the assumption on the collection $\{P_1, \dots, P_k\}$, it suffices to show that

$$\|\mu\|_{U^s[M]} \ll_A (\log M)^{-A}.$$

This bound is precisely Proposition 4.1. □

The rest of the paper is devoted to the proof of Theorem 1.3(2) (and the deduction of Theorem 1.1 from it).

6 | A LOCAL-TO-GLOBAL PRINCIPLE FOR MEAN VALUES

The main result of this section is Proposition 6.1, giving a factorisation result for mean values of a certain class of arithmetic functions that are close to 1 on average, which, as a corollary, will be applied to correlations of the function Λ_w , given in (4.1). This can be viewed as a local-to-global principle, since the main term in the asymptotic factorises according to the contribution of individual primes to the arithmetic function.

Proposition 6.1 (A local-to-global principle for mean values). *Let $C \geq 1$ and $N \geq 2$. Let P be a non-constant polynomial with integer coefficients having no fixed prime divisor. Let P have degree at most C . Let $f : \mathbb{N}^2 \rightarrow \mathbb{C}$ be a function satisfying the following.*

- (1) *The function $n \mapsto f(p, n)$ is p -periodic for any prime p .*
- (2) *We have $f(p, n) = 1$ for $p > \exp((\log N)^{1/2})$.*
- (3) *We have $|f(p, n) - 1| \leq 1_{p|P(n)} + C/p$ for all primes $p \leq \exp((\log N)^{1/2})$ and all $n \in \mathbb{N}$.*

Then, we have

$$\mathbb{E}_{n \leq N} \prod_p f(p, n) = \prod_p \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} f(p, n) + O_C(\exp(-(\log N)^{1/2}/10)). \tag{6.1}$$

Proof. We may assume without loss of generality that $C \geq 2$. Let $w = \exp((\log N)^{1/2})$. Then, by assumption (2), it suffices to show that

$$\mathbb{E}_{n \leq N} \prod_{p \leq w} f(p, n) = \prod_{p \leq w} \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} f(p, n) + O_C(\exp(-(\log N)^{1/2}/10)). \tag{6.2}$$

Write $f(p, n) = 1 + a(p, n)$. Then, by assumptions (1) and (3), $n \mapsto a(p, n)$ is p -periodic for every $n \in \mathbb{N}$ and $|a(p, n)| \leq 1_{p|P(n)} + C/p$. Writing $W = \prod_{p \leq w} p$, we now obtain

$$\prod_{p \leq w} f(p, n) = \sum_{d|W} A(d, n),$$

where $A(d, n) := \prod_{p|d} a(p, n)$. The function $d \mapsto A(d, n)$ is multiplicative, and for squarefree d we have the crude estimate

$$|A(d, n)| \leq \prod_{p|d} \left(1 + \frac{C}{p}\right) \prod_{p|d} \frac{C}{p} \leq C^{\omega(d)} \prod_{p|d} \frac{1}{p} = \frac{C^{\omega(d)}(d, P(n))}{d}.$$

Let $s = 2(\log N)^{1/2}/5$. Then, we can split

$$\prod_{p \leq w} f(p, n) = \sum_{\substack{d|W \\ d \leq w^s}} A(d, n) + O\left(\sum_{\substack{d|W \\ d > w^s}} \frac{C^{\omega(d)}(d, P(n))}{d}\right).$$

Averaging over $n \leq N$ gives

$$\mathbb{E}_{n \leq N} \prod_{p \leq w} f(p, n) = \sum_{\substack{d|W \\ d \leq w^s}} \mathbb{E}_{n \leq N} A(d, n) + O\left(\mathbb{E}_{n \leq N} \sum_{\substack{d|W \\ d > w^s}} \frac{C^{\omega(d)}(d, P(n))}{d}\right). \tag{6.3}$$

Using Rankin’s trick and Mertens’s bound, for any $n \in [N]$, we can estimate

$$\begin{aligned} \sum_{\substack{d|W \\ d > w^s}} \frac{C^{\omega(d)}(d, P(n))}{d} &\leq e^{-s} \sum_{d|W} \frac{C^{\omega(d)}(d, P(n))}{d^{1-1/(\log w)}} \\ &= e^{-s} \prod_{p \leq w} \left(1 + \frac{C \cdot (p, P(n))}{p^{1-1/(\log w)}}\right) \\ &\leq e^{-s} \prod_{p \leq w} \left(1 + \frac{Ce \cdot (p, P(n))}{p}\right) \\ &\ll e^{-s} (\log w)^{Ce} (1 + Ce)^{\omega(P(n))} \\ &\leq e^{-s} (\log w)^{Ce} \tau(P(n))^B, \end{aligned} \tag{6.4}$$

where $B = B_C = (\log(1 + Ce))/(\log 2)$.

An elementary inequality due to Landreau [11] states that there exists K , depending only on C , such that for all $1 \leq n \leq (C + 1)N^{2C}$ we have

$$\tau(n)^B \leq K \sum_{\substack{d|n \\ d \leq N^{1/10}}} \tau(d)^K.$$

Using this and the fact (following from [4, Lemma 6.1]) that the congruence $P(x) \equiv 0 \pmod{p^r}$ has at most $\min\{Cp^{r-1}, Cp^{r(1-1/C)}\}$ solutions for any prime p and $r \geq 1$, we see that

$$\begin{aligned} \mathbb{E}_{n \leq N} \tau(P(n))^B &\ll_C \sum_{d \leq N^{1/10}} \tau(d)^K \mathbb{E}_{n \leq N} \mathbf{1}_{d|P(n)} \\ &\ll \sum_{d \geq 1} d^{-1/(\log N)} \tau(d)^K \mathbb{E}_{n \leq N} \mathbf{1}_{d|P(n)} \\ &\ll \prod_p \left(1 + O_C(p^{-1-1/\log N})\right) \\ &\ll_C (\log N)^{O_C(1)} \end{aligned} \tag{6.5}$$

by a standard bound for moments of the divisor function.

Combining (6.3), (6.4) and (6.5), we see that

$$\mathbb{E}_{n \leq N} \prod_{p \leq w} f(p, n) = \sum_{\substack{d|W \\ d \leq w^s}} \mathbb{E}_{n \leq N} A(d, n) + O_C(\exp(-(\log N)^{1/2}/10)). \tag{6.6}$$

Since $d \leq N^{2/5}$ in the above sum, we have

$$\mathbb{E}_{n \leq N} A(d, n) = \tilde{A}(d) + O(N^{-3/5+o(1)}),$$

where $\tilde{A}(d) := \mathbb{E}_{n \in \mathbb{Z}/d\mathbb{Z}} A(d, n)$. From the d -periodicity of $n \mapsto A(d, n)$ and the multiplicativity of $d \mapsto A(d, n)$, we see that \tilde{A} is a multiplicative function. Moreover, since $|A(d, n)| \leq C^{\omega(d)}(d, P(n))/d$, we have

$$|\tilde{A}(d)| \leq C^{\omega(d)} \prod_{p|d} (C + 1 - C/p) \leq (C + 1)^{2\omega(d)}.$$

We conclude that

$$\mathbb{E}_{n \leq N} \prod_{p \leq w} f(p, n) = \sum_{d|W} \tilde{A}(d) + O\left(\sum_{\substack{d|W \\ d > w^s}} \frac{(C + 1)^{2\omega(d)}}{d}\right) + O_C(\exp(-(\log N)^{1/2}/10)). \tag{6.7}$$

The sum in the error term can be bounded using, for example [4, Lemma 7.5], which yields

$$\sum_{\substack{d|W \\ d > w^s}} \frac{(C + 1)^{2\omega(d)}}{d} \ll_C e^{-s/4} \ll \exp(-(\log N)^{1/2}/10). \tag{6.8}$$

By the multiplicativity of \tilde{A} , the main term on the right-hand side of (6.7) becomes

$$\prod_{p \leq w} (1 + \tilde{A}(p)).$$

The claim (6.2) now follows by noting that $1 + \tilde{A}(p) = 1 + \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} a(p, n) = \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} f(p, n)$. \square

Corollary 6.2 (A factorisation lemma for correlations). *Let $k \in \mathbb{N}$ and $C \geq 1$. Let $N \geq 3$ and $w_1, \dots, w_k, Q \leq \exp((\log N)^{1/2}/100)$. Let $W_j = \prod_{p \leq w_j} p$. Let $a_1, \dots, a_k \in [-N^C, N^C]$ be integers. Then, we have*

$$\begin{aligned} & \mathbb{E}_{n \leq N} \prod_{j=1}^k \Lambda_{W_j}(Qn + a_j) \\ &= \prod_p \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_{(p, W_j)}(Qn + a_j) + O_{C,k}(\exp(-(\log N)^{1/2}/10)), \end{aligned} \tag{6.9}$$

where Λ_p for p prime is as in (1.3) and $\Lambda_1 := 1$.

Proof. This follows from Proposition 6.1, taking $f(p, n) = \prod_{j=1}^k 1_{(p, W_j) \mid Qn + a_j}$ there (and multiplying both sides of (6.9) by $\prod_p \prod_{j=1}^k \frac{\varphi((p, W_j))}{(p, W_j)}$). Properties (1)–(3) required of f in Proposition 6.1 are immediate (with $C \ll_k 1$ and $P(y) = \prod_{j=1}^k (Qy + a_j)$). \square

We end this section by proving a lemma that bounds the moments of a sum that arises as an error term when truncating the definition of Λ_W . This will be needed in Section 7.

Lemma 6.3. *Let $C \geq 1$. Let $w \geq 2$ and write $W = \prod_{p \leq w} p$. Then for any $N \geq 2$ and $s \geq 1$, we have*

$$\sum_{n \leq N} \left| \sum_{\substack{d \mid W \\ d > w^s}} \mu(d) 1_{d \mid n} \right|^C \ll_C N e^{-s} (\log w)^{2C+4}.$$

Proof. Let 2ℓ be the least even integer that is $\geq C$. Then, we have

$$\begin{aligned} \mathbb{E}_{n \leq N} \left| \sum_{\substack{d \mid W \\ d > w^s}} \mu(d) 1_{d \mid n} \right|^C &\leq \mathbb{E}_{n \leq N} \left| \sum_{\substack{d \mid W \\ d > w^s}} \mu(d) 1_{d \mid n} \right|^{2\ell} \\ &\leq \sum_{\substack{d_1, \dots, d_{2\ell} \mid W \\ d_1, \dots, d_{2\ell} > w^s}} \mathbb{E}_{n \leq N} \mu^2(d_1) \cdots \mu^2(d_{2\ell}) \prod_{j=1}^{2\ell} 1_{d_j \mid n} \\ &\leq \sum_{\substack{d_1, \dots, d_{2\ell} \mid W \\ d_1, \dots, d_{2\ell} > w^s}} \frac{\mu^2(d_1) \cdots \mu^2(d_{2\ell})}{[d_1, \dots, d_{2\ell}]} \\ &\leq e^{-s} \sum_{d_1, \dots, d_{2\ell} \mid W} \mu^2(d_1) \cdots \mu^2(d_{2\ell}) \frac{(d_1 \cdots d_{2\ell})^{1/(2\ell \log w)}}{[d_1, \dots, d_{2\ell}]}. \end{aligned}$$

Writing the previous expression as an Euler product and applying Mertens's theorem, we see that it is

$$\begin{aligned}
 &= e^{-s} \prod_{p \leq w} \left(1 + \sum_{(a_1, \dots, a_{2^\ell}) \in \{0,1\}^{2^\ell} \setminus \{\mathbf{0}\}} p^{(a_1 + \dots + a_{2^\ell}) / (2^\ell \log w) - \max_i \{a_i\}} \right) \\
 &\ll e^{-s} \prod_{p \leq w} \left(1 + (2^{2^\ell} - 1) \frac{e}{p} \right) \\
 &\ll_\ell e^{-s} (\log w)^{(2^{2^\ell} - 1)e} \\
 &\ll e^{-s} (\log w)^{2^{C+4}},
 \end{aligned}$$

as desired. □

7 | PRIME CORRELATIONS ALONG POLYNOMIAL PATTERNS

7.1 | Overview of proof

In this section, we prove Theorem 1.3(2) and deduce Theorem 1.1 from it. We begin with a brief sketch to indicate the main tools and why we need to work with the Siegel model, defined in Definition 4.4.

Suppose, for simplicity, that we want to estimate

$$S := \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) \Lambda(n + m^3) \Lambda(n + 2m^3) \tag{7.1}$$

up to an $O_A((\log N)^{-A})$ error. Let $w_3 = \exp((\log \log N)^{C_k})$ for a large constant C , and let $W_j = \prod_{p \leq w_j} p$. Writing $\Lambda = \tilde{\Lambda}_{W_3} + E_3$, we see that

$$\begin{aligned}
 S &= \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) \Lambda(n + m^3) \tilde{\Lambda}_{W_3}(n + 2m^3) \\
 &\quad + \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) \Lambda(n + m^3) E_3(n + 2m^3) \\
 &:= S_1 + S_2.
 \end{aligned}$$

Combining Propositions 3.1 and 4.6, we have $|S_2| \ll_A (\log N)^{-A}$. To estimate S_1 , we choose some w_2 that is moderately large in terms of w_3 ($w_2 = w_3^{(\log \log N)^C}$ works) and write $\Lambda = \tilde{\Lambda}_{W_2} + E_2$ to obtain

$$S_1 = \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) \tilde{\Lambda}_{W_2}(n + m^3) \tilde{\Lambda}_{W_3}(n + 2m^3)$$

$$\begin{aligned}
 &+ \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) E_2(n + m^3) \tilde{\Lambda}_{W_3}(n + 2m^3) \\
 &:= S_3 + S_4.
 \end{aligned}$$

Using Lemma 6.3, we see that $\tilde{\Lambda}_{W_3}(n)$ is up to negligible error (in the L^1 norm) equal to

$$(1 - \tilde{\chi}(n)n^{\tilde{\beta}-1}) \frac{W_3}{\varphi(W_3)} \sum_{\substack{t|W_3 \\ t \leq w_3^{(\log \log N)^2}}} \mu(t) 1_{t|n}.$$

Substituting this into S_4 and applying the triangle inequality, and denoting by $\tilde{q} \leq w_3$ the modulus of $\tilde{\chi}$, we see that for some $1 \leq a, t \leq \tilde{q}w_3^{(\log \log N)^2}$ we have

$$|S_4| \ll \frac{W_3}{\varphi(W_3)} \frac{\tilde{q}w_3^{(\log \log N)^2}}{N^4} \left| \sum_{m \leq N} \sum_{n \leq N^3} \Lambda(n) E_2(n + m^3) 1_{n+2m^3 \equiv a \pmod{t}} \right|.$$

Splitting the m variable into progressions modulo t and applying Propositions 3.1 and 4.6,[†] we conclude that $|S_4| \ll_A (\log N)^{-A}$. Performing again a similar splitting $\Lambda = \tilde{\Lambda}_{W_1} + E_1$ with w_1 large in terms of w_2 ($w_1 = w_2^{(\log \log N)^c}$ works) and arguing as above, we conclude that

$$S = \frac{1}{N^4} \sum_{m \leq N} \sum_{n \leq N^3} \tilde{\Lambda}_{W_1}(n) \tilde{\Lambda}_{W_2}(n + m^3) \tilde{\Lambda}_{W_3}(n + 2m^3) + O_A((\log N)^{-A}). \tag{7.2}$$

Now the remaining task is to evaluate the correlations of the Siegel model with error terms that save an arbitrary power of logarithm. Writing $\tilde{\Lambda}_{W_j} = \Lambda_{W_j} + E'_j$, the right-hand side of (7.2) splits into a sum of eight averages. The average involving each of the functions $\Lambda_{W_1}, \Lambda_{W_2}$ and Λ_{W_3} can be evaluated using Corollary 6.2, and it gives us our main term. All the other averages involving at least one copy of E'_j involve the correlations of a Dirichlet character of large conductor, and (after some work) they can be shown to be small by using the Weil bound.

If we tried to run the same argument with the simpler model Λ_W , we would run into serious trouble, since we can only save an arbitrary power of logarithm in the Gowers norm $\|\Lambda - \Lambda_{W_j}\|_{U^s[N]}$, whereas we would need to save at least $w_{j+1}^{C(\log \log N)^2}$, which is bigger since w_3 needs to be at least a large power of logarithm. The use of the Siegel model rectifies this, since we have a much better quasipolynomial bound on $\|\Lambda - \tilde{\Lambda}_{W_j}\|_{U^k[N]}$ as a function of w_j .

7.2 | Correlations of the Siegel model

In what follows, we will need the following simple lemma.

[†]Note that we are applying Proposition 3.1 with the coefficients of the polynomials involved bounded by $O(t^3)$. Hence, Proposition 3.1 involves a loss of $t^{O(1)}$, but we win back this loss thanks to the strong bound on the Gowers norm of E_2 .

Lemma 7.1. *Let $k \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed polynomials. Let $\beta_p(m), \beta'_p(n)$ be as in (1.5). For any $N \geq 2, m \in [N], n \in [N^d]$ and $A > 0$, we have*

$$\prod_{p > (\log N)^{2A}} \beta_p(m) = 1 + O_A((\log N)^{-A}), \quad \prod_{p > (\log N)^{2A}} \beta'_p(n) = 1 + O_A((\log N)^{-A}).$$

Proof. Note first that

$$\left(\frac{p}{p-1}\right)^k \left(1 - \frac{k}{p}\right) \leq \beta_p(m) \leq \begin{cases} \left(\frac{p}{p-1}\right)^k \left(1 - \frac{k}{p}\right), & p \nmid \prod_{1 \leq i < j \leq k} (P_i(m) - P_j(m)) \\ \left(\frac{p}{p-1}\right)^k, & p \mid \prod_{1 \leq i < j \leq k} (P_i(m) - P_j(m)). \end{cases}$$

Since $\prod_{p > (\log N)^{2A}} \left(\frac{p}{p-1}\right)^k \left(1 - \frac{k}{p}\right) = 1 + O_A((\log N)^{-2A})$, the claim for $\beta_p(m)$ follows by observing that for any $m \in [N]$ and $1 \leq i < j \leq k$ we have

$$\prod_{\substack{p \mid P_i(m) - P_j(m) \\ p > (\log N)^{2A}}} \left(\frac{p}{p-1}\right) \leq \exp \left(\sum_{\substack{p \mid P_i(m) - P_j(m) \\ p > (\log N)^{2A}}} \frac{1}{p-1} \right) \leq \exp \left(\frac{2 \log(|P_i(m) - P_j(m)| + 2)}{(\log N)^{2A} - 1} \right) \ll (\log N)^{-A}.$$

The argument for $\beta'_p(n)$ is analogous. □

We proceed to calculate the correlations of the Siegel model $\tilde{\Lambda}_W$ (given in Definition (4.4)) along polynomial progressions; this will give the main term in Theorem 1.3(2).

Lemma 7.2 (Siegel model correlations). *Let $k, d \in \mathbb{N}$ be fixed, and let $P_1, \dots, P_k \in \mathbb{Z}[y]$ be fixed polynomials of degree at most d . Let $N \geq 3, A > 0$ and $(\log N)^A \leq w_1, \dots, w_k \leq \exp((\log N)^{1/2}/(100k))$. Let $W_j = \prod_{p \leq w_j} P$.*

(1) *We have*

$$\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \tilde{\Lambda}_{W_j}(n + P_j(m)) = \prod_p \beta_p(m) + O_A((\log N)^{-A})$$

for all but $\ll_A N/(\log N)^A$ integers $m \in [N]$.

(2) *We have*

$$\mathbb{E}_{m \leq N} \prod_{j=1}^k \tilde{\Lambda}_{W_j}(n + P_j(m)) = \prod_p \beta'_p(n) + O_A((\log N)^{-A})$$

for all but $\ll_A N^d/(\log N)^A$ integers $n \in [N^d]$.

Proof. **Proof of part (1).** By definition, we can decompose

$$\tilde{\Lambda}_{W_j}(n) = \Lambda_{W_j}(n) - \Lambda_{W_j}(n)|n|^{\tilde{\beta}_j-1}\tilde{\chi}_j(|n|).$$

Substituting this into the left-hand side of the first statement, we see that it suffices to show that

$$\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \Lambda_{W_j}(n + P_j(m)) = \prod_p \beta_p(m) + O_A((\log N)^{-A}) \tag{7.3}$$

for all $m \in [N]$, and that, for any non-empty set $J \subset [k]$, we have

$$\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \Lambda_{W_j}(n + P_j(m)) \prod_{j \in J} |n + P_j(m)|^{\tilde{\beta}_j-1} \tilde{\chi}_j(|n + P_j(m)|) \ll_A (\log N)^{-A} \tag{7.4}$$

for all but $\ll_A N/(\log N)^A$ integers $m \in [N]$. Note that (7.3) follows directly from Corollary 6.2 in view of Lemma 7.1.

We are left with showing (7.4). Using the identity

$$y^{\tilde{\beta}-1} = 1 + \int_1^{CN^d} (\tilde{\beta} - 1)u^{\tilde{\beta}-2} 1_{u \leq y} du \tag{7.5}$$

for $1 \leq y \leq CN^d$ and exchanging the order of integration and summation, we reduce to showing that for all but $\ll_A N/(\log N)^A$ choices of $m \in [N]$ we have

$$\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \Lambda_{W_j}(n + P_j(m)) \prod_{j \in J} 1_{|n+P_j(m)| \geq u_j} \tilde{\chi}_j(|n + P_j(m)|) \ll_A (\log N)^{-A}$$

for all $u_j \in [1, N^d]$. Since the condition $|n + P_j(m)| \geq u_j$ is equivalent to n belonging to a union of two intervals and since $\tilde{\chi}_j(-n) = \tilde{\chi}_j(-1)\tilde{\chi}_j(n)$, we reduce to showing that for all but $\ll_A N/(\log N)^A$ choices of $m \in [N]$ we have

$$\mathbb{E}_{n \leq t} \prod_{j=1}^k \Lambda_{W_j}(n + P_j(m)) \prod_{j \in J} \tilde{\chi}_j(n + P_j(m)) \ll_A (\log N)^{-A} \tag{7.6}$$

for all $t \in [N^d/(\log N)^A, N^d]$.

Let

$$Q = [\tilde{q}_1, \dots, \tilde{q}_k],$$

where we recall that $\tilde{q}_j \leq w_j \leq \exp((\log N)^{1/2}/(100k))$ is the modulus of the character $\tilde{\chi}_j$. Then

$$(\log N)^A \ll_A Q \leq \exp((\log N)^{1/2}/100), \tag{7.7}$$

with the lower bound coming from Siegel’s bound. Splitting the n variable in (7.6) into progressions (mod Q) it suffices to show that for all $1 \leq a \leq Q$ and all but $\ll_A N/(\log N)^A$ integers $m \in [N]$ we have

$$\mathbb{E}_{a \leq Q} \prod_{j \in \mathcal{J}} \tilde{\chi}_j(a + P_j(m)) \mathbb{E}_{n \leq t/Q} \prod_{j=1}^k \Lambda_{W_j}(Qn + a + P_j(m)) \ll_A (\log N)^{-A}$$

for all $t \in [N^d/(\log N)^A, N^d]$. By Corollary 6.2, the inner average over n here is

$$\prod_p \mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_{(p, W_j)}(Qn + a + P_j(m)) + O_{d,k}(\exp(-(\log N)^{1/2})/10). \tag{7.8}$$

If $p \nmid Q$, the factor depending on p here is by a change of variables equal to

$$\mathbb{E}_{n \in \mathbb{Z}/p\mathbb{Z}} \prod_{j=1}^k \Lambda_{(p, W_j)}(n + P_j(m)),$$

and if instead $p \mid Q$ it is equal to

$$\prod_{j=1}^k \Lambda_{(p, W_j)}(a + P_j(m)).$$

Now, separating the contributions of $p \mid Q$ and $p \nmid Q$ in (7.8), it suffices to show that

$$\mathbb{E}_{a \leq Q} \prod_{j \in \mathcal{J}} \tilde{\chi}_j(a + P_j(m)) \prod_{p \mid Q} \prod_{j=1}^k 1_{a+P_j(m) \not\equiv 0 \pmod{(p, W_j)}} \ll_A (\log N)^{-A} \tag{7.9}$$

for all but $\ll_A N/(\log N)^A$ integers $m \in [N]$.

For any $j \in [k]$ and any prime power p^r , write $\tilde{\chi}_{j,p^r}$ for the character (mod (p^r, \tilde{q}_j)) induced by $\tilde{\chi}_j$ (mod \tilde{q}_j). Then, by the Chinese remainder theorem, (7.9) factorises as

$$\prod_{p^r \mid \mid Q} \mathbb{E}_{a \in \mathbb{Z}/p^r\mathbb{Z}} \prod_{j \in \mathcal{J}} \tilde{\chi}_{j,p^r}(a + P_j(m)) \prod_{j=1}^k 1_{a+P_j(m) \not\equiv 0 \pmod{(p, W_j)}}. \tag{7.10}$$

Since $\tilde{\chi}_j$ are primitive real characters, their moduli are of the form $2^{b_j} q'_j$ with q'_j squarefree and $b_j \leq 3$. Hence, for any prime $p > 2$ the average in (7.10) equals to

$$\mathbb{E}_{a \in \mathbb{Z}/p\mathbb{Z}} \prod_{j \in \mathcal{J}} \tilde{\chi}_{j,p}(a + P_j(m)) + O_{d,k}(1/p). \tag{7.11}$$

From the Weil bound [10, Corollary 11.24], the average in (7.11) is $\leq Kp^{-1/2}$ in modulus for some constant $K = K_{d,k}$ unless there exist $i \neq j$ such that $P_i(m) - P_j(m) \equiv 0 \pmod{p}$. Hence, for some $1 \leq i < j \leq k$, the expression (7.10) is

$$\ll \left(\prod_{p|Q} K p^{-1/2} \right) \left(\prod_{\substack{p|P_i(m)-P_j(m) \\ p|Q}} p^{1/2} \right)^{k^2} \ll_{d,k} Q^{-1/2+o(1)} \left(\prod_{\substack{p|P_i(m)-P_j(m) \\ p|Q}} p^{1/2} \right)^{k^2}, \tag{7.12}$$

since $Q/2^b$ is squarefree for some $b \leq 3$. From (7.7), we see that $Q^{-1/10} \ll_A (\log N)^{-A}$, so it suffices to show that for any non-zero polynomial $R \in \mathbb{Z}[y]$ for all but $\ll_{A,R} N/(\log N)^A$ integers $m \in [N]$ we have

$$\prod_{\substack{p|R(m) \\ p|Q}} p \leq Q^{1/(3k^2)}. \tag{7.13}$$

Note that, by Lagrange’s theorem, for any polynomial $R \in \mathbb{Z}[y]$ of degree $d \in \mathbb{N}$, the congruence $R(m) \equiv 0 \pmod{p}$ has $\leq d$ solutions for all primes p that are large enough in terms of R . Now, by Markov’s inequality, the number of $m \in [N]$ violating (7.13) is

$$\begin{aligned} &\leq Q^{-1/(3k^2)} \sum_{m \leq N} (R(m), Q) \\ &\leq Q^{-1/(3k^2)} \sum_{t|Q} t \sum_{\substack{m \leq N \\ t|R(m)}} 1 \\ &\ll_R Q^{-1/(3k^2)} \sum_{t|Q} d^{\Omega(t)} N \\ &\leq Q^{-1/(3k^2)} \tau(Q) d^{\Omega(Q)} N \\ &\ll Q^{-1/(3k^2)+o(1)} N, \end{aligned}$$

since $\Omega(Q) \leq \omega(Q) + O(1) \ll (\log Q)/(\log \log Q)$ by the fact that $Q/2^b$ is squarefree for some $b \leq 3$. By (7.7), we conclude that the number of $m \in [N]$ not satisfying (7.13) is $\ll_{A,R} N/(\log N)^A$. Hence, these integers $m \in [N]$, may be included in the exceptional set of m .

Proof of part (2). The proof proceeds the same way as for Lemma 7.2(1), up to a swapping of the averaging variable, until (in analogy with (7.9)) we are left with showing that

$$\mathbb{E}_{b \leq Q} \prod_{j \in J} \tilde{\chi}_j(n + P_j(b)) \prod_{p|Q} \prod_{j=1}^k 1_{n+P_j(b) \not\equiv 0 \pmod{(p, W_j)}} \ll_A (\log N)^{-A} \tag{7.14}$$

for all but $\ll_A N^d/(\log N)^A$ integers $n \in [N^d]$.

As in the proof of part (1), for any $j \in [k]$ and any prime power p^r , write $\tilde{\chi}_{j,p^r}$ for the character $\pmod{(p^r, \tilde{q}_j)}$ induced by $\tilde{\chi}_j \pmod{\tilde{q}_j}$. Then, by the Chinese remainder theorem, (7.9) factorises as

$$\prod_{p^r || Q} \mathbb{E}_{b \in \mathbb{Z}/p^r \mathbb{Z}} \prod_{j \in J} \tilde{\chi}_{j,p^r}(n + P_j(b)) \prod_{j=1}^k 1_{n+P_j(b) \not\equiv 0 \pmod{(p, W_j)}}. \tag{7.15}$$

Since $\tilde{\chi}_j$ are primitive real characters, their moduli are of the form $2^{b_j} q'_j$ with q'_j squarefree and $b_j \leq 3$. Hence, for any prime $p > 2$ the average in (7.15) equals to

$$\mathbb{E}_{b \in \mathbb{Z}/p\mathbb{Z}} \prod_{j \in \mathcal{J}} \tilde{\chi}_{j,p}(n + P_j(b)) + O_{d,k}(1/p), \tag{7.16}$$

where the error term comes from the fact that $n + P_j(b) \equiv 0 \pmod{p}$ happens for density $\ll_{d,k} 1/p$ of $b \in \mathbb{Z}/p\mathbb{Z}$.

By the Weil bound, (7.16) is $\leq Kp^{-1/2}$ in modulus for some constant $K = K_{d,k}$ unless the polynomial $f_n(y) := (P_1(y) + n) \cdots (P_k(y) + n)$ is of the form $cg(y)^2$ modulo p for some $c \in \mathbb{Z}/p\mathbb{Z}$ and some polynomial g . This can only happen if p divides the discriminant $\Delta(n)$ of f_n . The discriminant Δ is some polynomial with integer coefficients. Hence, (7.15) becomes

$$\ll \left(\prod_{p|Q} Kp^{-1/2} \right) \left(\prod_{\substack{p|\Delta(n) \\ p|Q}} p^{1/2} \right) \ll_{d,k} Q^{-1/2+o(1)} \left(\prod_{\substack{p|\Delta(n) \\ p|Q}} p^{1/2} \right).$$

Now we may use (7.13) to conclude that this is $\ll Q^{-1/10} \ll_A (\log N)^{-A}$ for all but $\ll_A N^d/(\log N)^A$ integers $n \in [N^d]$. □

7.3 | Proof of Theorem 1.3(2)

We are now ready to prove Theorem 1.3(2) and to deduce Theorem 1.1 from it. We begin with the latter task.

Proof of Theorem 1.1 assuming Theorem 1.3(2). Using the triangle inequality, we only need to prove that

$$\mathbb{E}_{m \leq N} \prod_p \beta_p(m) = \prod_p \beta_p + O_A((\log N)^{-A}).$$

Since $\beta_p(m), \beta_p = 1 + O_{d,k}(1/p^2)$ and $\beta_p = \mathbb{E}_{m \in \mathbb{Z}/p\mathbb{Z}} \beta_p(m)$, it in fact suffices to show that

$$\mathbb{E}_{m \leq N} \prod_{p \leq \exp((\log N)^{1/2})} \beta_p(m) = \prod_{p \leq \exp((\log N)^{1/2})} \mathbb{E}_{m \in \mathbb{Z}/p\mathbb{Z}} \beta_p(m) + O_A((\log N)^{-A}).$$

This estimate follows from Proposition 6.1 with $f(p, n) = \beta_p(n)1_{p \leq \exp((\log N)^{1/2})}$; properties (1)–(3) required by the proposition are easily verified (with $C \ll_k 1$ and with P being $\prod_{1 \leq i < j \leq k} (P_i - P_j)$). □

Proof of Theorem 1.3(2). Shifting the collection $\{P_1, \dots, P_k\}$ by a constant if necessary, we may assume that none of the P_i are the zero polynomial.

For $j \in [k]$, let

$$w_j := \exp((\log \log N)^{c^{k+1-j}}), \tag{7.17}$$

where $C = C(d, k)$ is a large enough constant. Then

$$w_{j+1} = \exp((\log w_j)^{1/C}),$$

which will eventually allow us to apply Proposition 4.6.

For some unimodular $\theta_1, \theta_2 : \mathbb{Z} \rightarrow \mathbb{C}$, we can write

$$\begin{aligned} & \mathbb{E}_{m \leq N} \left| \mathbb{E}_{n \leq N^d} \prod_{j=1}^k \Lambda(n + P_j(m)) - \prod_p \beta_p(m) \right| \\ &= \mathbb{E}_{m \leq N} \theta_1(m) \left(\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \Lambda(n + P_j(m)) - \prod_p \beta_p(m) \right) \end{aligned} \tag{7.18}$$

and

$$\begin{aligned} & \mathbb{E}_{n \leq N^d} \left| \mathbb{E}_{m \leq N} \prod_{j=1}^k \Lambda(n + P_j(m)) - \prod_p \beta'_p(n) \right| \\ &= \mathbb{E}_{n \leq N^d} \theta_2(n) \left(\mathbb{E}_{m \leq N} \prod_{j=1}^k \Lambda(n + P_j(m)) - \prod_p \beta'_p(n) \right). \end{aligned} \tag{7.19}$$

From Lemma 7.1 and the estimates $\beta_p(m), \beta'_p(m) = 1 + O_k(1/p)$, we see that the expressions inside the brackets here are $\ll (\log N)^{O_k(1)}$.

We split the average on the right-hand side of (7.18), (7.19) into a sum of 2^k different averages by substituting

$$\Lambda = \tilde{\Lambda}_{W_j} + E_j$$

in the j -th factor in (7.18), where $\tilde{\Lambda}_{W_j}$ is as in Definition 4.4. Note for later use that

$$|E_j(n)| \ll \log |n| + \log N. \tag{7.20}$$

Then it suffices to show that

$$\mathbb{E}_{n \leq N^d} \prod_{j=1}^k \tilde{\Lambda}_{W_j}(n + P_j(m)) = \prod_p \beta_p(m) + O_A((\log N)^{-A}) \tag{7.21}$$

for all but $\ll_A N/(\log N)^A$ integers $m \in [N]$, and that

$$\mathbb{E}_{m \leq N} \prod_{j=1}^k \tilde{\Lambda}_{W_j}(n + P_j(m)) = \prod_p \beta'_p(m) + O_A((\log N)^{-A}) \tag{7.22}$$

for all but $\ll_A N^d/(\log N)^A$ integers $n \in [N^d]$, and that for any $0 \leq \ell \leq k - 1$ and any unimodular $\theta_1, \theta_2 : \mathbb{Z} \rightarrow \mathbb{C}$ we have

$$\mathbb{E}_{n \leq N^d} \mathbb{E}_{m \leq N} \theta_1(m) \theta_2(n) \prod_{j=1}^{k-\ell-1} \Lambda(n + P_j(m)) E_{k-\ell}(n + P_{k-\ell}(m)) \prod_{j=k-\ell+1}^k \tilde{\Lambda}_{W_j}(n + P_j(m)) \tag{7.23}$$

$$\ll_A (\log N)^{-A}.$$

The estimates (7.21), (7.22) follow directly from Lemma 7.2. Hence, the remaining task is proving (7.23).

Suppose first that $\ell = 0$. In this case, the claim (7.23) simplifies as

$$\mathbb{E}_{n \leq N^d} \mathbb{E}_{m \leq N} \theta_1(m) \theta_2(n) \prod_{j=1}^{k-1} \Lambda(n + P_j(m)) \cdot E_k(n + P_k(m)) \ll_A (\log N)^{-A}.$$

By Proposition 3.1 (and the bounds $\Lambda(n) \ll \log |n|$ and (7.20)), this holds unless

$$\|E_k\|_{U^s[CN^d]} \gg (\log N)^{-O_{A,d}(1)}$$

for some $1 \leq s, C \ll_{P_1, \dots, P_k} 1$. But since $E_k = \Lambda - \tilde{\Lambda}_{W_k}$, this contradicts Proposition 4.6 (recalling (7.17) and that $C = C(d, k)$ is large enough there). We may suppose from now on that $1 \leq \ell \leq k - 1$.

Let

$$r = (\log \log N)^2.$$

For $W, V \geq 1$, define the truncated function

$$\Lambda_{W, \leq V}(n) := \frac{W}{\varphi(W)} \sum_{\substack{d|W \\ d \leq V}} \mu(d) 1_{d|n}; \tag{7.24}$$

then by the Möbius inversion formula we have $\Lambda_W(n) = \Lambda_{W, \leq V}(n)$ for any $V \geq n$. Write

$$\tilde{\Lambda}_{W_j}(n) = \Lambda_{W_j, \leq w_j^r}(n) \cdot (1 - |n|^{\tilde{\beta}-1} \tilde{\chi}(|n|)) + \mathcal{E}_j(n). \tag{7.25}$$

Now split (7.23) into a sum of 2^ℓ different averages by writing (7.25) in the j -th factor of (7.23) for $k - \ell + 1 \leq j \leq k$. Using the bounds $|\Lambda_{W_j}(n)|, |\mathcal{E}_j(n)| \ll_{C,d} \log N$ for $|n| \leq CN^d$ and

$$\mathbb{E}_{n \leq CN^d} |\Lambda_{W_j, \leq w_j^r}(n)|^k \leq \left(\frac{W_j}{\varphi(W_j)} \right)^k \mathbb{E}_{n \leq CN^d} \tau(n)^k \ll_C (\log N)^{2k-1+o(1)}$$

and Hölder's inequality, it suffices to show that

$$\mathbb{E}_{n \leq N^d} \mathbb{E}_{m \leq N} \theta_1(m) \theta_2(n) \prod_{j=1}^{k-\ell-1} \Lambda(n + P_j(m)) \cdot E_{k-\ell}(n + P_{k-\ell}(m))$$

$$\begin{aligned} &\cdot \prod_{j=k-\ell+1}^k \Lambda_{W_j, \leq w_j^r}(n + P_j(m))(1 - |n + P_j(m)|^{\tilde{\beta}_j-1} \tilde{\chi}_j(|n + P_j(m)|)) \\ &\ll_A (\log N)^{-A}, \end{aligned} \tag{7.26}$$

and that for any $j \in [k]$ we have

$$\mathbb{E}_{n \leq N^d} \mathbb{E}_{m \leq N} |\mathcal{E}_j(n + P_j(m))|^k \ll_A (\log N)^{-A}. \tag{7.27}$$

The bound (7.27) follows by making the change of variables $n' = n + P_j(m)$ and applying Lemma 6.3 with $s = r = (\log \log N)^2$. The remaining task is then to prove (7.26).

For proving (7.26), our aim is to split m and n into progressions in (7.26) to make the last product over j there constant, after which Proposition 3.1 can be applied. Expanding out the definition of $\Lambda_{W_j, \leq w_j^r}$, the estimate (7.26) reduces to showing that

$$\begin{aligned} &\mathbb{E}_{n \leq N^d} \mathbb{E}_{m \leq N} \theta_1(m) \theta_2(n) \prod_{j=1}^{k-\ell-1} \Lambda(n + P_j(m)) \cdot E_{k-\ell}(n + P_{k-\ell}(m)) \\ &\cdot \prod_{j=k-\ell+1}^k 1_{d_j | n + P_j(m)} (1 - |n + P_j(m)|^{\tilde{\beta}_j-1} \tilde{\chi}_j(|n + P_j(m)|)) \\ &\ll_A \frac{(\log N)^{-A}}{[d_{k-\ell+1}, \dots, d_k]} \end{aligned}$$

for any natural numbers $d_j \leq w_j^r$.

Using (7.5) and exchanging the order of integration and summation, and splitting the n variable into short segments, we reduce to

$$\begin{aligned} &\mathbb{E}_{m \leq N} \theta_1(m) \theta_2(n) \mathbb{E}_{n \in I} \prod_{j=1}^{k-\ell-1} \Lambda(n + P_j(m)) \cdot E_{k-\ell}(n + P_{k-\ell}(m)) \prod_{j=k-\ell+1}^k 1_{d_j | n + P_j(m)} \tilde{\chi}_j(n + P_j(m)) \\ &\ll_A \frac{(\log N)^{-A}}{[d_{k-\ell+1}, \dots, d_k]} \end{aligned} \tag{7.28}$$

for any interval $I \subset [N^d/(\log N)^A, N^d]$ of length $N^d/(\log N)^{2A}$ and all unimodular weights $\theta_1, \theta_2 : \mathbb{Z} \rightarrow \mathbb{C}$.

Let

$$Q' = \prod_{j=k-\ell+1}^k \tilde{q}_j, \quad D = [d_{k-\ell+1}, \dots, d_k].$$

Splitting the averaging variables in (7.28) into residue classes (mod DQ'), it suffices to show that for any $1 \leq a_1, a_2 \leq DQ'$ we have

$$\mathbb{E}_{n \in I} \mathbb{E}_{m \leq N/(DQ')} 1_{n + P_1(DQ'm + a_2) \equiv a_1 + P_1(a_2) \pmod{DQ'}}$$

$$\begin{aligned} & \cdot \theta_1(m)\theta_2(n) \prod_{j=1}^{k-\ell-1} \Lambda(n + P_j(DQ' m + a_2)) E_{k-\ell}(n + P_{k-\ell}(DQ' m + a_2)) \\ & \ll_A \frac{(\log N)^{-A}}{DQ'}. \end{aligned} \tag{7.29}$$

By splitting the n average into shorter segments if necessary, it suffices to show that for all $1 \leq a_1, a_2 \leq DQ'$ and $x \in [N^d]$ we have

$$\begin{aligned} & \mathbb{E}_{n \leq (N/(DQ'))^d} \mathbb{E}_{m \leq N/(DQ')} \mathbf{1}_{x+n+P_1(DQ' m+a_2) \equiv a_1+P_1(a_2) \pmod{DQ'}} \\ & \cdot \theta_1(m)\theta_2(n) \prod_{j=1}^{k-\ell-1} \Lambda(x + n + P_j(DQ' m + a_2)) E_{k-\ell}(x + n + P_{k-\ell}(DQ' m + a_2)) \\ & \ll_A \frac{(\log N)^{-A}}{DQ'}. \end{aligned} \tag{7.30}$$

We now apply Proposition 3.1 to the collection of $k - \ell$ polynomials $(P_1(DQ'y + a_2), \dots, P_{k-\ell}(DQ'y + a_2))$. Let $m_1, \dots, m_{k-\ell}$ be the leading coefficients of these polynomials. Then $m_i \asymp_{P_1, \dots, P_k} (DQ')^d$ for all $1 \leq i \leq k - \ell$. In particular, this means that $m_i \asymp_{P_1, \dots, P_k} c_{k-\ell}$ for all $1 \leq i \leq k - \ell$. In addition, by the assumption on P_i , we know that the m_i are distinct.

Now, by Proposition 3.1 (and the bounds $\Lambda(n) \ll \log |n|$ and (7.20)) we see that (7.30) holds unless

$$\|E_{k-\ell}(x + \cdot)\|_{U^s[-C(N/(DQ'))^d, C(N/(DQ'))^d]} \gg (\log N)^{-O_{A,d,k}(1)} (DQ')^{-O_{d,k}(1)} \tag{7.31}$$

for some $1 \leq s \ll_{d,k} 1$ and $1 \leq C \ll_{P_1, \dots, P_k} 1$. By a change of variables, the previous inequality implies that

$$\|E_{k-\ell} \cdot \mathbf{1}_{[-x-C(N/(DQ'))^d, -x+C(N/(DQ'))^d]}\|_{U^s[-CN^d, CN^d]} \gg (\log N)^{-O_{A,d,k}(1)} (DQ')^{-O_{d,k}(1) - \frac{s+1}{2s}}.$$

Let $B = B(d, k)$ be large. Using Vinogradov’s Fourier expansion (Lemma 2.1) to approximate $\mathbf{1}_{[-x, C(N/(DQ'))^d - x]}$ by a trigonometric polynomial with $O((DQ')^B)$ terms and with coefficients bounded by $O(1)$, we conclude that

$$\|E_{k-\ell}\|_{U^s[-CN^d, CN^d]} \gg (\log N)^{-O_{A,d,k}(1)} (DQ')^{-O_{d,k}(1)}.$$

Since $E_{k-\ell}$ is an even function, this yields

$$\|E_{k-\ell}\|_{U^s[CN^d]} \gg (\log N)^{-O_{A,d,k}(1)} (DQ')^{-O_{d,k}(1)}.$$

Let c_k be the constant in Proposition 4.6. We may assume that $C = C(d, k)$ in (7.17) is chosen large enough in terms of c_k . Then we have $DQ' \ll w_{k-\ell+1}^{2r} < \exp((\log w_{k-\ell})^{c_k/2})$. Now, recalling that $E_{k-\ell} = \Lambda - \tilde{\Lambda}_{W_{k-\ell}}$ we obtain a contradiction with Proposition 4.6. Hence, (7.31) cannot hold, so the proof is complete. \square

ACKNOWLEDGEMENTS

LM was supported by the Swedish Research Council (VR) Grant 2021-04526. JT was supported by funding from European Union's Horizon Europe research and innovation programme under Marie Skłodowska-Curie Grant agreement no. 101058904 and ERC Grant agreement no. 101162746. MW was supported by the Academy of Finland Grant no. 333707. Part of this material is based upon work supported by the Swedish Research Council under Grant no. 2021-06594 while the authors were in residence at Institut Mittag-Leffler in Djursholm, Sweden, during Spring 2024.

JOURNAL INFORMATION

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. S. Baier and L. Zhao, *Primes in quadratic progressions on average*, Math. Ann. **338** (2007), no. 4, 963–982.
2. F. Balestrieri and N. Rome, *Average Bateman-Horn for Kummer polynomials*, Acta Arith. **207** (2023), no. 4, 315–350.
3. P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367.
4. T. Browning, E. Sofos, and J. Teräväinen, *Bateman-Horn, polynomial Chowla and the Hasse principle with probability 1*, arXiv:2212.10373, 2022.
5. J. Friedlander and H. Iwaniec, *Opera de Cribro*, American Mathematical Society Colloquium Publications, vol. 57, American Mathematical Society, Providence, RI, 2010.
6. B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547.
7. B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), no. 3, 1753–1850.
8. B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) **175** (2012), no. 2, 541–566.
9. B. Green, T. Tao, and T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$ -norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372.
10. H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.
11. B. Landreau, *A new proof of a theorem of van der Corput*, Bull. Lond. Math. Soc. **21** (1989), no. 4, 366–368.
12. J. Leng, *Efficient equidistribution of nilsequences*, arXiv:2312.10772, 2024.
13. J. Leng, A. Sah, and M. Sawhney, *Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$ -norm*, arXiv:2402.17994, 2024.
14. F. Manners, *Quantitative bounds in the inverse theorem for the Gowers U^{s+1} -norms over cyclic groups*, arXiv:1811.00718, 2018.
15. K. Matomäki, M. Radziwiłł, T. Tao, J. Teräväinen, and T. Ziegler, *Higher uniformity of bounded multiplicative functions in short intervals on average*, Ann. of Math. (2) **197** (2023), no. 2, 739–857.
16. S. Peluse, *Bounds for sets with no polynomial progressions*, Forum Math. Pi **8** (2020), e16, 55.
17. J. Pintz, *Elementary methods in the theory of L-functions. V. The theorems of Landau and Page*, Acta Arith. **32** (1977), no. 2, 163–171.
18. T. Tao and J. Teräväinen, *Quantitative bounds for Gowers uniformity of the Möbius and von Mangoldt functions*, J. Eur. Math. Soc. **27** (2025), no. 4, 1321–1384.
19. T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Math. **201** (2008), no. 2, 213–305.
20. T. Tao and T. Ziegler, *Polynomial patterns in the primes*, Forum Math. Pi **6** (2018), e1, 60.

21. I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Dover Publications, Inc., Mineola, NY, 2004. Translated from the Russian, revised and annotated by K. F. Roth and Anne Davenport, Reprint of the 1954 translation.
22. N. H. Zhou, *Primes in higher-order progressions on average*, Int. J. Number Theory **14** (2018), no. 7, 1943–1959.