ON ANQIE ENTROPY AND MÖBIUS DISJOINTNESS OF A CLASS OF EXPONENTIAL FUNCTIONS

WEICHEN GU AND FEI WEI

Abstract. We show that a certain class of exponential functions have zero anqie entropy and study the Möbius disjointness of them. We show a result on the correlation of the Möbius function with polynomial phases in almost all short intervals and arithmetic progressions simultaneously. As an application, we give an equivalent condition of Sarnak’s Möbius disjointness conjecture in terms of the correlation between the Möbius function and arithmetic functions realized in any deterministic flow in short intervals on average.

1. Introduction

Let \( \mathbb{N} = \{0, 1, 2, \ldots \} \) denote the set of natural numbers. Functions from \( \mathbb{N} \) to \( \mathbb{C} \) are called arithmetic functions. Many problems in number theory are related to arithmetic (additive or multiplicative) structures of \( \mathbb{N} \). They can often be reformulated in terms of properties of arithmetic functions. For example, let \( \mu(n) : \mathbb{N} \to \{-1, 0, 1\} \) be the Möbius function, that is, \( \mu(n) = 0 \) when \( n \) is not square free (i.e., divisible by a nontrivial square), \( 1 \) when \( n \) is the product of even number of distinct primes and \( -1 \) when \( n \) is an odd number of products. It is well known that the Prime Number Theorem is equivalent to \( \sum_{n \leq x} \mu(n) = o(x) \); the Riemann hypothesis holds if and only if \( \sum_{n \leq x} \mu(n) = o(x^{1/2 + \epsilon}) \), for any \( \epsilon > 0 \). Here, the notation “\( f(x) = o(g(x)) \)” means \( \lim_{x \to \infty} f(x)/g(x) = 0 \). An arithmetic function \( f(n) \) is said to be disjoint from another one \( g(n) \) if \( \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(n)g(n) = 0 \). Sarnak ([Sar09], see also [LS15]) conjectured that the Möbius function is disjoint from all arithmetic functions realized in any topological dynamical system with zero topological entropy. More specifically,

**Conjecture 1.1** (Sarnak’s Möbius disjointness conjecture (SMDC)). Let \( \mathcal{X} \) be a compact Hausdorff space and \( T \) a continuous map on \( \mathcal{X} \) with zero topological entropy, then

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)F(T^{n}x_0) = 0
\]

for any \( x_0 \in \mathcal{X} \) and \( F \in C(\mathcal{X}) \).

A topological dynamical system with zero topological entropy is also known as a deterministic flow. There are a lot of progress made on Conjecture 1.1, while it is still open in its most general form. We simply refer to survey papers [FKPL18], [KPL20] for the progress in this area, and discuss only the results that are more related to this paper.
To use tools from operator algebra to study Sarnak’s conjecture, in the survey paper [Ge16], Ge introduced the notion of anqie entropy, which is given by the topological entropy of a topological dynamical system with a transitive map corresponding to the addition in $\mathbb{N}$. The anqie entropy of a bounded arithmetic function can be used to measure its complexity. This point is more clear for functions which take only finitely many values. In this case, the anqie entropy of such a function $f$ is determined by the number of different $J$-blocks appearing in the sequence $\{f(n)\}_{n=0}^{\infty}$. Specifically, let $\mathcal{B}_J(f)$ denote the set of all $J$-blocks occurring in $f$, i.e., $\mathcal{B}_J(f) = \{(f(n), f(n+1), \ldots, f(n+J-1)) : n \geq 0\}$, then the anqie entropy of $f(n)$ equals ([Wei18a, Lemma 6.1])

$$\lim_{J \to \infty} \frac{\log |\mathcal{B}_J(f)|}{J},$$

(1.1)

where $|\mathcal{B}_J(f)|$ is the cardinality of the set $\mathcal{B}_J(f)$. So if $f(n)$ is a random function, then the anqie entropy may attain to the maximal value $\log |f(\mathbb{N})|$. If $f(n)$ is a simple function, such as a periodic function, then the anqie entropy is zero.

The anqie entropy has nice properties. Here we list two properties which will be used in our paper. One is about the algebraical operations (see Proposition 1.2) and the other one is about the lower semi-continuity (see Theorem 1.3 below). These results are claimed in [Ge16, Section 4]. We also refer readers to [Wei18a] for proofs. Throughout this paper, we use $\mathcal{E}(f)$ to denote the anqie entropy of any bounded arithmetic function $f(n)$.

**Proposition 1.2.** For any bounded arithmetic functions $f, g$ and a continuous function $\phi(z)$ in $\mathbb{C}$, we have

$$|\mathcal{E}(f) - \mathcal{E}(g)| \leq \mathcal{E}(f \pm g) \leq \mathcal{E}(f) + \mathcal{E}(g),$$

$$\mathcal{E}(f \cdot g) \leq \mathcal{E}(f) + \mathcal{E}(g), \quad \mathcal{E}(\phi(f)) \leq \mathcal{E}(f).$$

**Theorem 1.3.** If $\{f_N(n)\}_{N=0}^{\infty}$ is a sequence of bounded arithmetic functions converging to $f(n)$ uniformly with respect to $n \in \mathbb{N}$, then $\lim \inf_{N \to \infty} \mathcal{E}(f_N) \geq \mathcal{E}(f)$.

In the language of anqie entropy, Sarnak’s conjecture is equivalent to that the M"obius function is disjoint from all bounded arithmetic functions with zero anqie entropy.

**Theorem 1.4.** (See [Ge16, Section 6]) Sarnak’s M"obius disjointness conjecture holds if and only if for any bounded arithmetic function $f(n)$ with zero anqie entropy,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)f(n) = 0.$$

In this paper, we shall prove that exponential functions $e(f(n))$ with $f(n)$ satisfying equation (1.3) below have zero anqie entropy. So by the above theorem, SMDC implies that these functions are disjoint from the M"obius function. Moreover, we also give some necessary conditions of SMDC in terms of the correlation between $\mu(n)$ and $e(f(n))$ in short arithmetic progressions. As an application, we obtain an equivalent condition of Sarnak’s conjecture that M"obius does not correlate with any arithmetic function realized in any zero topological entropy dynamics in short intervals on average (see Theorem 1.10 below). After that, we investigate the M"obius
disjointness of \( e(f(n)) \) unconditionally. In this process, we show a result on the correlation of the M"obius function with polynomial phases in almost all short arithmetic progressions (see Theorem 1.14 below). In the following, we expand the above introduction into three subsections. Before more statements, we list some notation.

1.1. **Notation.** Throughout, for any real-valued arithmetic function \( f(n) \), we denote by \( e(f(n)) \) the exponential function \( \exp(2\pi i f(n)) \), where \( i \) is the imaginary unit. When we refer to the exponential function of \( f(n) \) in this paper, we always mean \( e(f(n)) \). Sometimes we call \( e(f(n)) \) an \( f(n) \) phase. If \( S \) is a statement, we use \( 1_S \) to denote the indicator, that is \( 1_S = 1 \) when \( S \) is true and \( 1_S = 0 \) when \( S \) is false. We also denote \( 1_A(n) = 1_{n \in A} \) for any subset \( A \) of \( \mathbb{N} \). For any finite set \( C \), we use \( |C| \) to denote the carnality of \( C \). When \( x \in \mathbb{R} \), we use \( \{ x \} \) and \( \lfloor x \rfloor \) to denote the fractional part and the integer part functions, respectively. We use \( \| x \|_{\mathbb{R}/\mathbb{Z}} \) to denote the distance between \( x \) and the set \( \mathbb{Z} \), i.e., \( \| x \|_{\mathbb{R}/\mathbb{Z}} = \min(\{ x \}, 1 - \{ x \}) \). For ease of notation, we drop the subscript and write simply \( \| x \| \). The difference operator \( \Delta \) is defined on the set of all arithmetic functions, mapping \( f(\cdot) \) to \( f(\cdot + 1) - f(\cdot) \). The \( k \)-th difference operator \( \Delta^k \) is defined by the composition of \( \Delta \) with \( k \) times. For two arithmetic functions \( f(n) \) and \( g(n) \), \( f \ll g \) means that there is an absolute constant \( c \) such that \( |f| \leq c|g| \), and \( f = g + O(h) \) means \( f - g \ll h \).

1.2. **Anqie entropy of a class of exponential functions.** Our first main result is stated as follows.

**Theorem 1.5.** Let \( w \) be a positive integer and \( \mathbb{N} = S_1 \cup S_2 \cup \cdots \cup S_w \) be a partition of \( \mathbb{N} \), where each \( 1_{S_v}(n) \) has zero anqie entropy. Suppose that \( p_1(y), \ldots, p_w(y) \) are polynomials in \( \mathbb{R}[y] \). Let

\[
 g(n) = \sum_{v=1}^{w} 1_{S_v}(n)p_v(n). \tag{1.2}
\]

Suppose that \( f(n) \) is an arithmetic function satisfying

\[
 \lim_{n \to \infty} \| \Delta^k f(n) - g(n) \| = 0 \tag{1.3}
\]

for some \( k \in \mathbb{N} \), then \( \mathbb{A}(e(f(n))) = 0 \).

There are many examples of arithmetic functions satisfying equation (1.3), such as polynomials, sub-polynomials (e.g., \( n^\gamma \) with \( 0 < \gamma < 1 \)), and \( n^\alpha (\log n)^\beta \) (\( \alpha, \beta \in \mathbb{R} \)). Taking \( g = 0 \) in Theorem 1.5, we have that the exponentials of these functions are of zero anqie entropy. To give an example with \( g \neq 0 \), we show that the bracket polynomial phase \( e(\sqrt{3}n\{\sqrt{3}n\}) \) satisfies equation (1.3) and then has zero anqie entropy (see Example 3.3). We also consider whether the exponential function of any concatenation of polynomials has zero anqie entropy.

**Definition 1.6.** Let \( \{N_i\}_{i=0}^\infty \) be an increasing sequence of natural numbers with \( N_0 = 0 \) and \( \lim_{i \to \infty} (N_{i+1} - N_i) = \infty \). Let \( \{f_i(n)\}_{i=0}^\infty \) be a sequence of arithmetic functions. Define \( f(n) \) to be the arithmetic function given by \( f(n) = f_i(n) \) when \( N_i \leq n < N_{i+1} \) for \( i = 0, 1, \ldots \). The function \( f(n) \) is called the **concatenation of** \( \{f_i(n)\}_{i=0}^\infty \) **with respect to the sequence** \( \{N_i\}_{i=0}^\infty \).
For the exponential functions of concatennations, we obtain the following result.

**Theorem 1.7.** Let $g(n)$ be defined as in Theorem 1.5. Suppose that $\{f_i(n)\}_{i=0}^{\infty}$ is a sequence of real-valued arithmetic functions such that

$$
\lim_{n \to \infty} \sup_{i \in \mathbb{N}} \| \triangle^k f_i(n) - g(n) \| = 0
$$

(1.4)
holds for some $k \in \mathbb{N}$. Then for any concatenation $f(n)$ of $\{f_i(n)\}_{i=0}^{\infty}$, $\mathbb{E}(e(f(n))) = 0$.

We remark that polynomial, sub-polynomial and bracket polynomial pases are known as nilsequences (see [GT12b]), while many concatennations of polynomial phases are not nilsequences. Our main idea to prove Theorem 1.5 is to construct a sequence of arithmetic functions $\{f_N(n)\}_{N=0}^{\infty}$ with finite ranges uniformly converging to $f(n)$. Then it suffices to show that $\mathbb{E}(f_N) = 0$ for $N$ large enough by Theorem 1.3. In this process, we develop some methods to compute $\mathbb{E}(f_N)$. For example, to apply formula (1.1), we use the cardinality of pieces of $\mathbb{R}^k$ cut by hyperplanes to bound the cardinality of blocks with given lengths occurring in the sequence $\{f_N(n)\}_{n=0}^{\infty}$. We refer readers to Section 3 for more details.

1.3. **Connection with Sarnak’s conjecture.** By Theorem 1.4, we know that SMDC is equivalent to that the Möbius function is disjoint from any arithmetic function with zero anqie entropy. Note that the product of any two functions with zero anqie entropy also has zero anqie entropy by Proposition 1.2, so by Theorem 1.5, we have the following result.

**Proposition 1.8.** With the same assumptions as in Theorem 1.5, suppose that $f(n)$ satisfies condition (1.3). Then SMDC implies that, for any arithmetic function $\eta(n)$ with $\mathbb{E}(\eta) = 0$,

$$
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(n)\eta(n)e(f(n)) = 0.
$$

Interestingly, SMDC implies that $\mu(n)\eta(n)$ does not correlate with $e(f(n))$ in short intervals on average when the $k$-th difference of $f(n)$ satisfies certain conditions. Precisely,

**Theorem 1.9.** Let $g(n)$ be defined as in Theorem 1.5. Suppose that $\mathcal{D}$ is a family of real-valued arithmetic functions such that

$$
\lim_{n \to \infty} \sup_{f \in \mathcal{D}} \| \triangle^k f(n) - g(n) \| = 0
$$

(1.5)
holds for some $k \geq 1$. Let $X, h \geq 1$. Then SMDC implies that, for any arithmetic function $\eta(n)$ with zero anqie entropy,

$$
\lim_{h \to \infty} \lim_{X \to \infty} \frac{1}{Xh} \int_{X}^{2X} \sup_{f \in \mathcal{D}} \left| \sum_{x \leq n < x+h} \mu(n)\eta(n)e(f(n)) \right| dx = 0.
$$

(1.6)

Applying the above theorem, we obtain our second main result, which gives an equivalent condition of SMDC.
Theorem 1.10. Let $X, h \geq 1$. Then SMDC holds if and only if for any bounded arithmetic function $f$ with $\mathcal{A}(f) = 0$,

$$
\lim_{X \to \infty} \limsup_{h \to \infty} \frac{1}{Xh} \int_X^{2X} \left| \sum_{x \leq n < x+h} \mu(n)f(n) \right| dx = 0. \tag{1.7}
$$

Equivalently, for any compact Hausdorff space $X$ and any continuous map $T$ on $X$ with zero topological entropy, we have

$$
\lim_{X \to \infty} \limsup_{h \to \infty} \frac{1}{Xh} \int_X^{2X} \left| \sum_{x \leq n < x+h} \mu(n)F(T^n x_0) \right| dx = 0. \tag{1.8}
$$

for all $x_0 \in X$ and $F \in C(X)$.

As we know, Sarnak’s conjecture states the correlation between Möbius and arithmetic functions realized in any deterministic flow in long intervals. It is interesting to see from the above theorem that Sarnak’s conjecture essentially states this correlation in almost all short intervals.

Recently He and Wang [HW19] showed equation (1.8) holds for any nilmanifold $X = G/\Gamma$, any 1-Lipschitz continuous function $F$, and any $T$ with $T_x = g \cdot x_0$, where $g \in G$ and $x_0 \in G/\Gamma$.

If we take $\eta(n) = 1_{n \equiv a (\bmod q)}(n)$ and $D$ the set of all polynomials of degrees less than a given positive integer in Theorem 1.9, then we have the following corollary.

Corollary 1.11. Let $k \geq 1$ be a given integer. Denote by $D_k$ the set of all polynomials in $\mathbb{R}[y]$ of degrees less than $k$. Let $q$ and $a$ be given integers with $q \geq 1$ and $0 \leq a \leq q - 1$. Let $X, h \geq 1$. Then SMDC implies

$$
\lim_{X \to \infty} \limsup_{h \to \infty} \frac{1}{Xh} \int_X^{2X} \sup_{p(y) \in D_k} \left| \sum_{x \leq n < x+h \atop n \equiv a (\bmod q)} \mu(n)e(p(n)) \right| dx = 0. \tag{1.9}
$$

As observed in [Tao17] equation (1.9) is implied by the local higher order Fourier uniformity conjecture, which is deduced from the Chowla conjecture ([Cho65], see also [Ng08]). It is known that Chowla’s conjecture implies SMDC ([Sar09]). Corollary 1.11 shows that equation (1.9) can be deduced from SMDC.

In the following, we list some results relevant to equation (1.9). For $k = 1$, equation (1.9) has been obtained from the work of Matomäki-Radziwill ([MR16]). For $k \geq 2$, it is open whether (1.9) holds, while recently Matomäki-Radziwill-Tao-Teräväinen-Ziegler in [MRTTZ20] established equation (1.9) when $h = X^\theta$ for any fixed $\theta > 0$. Without taking the average on $X$ in equation (1.9), let $h = X^\theta$, the case that $k = 1$ and $\theta > 0.55$ was obtained by Matomäki-Teräväinen in [MT19]; the case that $k = 2$ and $\theta > 5/8$ was previously established by Zhan in [Zhan91] and extended to $\theta > 3/5$ in [MT19]; the case that $k \geq 2$ and $\theta > 2/3$ was obtained by Matomäki-Shao in [MS19] (also see Huang [Huang16] for related results on $\Lambda(n)$ instead of $\mu(n)$).
1.4. Disjointness of Möbius from $e(f(n))$ with the $k$-th difference of $f(n)$ tending to zero. The disjointness of Möbius from exponential functions is important and has been extensively studied in number theory. For example, the disjointness of $\mu$ from $e(n\alpha)$, for any $\alpha \in \mathbb{R}$, is closely related to the estimate of exponential sums in prime variables, from which one can deduce Vinogradov’s three primes theorem. In this paper, we shall study Möbius disjointness of $e(f(n))$ with $\lim_{n \to \infty} \| \Delta^k f(n) \| = 0$ for some natural number $k$, which is a necessary condition of SMDC by Proposition 1.8. It is obvious that $\mu(n)$ is disjoint from $e(f(n))$ for the case $k = 0$ by the Prime Number Theorem. For the case that $f(n)$ is a polynomial of degree $d$ ($k = d+1$), the Möbius disjointness of $e(f(n))$ has been established by Davenport [Dav37] when $d = 1$ and by Hua [Hua65] when $d \geq 2$. For general $f(n)$, when $k = 1$, we have

**Theorem 1.12.** Suppose that $f(n)$ is a real-valued arithmetic function and $c \in \mathbb{R}$ a constant satisfying

$$\lim_{n \to \infty} \| \Delta f(n) - c \| = 0.$$  

Then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f(n)) = 0.$$  

For $k = 2$, we have the following result.

**Theorem 1.13.** Suppose that $f(n)$ is a real-valued arithmetic function satisfying that the set \{ $e(\Delta^2 f(n)) : n = 0, 1, \ldots$ \} has finitely many limit points, and

$$\lim_{n \to \infty} \| \Delta^2 f(n) - c \| = 0$$

for some constant $c \in \mathbb{R}$. Then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f(n)) = 0.$$  

We remark that Theorem 1.12 is a special case of Theorem 1.13. They are both applications of our third main result stated below on a discorrelation estimate between $\mu$ and polynomial phases on almost all arithmetic progressions $a(mod s)$ and almost all short intervals simultaneously.

**Theorem 1.14.** Let $s \geq 1$ and $h \geq 3$ be integers. Let $d \geq 0$ be an integer. Suppose $p(x) \in \mathbb{R}[x]$ of degree $d$. Then for $N$ large enough with $hs \leq (\log N)^{1/32}$,

$$\frac{1}{Ns} \sum_{a=1}^{s} \sum_{n=1}^{N} \left\lfloor \frac{1}{h} \sum_{m=n+1 \atop m \equiv a(mod s)}^{n+h} \mu(m)e(p(m)) \right\rfloor \ll (\log \log(s+2)) \frac{(\log \log h)^3}{(\log h)^{1/2}}, \quad (1.10)$$

where the implied constant at most depends on $d$.

We list some results related to formula (1.10). For $s = 1$, as a consequence of the recent breakthrough of Matomäki-Radziwiłł [MR16] and Matomäki-Radziwiłł-Tao [MRT15], formula (1.10) holds for linear functions $p(x)$. As mentioned previously below Theorem 1.10, He-Wang showed
formula (1.10) with $e(p(m))$ replaced by any nilsequence, see also [EALdlR17], [FFKPL19] for earlier partial results in this direction.

For general $s$, Kanigowski, Lemańczyk and Radziwiłł [KLR19] showed formula (1.10) when $p(x)$ is a constant (i.e., $d = 0$). Here we show formula (1.10) holds for all polynomials $p(x)$. In fact, we study a general version of Theorem 1.10 replacing $\mu$ by multiplicative functions. This will be given in Theorem 5.1.

Our main issue to prove Theorem 1.14 is to split the estimate into major and minor arc cases. In the major arc case, we use the second author’s result [Wei18b, Theorem 1.7] on self-correlations of multiplicative functions in short arithmetic progressions on average. In the minor arc case, we use a variant of the arguments of [HW19] applied to polynomials $p(x) \in \mathbb{R}[x]$. We remark that if we use the result of [HW19] directly to prove formula (1.10) by taking the multiplicative function $\beta(n) = \mu(n)\chi(n)$ with $\chi$ a Dirichlet character, then we have that the coefficient before $\frac{(\log \log h)^3}{(\log h)^{3/2}}$ is the quantity $s$. Due to our method, this result is not enough to prove Theorem 1.13.

We also refer to [MV18] for a different method to prove this case.

Now, let us return to the further study of the Möbius disjointness of $e(f(n))$. For $k \geq 2$, we obtain that $\mu$ is disjoint from $e(f(n))$ when the $k$-th difference of $f(n)$ decays to zero at a certain rate.

**Theorem 1.16.** Let $\tau \in (5/8, 1)$ and $k \geq 2$. Let $f(n)$ be a real-valued arithmetic function satisfying when $n$ large enough,

$$\|\Delta^k f(n)\| \leq \frac{C}{\exp((\log n)^\tau)}$$

(1.12)
for some positive constant $C$. Then
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f(n)) = 0.
\]

The first ingredient of our proof of the above result is that $e(f(n))$ with the $k$-th difference of $f(n)$ decaying to zero can be approximated uniformly by certain concatenations of polynomial phases. The second one is Matomäki-Radziwiłł-Tao-Teräväinen-Ziegler’s estimate [MRTTZ20] on averages of the correlation of multiplicative functions with polynomial phases in short intervals.

In the following, we provide a sufficient condition that $\mu$ is disjoint from $e(f(n))$ with the $k$-th difference of $f(n)$ tending to zero.

**Proposition 1.17.** Let $k$ be a given positive integer. Assume that equation (1.9) holds, then for any $f(n)$ satisfying $\lim_{n \to \infty} \|\triangle^k f(n)\| = 0$, we have
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f(n)) = 0.
\]

**Remark 1.18.** All the results listed in this section also hold if $\mu$ is replaced by many other multiplicative functions, such as the Liouville function and $\mu(n)\chi(n)$, where $\chi$ is a given Dirichlet character.

This paper is organized as follows. In Section 2, we prove some preliminary results on computing the anqie entropy of arithmetic functions with finite ranges. In Section 3, we prove that a certain class of exponential functions have zero anqie entropy. Theorems 1.5 and 1.7 are proved in this section. In Section 4, we give some necessary conditions of Sarnak’s Möbius disjointness conjecture. Proofs of proposition 1.8, Theorems 1.9 and 1.10 are present. In Section 5, we prove Theorem 1.14. In Section 6, we study the Möbius disjointness of exponential functions of arithmetic functions with the $k$-th differences tending to zero. Theorems 1.12, 1.13, 1.16, and Propositions 1.15, 1.17 are shown.

**Acknowledgments.** This research was begun when we visit Harvard University during the 2018-2019 academic year. We sincerely thank Professor Arthur Jaffe and Professor Jinxin Xue for their support. We would like to thank Professor Liming Ge for valuable discussions. This research was supported by Grant TRT 0159 from the Templeton Religion Trust. F. W. is supported by the fellowship of China Postdoctoral Science Foundation 2020M670273.

2. Preliminaries

2.1. Preliminary results for Sections 3 and 4. In this subsection, we show some preliminary results on computing the anqie entropy of arithmetic functions with finite ranges. Let us first recall some basic concepts in symbolic dynamical systems. For a finite set $A$, a block over $A$ is a finite sequence of symbols from $A$. A $J$-block is a block of length $J$ ($J \geq 1$). For any given (finite or infinite) sequence $x = (x_0, x_1, \ldots)$ of symbols from $A$, we say that a block $w$ occurs in $x$ or $x$ contains $w$ if there are natural numbers $i$, $j$ with $i \leq j$, such that $(x_i, \ldots, x_j) = w$. 
A concatenation of two blocks \( w_1 = (a_1, \ldots, a_k) \) and \( w_2 = (b_1, \ldots, b_l) \) over \( \mathbb{A} \) is the block \( w_1w_2 = (a_1, \ldots, a_k, b_1, \ldots, b_l) \).

Now suppose that \( f \in l^\infty(\mathbb{N}) \) has finite range. We can view \( f \) as a sequence \( \{f(n)\}_{n=0}^\infty \) in \( f(\mathbb{N})^\mathbb{N} \). Let \( B_J(f) \) denote the set of all \( J \)-blocks that occur in \( f \), i.e.,

\[
B_J(f) = \{(f(n), f(n+1), \ldots, f(n+J-1)) : n \geq 0 \}.
\]

Based on these notation, a \( J \)-block of the form

\[
(f(lJ), f(lJ+1), \ldots, f(lJ+J-1)),
\]

for some \( l \in \mathbb{N} \) is called a regular \( J \)-block in \( f \). Denote the set of all regular \( J \)-blocks in \( f \) by \( B^r_J(f) \). A \( J \)-block, which occurs infinitely many times in the sequence \( \{f(n)\}_{n=0}^\infty \), is called an effective \( J \)-block in \( f \). Denote the set of all such blocks in \( f \) by \( B^e_J(f) \). A \( J \)-block \( (a_0, a_1, \ldots, a_{J-1}) \) is called regularly effective in \( f \) if there are infinitely many natural numbers \( n \) such that

\[
(a_0, a_1, \ldots, a_{J-1}) = (f(nJ), \ldots, f(nJ+J-1)).
\]

The set of all regularly effective blocks in \( f \) is denoted by \( B^e_J(f) \).

Throughout the paper, for any finite set \( S \), \( |S| \) is the cardinality of \( S \). For a function \( f \) taking finitely many values, as we stated in Section 1,

\[
\mathcal{A}(f) = \lim_{J \to \infty} \frac{\log |B_J(f)|}{J}. \tag{2.1}
\]

In the following, we show that \( \mathcal{A}(f) \) also can be computed through the cardinality of \( B^e_J(f) \), \( B^r_J(f) \) or \( B^{e,r}_J(f) \).

**Theorem 2.1.** Let \( f(n) \) be an arithmetic function with finite range. Then

\[
\mathcal{A}(f) = \lim_{J \to \infty} \frac{\log |B^r_J(f)|}{J} = \lim_{J \to \infty} \frac{\log |B^{e,r}_J(f)|}{J} = \lim_{J \to \infty} \frac{\log |B^e_J(f)|}{J}. \tag{2.2}
\]

**Proof.** We first show that the first equality in equation (2.2) holds. On one hand, \( B^r_J(f) \subseteq B_J(f) \), so by formula (2.1),

\[
\mathcal{A}(f) \geq \limsup_{J \to \infty} \frac{\log |B^r_J(f)|}{J}.
\]

On the other hand, given \( J \geq 1 \), for any \( l \geq 1 \) and any \((lJ)\)-block \( w \) occurring in \( f \), there is a concatenation of certain \( l+1 \) successive regular \( J \)-blocks in \( f \) containing \( w \). Thus \( |B_{lJ}(f)| \leq J|B^r_J(f)|^{l+1} \). This implies that

\[
\mathcal{A}(f) = \lim_{l \to \infty} \frac{\log |B_{lJ}(f)|}{lJ} \leq \frac{\log |B^r_J(f)|}{J}.
\]

We then have

\[
\mathcal{A}(f) \leq \liminf_{J \to \infty} \frac{\log |B^r_J(f)|}{J}.
\]

So \( \lim_{J \to \infty} \frac{\log |B^r_J(f)|}{J} \) exists and equals \( \mathcal{A}(f) \).
Next we show that the second equality in equation (2.2) holds, i.e.,
\[
\lim_{J \to \infty} \frac{\log |B_r^J(f)|}{J} = \lim_{J \to \infty} \frac{\log |B_{e,r}^J(f)|}{J}.
\] (2.3)

Since \(B_{e,r}^J(f) \subseteq B_r^J(f)\),
\[
\lim_{J \to \infty} \frac{\log |B_r^J(f)|}{J} \geq \limsup_{J \to \infty} \frac{\log |B_{e,r}^J(f)|}{J}.
\] (2.4)

So we only need to show that
\[
\lim_{J \to \infty} \frac{\log |B_r^J(f)|}{J} \leq \liminf_{J \to \infty} \frac{\log |B_{e,r}^J(f)|}{J}.
\] (2.5)

In fact, for any given \(J \geq 1\), since the set \(B_r^J(f) \setminus B_{e,r}^J(f)\) is finite, there is an integer \(l_J \geq 1\) such that all regular \(J\)-blocks in the set
\[
\{(f(nJ), \ldots, f(nJ + J - 1)) : n \geq l_J\}
\]
are regularly effective \(J\)-blocks in \(f\). Then for each \(l > l_J\), there is at most one regular \((lJ)\)-block in \(f\) which is not a concatenation of regularly effective \(J\)-blocks in \(f\). This implies \(|B_r^{lJ}(f)| \leq |B_{e,r}^{lJ}(f)|^l + 1\). Therefore
\[
\mathcal{E}(f) = \lim_{l \to \infty} \frac{\log |B_{e,r}^{lJ}(f)|}{lJ} \leq \lim_{l \to \infty} \frac{\log(|B_{e,r}^{lJ}(f)|^l + 1)}{lJ} = \frac{\log |B_{e,r}^{lJ}(f)|}{J}
\]
holds. Letting \(J \to \infty\), we obtain formula (2.5).

At last, note that
\[
B_{e,r}^{lJ}(f) \subseteq B_{e}^{lJ}(f) \subseteq B_f(f),
\]
then
\[
\lim_{J \to \infty} \frac{\log |B_{e}^{lJ}(f)|}{J} \leq \liminf_{J \to \infty} \frac{\log |B_{e,r}^{lJ}(f)|}{J}, \quad \limsup_{J \to \infty} \frac{\log |B_r^{lJ}(f)|}{J} \leq \lim_{J \to \infty} \frac{\log |B_f(f)|}{J}.
\]

From formula (2.1) and the second equality in (2.2), we have
\[
\mathcal{E}(f) = \lim_{J \to \infty} \frac{\log |B_f(f)|}{J} = \lim_{J \to \infty} \frac{\log |B_{e,r}^{lJ}(f)|}{J}.
\]
Then \(\lim_{J \to \infty} \frac{\log |B_{e,r}^{lJ}(f)|}{J}\) exists and equals \(\mathcal{E}(f)\). So the third equality in equation (2.2) holds. □

To estimate the cardinality of the set of \(J\)-blocks occurring in certain sequences, we introduce the following notion.

**Definition 2.2.** Let \(k, m \geq 1\) be integers and \(c_1, \ldots, c_m \in \mathbb{R}\) be constants. Suppose that for \(j = 1, \ldots, m\), \(F_j\) is a non-zero linear function of \(x_1, \ldots, x_k\) and \(H_j\) is the hyperplane in \(\mathbb{R}^k\) given by \(F_j(x_1, \ldots, x_k) = c_j\). Denote by
\[
H_j^+ = \{(x_1, \ldots, x_k) \in \mathbb{R}^k : F_j(x_1, \ldots, x_k) > c_j\},
\]
\[
H_j^- = \{(x_1, \ldots, x_k) \in \mathbb{R}^k : F_j(x_1, \ldots, x_k) < c_j\}.
\]
A non-empty subset $P$ of $\mathbb{R}^k$ of the following form

$$P = P_1 \cap P_2 \cap \cdots \cap P_m,$$

where each $P_j \in \{H_j^+, H_j^-, H_j\}$, is called a piece of $\mathbb{R}^k$ cut by $H_1, ..., H_m$.

In the following lemma we give an upper bound for the cardinality of pieces of $\mathbb{R}^k$ cut by hyperplanes.

**Lemma 2.3.** Let $k, m \geq 1$ be integers. Suppose that $H_1, ..., H_m$ are hyperplanes in $\mathbb{R}^k$. Let $C(H_1, ..., H_m, k)$ denote the cardinality of pieces of $\mathbb{R}^k$ cut by $H_1, ..., H_m$ and $W(m, k)$ denote the maximal value of $C(H_1, ..., H_m, k)$ when $H_1, ..., H_m$ go through all the possible hyperplanes. Then

$$W(m, k) \leq \sum_{j=0}^{k} 2^j \binom{m}{j}.$$

In particular, $W(m, k) \leq (k + 1)2^km^k$.

**Proof.** Notice that $W(1, k) = 3$ for any $k \geq 1$. If we have

$$W(m, k) \leq W(m-1, k) + 2W(m-1, k-1),$$

then one can easily draw the conclusion by induction on $m$.

Now we show formula (2.6) holds for $m \geq 2$. Let $\psi$ be the map from the set of all pieces of $\mathbb{R}^k$ cut by $H_1, ..., H_m$, denoted by $\mathcal{P}$, onto the set of all pieces of $\mathbb{R}^k$ cut by $H_2, ..., H_m$, denoted by $\tilde{\mathcal{P}}$, given by

$$\psi : P_1 \cap P_2 \cap \cdots \cap P_m \mapsto P_2 \cap \cdots \cap P_m.$$  

Then for any piece $D \in \tilde{\mathcal{P}}$, $|\psi^{-1}(D)| \leq 3$. Now we show that if $|\psi^{-1}(D)| \geq 2$, then $H_1 \cap D$ is nonempty. In fact, the only case we need to consider is when both $H_1^+ \cap D$ and $H_1^- \cap D$ are nonempty. In this case, suppose $p_1 \in H_1^+ \cap D$ and $p_2 \in H_1^- \cap D$, then there is a point $p_0 \in H_1 \cap D$ by the convexity of $D$. Hence, $|\mathcal{P}| - |\tilde{\mathcal{P}}| = C(H_1, ..., H_m, k) - C(H_2, ..., H_m, k)$ does not exceed the cardinality of pieces of the form $H_1 \cap P_2 \cap \cdots \cap P_m$ in $\mathcal{P}$ times 2. In the following, we estimate this cardinality.

For each $j = 2, ..., m$, $H_1 \cap H_j$ is a hyperplane in $H_1$, or empty, or equal to $H_1$. Denote $G_1, ..., G_n$ to be the ones which are hyperplanes in $H_1$. Then $n \leq m - 1$. We claim that if $H_1 \cap P_2 \cap \cdots \cap P_m = (H_1 \cap P_2) \cap \cdots \cap (H_1 \cap P_m) \neq \emptyset$, then $(H_1 \cap P_2) \cap \cdots \cap (H_1 \cap P_m)$ is a piece of $H_1$ cut by $G_1, ..., G_n$. In fact, for $j = 2, ..., m$, there are at most three cases. When $H_1 \cap H_j = \emptyset$, then $H_1 \cap H_j^+ = H_1$ or $H_1 \cap H_j^- = H_1$ and then $P_j = H_j^+$ or $H_j^-$, respectively. When $H_1 \cap H_j = H_1$, then $P_j = H_j$ and $H_1 \cap P_j = H_1$. When $H_1 \cap H_j = G_{i_j}$ is a hyperplane in $H_1$, then $H_1 \cap H_j^+ = G_{i_j}^+$ and $H_1 \cap H_j^- = G_{i_j}^-$. Hence $(H_1 \cap P_2) \cap \cdots \cap (H_1 \cap P_m)$ is of the form $T_1 \cap \cdots \cap T_n$, where each $T_i \in \{G_i, G_i^+, G_i^-\}$.

So the cardinality of pieces of the form $H_1 \cap P_2 \cap \cdots \cap P_m$ in $\mathcal{P}$ does not exceed $C(G_1, ..., G_n, k-1)$ which is at most $W(n, k-1) \leq W(m-1, k-1)$. Therefore the inequality (2.6) holds and the proof is completed. $\square$
2.2. Preliminary results for Section 5. In this subsection, we shall list some lemmas and prove some results which will be used in the proof of Theorem 5.1. In the following, we use \([N]\) to denote the set \(\{1, \ldots, N\}\) and \(\alpha \mathbb{Z}\) to denote an equivalence class in \(\mathbb{R}/\mathbb{Z}\) for any \(\alpha \in \mathbb{R}\), i.e., \(\alpha (\text{mod } \mathbb{Z})\). We first recall some definitions.

**Definition 2.4.** ([GT12a, Definition 2.7]) Let \(d \geq 0\) and \(f : \mathbb{Z} \to \mathbb{R}/\mathbb{Z}\) be a polynomial of degree \(d\). Write \(f(n) = \sum_{i=0}^{d} \alpha_i n^i\) with each \(\alpha_i \in \mathbb{R}/\mathbb{Z}\). For any integer \(N \geq 1\), the smoothness norm of \(f\) is defined by
\[
\|f\|_{C^\infty[N]} := \max_{1 \leq i \leq d} N^i \|\alpha_i\|_{\mathbb{R}/\mathbb{Z}}.
\]

**Definition 2.5.** (See e.g., [GT12a, Definition 1.2]) Let \(G/\Gamma\) be a nilmanifold endowed with the unique normalized Haar measure. Let \(\delta > 0\) and \(A\) a finite arithmetic progression in \(\mathbb{Z}\). A finite sequence \(\{x_n\}_{n \in A}\) in \(G/\Gamma\) is said to be \(\delta\)-equidistributed in \(G/\Gamma\) if one has
\[
\left| \mathbb{E}_{n \in A} F(x_n) - \int_{G/\Gamma} F \right| \leq \delta \|F\|_{\text{Lip}}
\]
for all Lipschitz function \(F : G/\Gamma \to \mathbb{C}\), where
\[
\mathbb{E}_{n \in A} F(x_n) = \frac{1}{|A|} \sum_{n \in A} F(x_n)
\]
and
\[
\|F\|_{\text{Lip}} := \sup_{x \in G/\Gamma} |F(x)| + \sup_{x, y \in G/\Gamma, x \neq y} \frac{|F(x) - F(y)|}{d_{G/\Gamma}(x, y)}.
\]

And a finite sequence \(\{x_n\}_{n \in A}\) in \(G/\Gamma\) is said to be totally \(\delta\)-equidistributed in \(G/\Gamma\) if the sequence \(\{x_n\}_{n \in A'}\) is \(\delta\)-equidistributed in \(G/\Gamma\) for all arithmetic progressions \(A' \subseteq A\) of length at least \(\delta N\).

The next result gives a simple connection between the coefficients of a polynomial and its smoothness norm.

**Lemma 2.6.** [GT12b, Lemma 3.2] Suppose that \(f : \mathbb{Z} \to \mathbb{R}/\mathbb{Z}\) is a polynomial of the form \(\sum_{i=0}^{d} \beta_in^i\). Then there is a positive integer \(D = O_d(1)\) such that \(\|D\beta_i\| \ll_d N^{-i}\|f\|_{C^\infty[N]}\) for all \(i = 1, \ldots, d\).

The following is a “strong recurrence” result for polynomials \(f : \mathbb{Z} \to \mathbb{R}\).

**Lemma 2.7.** [GT12a, Lemma 4.5] Suppose that \(f : \mathbb{Z} \to \mathbb{R}\) is a polynomial of degree \(d \geq 0\). Suppose that \(N \geq 1\), \(\delta \in (0, 1/2)\) and \(\epsilon \in (0, \delta/2]\). If \(f(n)(\text{mod } \mathbb{Z})\) belongs to an interval \(I \subseteq \mathbb{R}/\mathbb{Z}\) of length \(\epsilon\) for at least \(\delta N\) integers of \(n \in [N]\). Then there is a \(D \in \mathbb{Z}\) with \(0 < |D| \ll \delta^{-O_d(1)}\) such that \(\|Df(\text{mod } \mathbb{Z})\|_{C^\infty[N]} \ll_d \epsilon^{O_d(1)}\).

The next one is a special case of the quantitative equidistribution result on polynomials \(f : \mathbb{Z} \to \mathbb{R}\) ([GT12a, Theorem 2.9]).
Lemma 2.8. Let $d \geq 0$, $N \geq 1$ be integers. Suppose that $f : \mathbb{Z} \to \mathbb{R}$ is a polynomial of degree $d$. Let $R \geq 10$ and $N \gg R^{O_d(1)}$. If $\{f(n)\}_{n \in [N]}$ is not totally $R^{-1}$-equidistributed in $\mathbb{R}/\mathbb{Z}$, then there is an integer $D$ with $0 < |D| \leq R^{O_d(1)}$ such that $\|Df(\mod Z)\|_{C^\infty[N]} \leq R^{O_d(1)}$.

Now, we prove a variant version of the above result for the equidistribution in arithmetic progressions that is adapted to our situation.

Proposition 2.9. Suppose that $f(n) = \sum_{i=0}^d \alpha_i n^i$ with each $\alpha_i \in \mathbb{R}/\mathbb{Z}$ and $g(n,h) = f(n+h)$. Let $R \geq 10$ and $N,H \in \mathbb{N}$ with $N,H \gg R^{O_d(1)}$. Let integers $s \geq 1$ and $1 \leq a \leq s$. Write $A_a = \{h : 1 \leq h \leq Hs, h \equiv a(\mod s)\}$. Then at least one of the following holds:

(i). $\{g(n,h)\}_{h \in A_a}$ is totally $R^{-1}$-equidistributed in $\mathbb{R}/\mathbb{Z}$ for all but at most $R^{-1}Ns$ values of $(n,a) \in [N] \times [s]$.

(ii). there is an integer $Q$ with $0 < |Q| \leq \tilde{R}^{O_d(1)}$ such that $\|Q\alpha_d^s\|_{\mathbb{R}/\mathbb{Z}} \leq \tilde{R}^{O_d(1)}H^{-1}N^{-(d-1)}$.

Proof. Note that $g(n,h) = \alpha_d(n+h)^d + \alpha_{d-1}(n+h)^{d-1} + \cdots + \alpha_0$. If (i) in the above claim does not hold, then there is an $a_0$ with $1 \leq a_0 \leq s$ such that for more than $R^{-1}N$ values of $n \in [N], \{g(n,h)\}_{h \in A_{a_0}}$ is not totally $R^{-1}$-equidistributed. Let $T(n,l) = g(n,(l-1)s + a_0) = \alpha_d(n + ls + a_0 - s)^d + \alpha_{d-1}(n + ls + a_0 - s)^{d-1} + \cdots + \alpha_0 =: \sum_{i=0}^d \beta_i l^i$. Then there are more than $R^{-1}N$ values of $n \in [N]$ such that $\{T(n,l)\}_{l \in [H]}$ is not totally $R^{-1}$-equidistributed. For each such $n$, by Lemma 2.8, there is an integer $D$ with $0 < |D| \leq \tilde{R}^{O_d(1)}$ such that

$$\|DT(n,\cdot)(\mod Z)\|_{C^\infty[H]} \ll \tilde{R}^{O_d(1)}.$$ (8.8)

By the pigeonhole principal, there is a common $D$ with $0 < |D| \leq \tilde{R}^{O_d(1)}$ such that formula (8.8) holds for at least $\tilde{R}^{-O_d(1)}N$ values of $n \in [N]$. It is not hard to check that $\beta_1 = \sum_{j=1}^d j\alpha_j(n + a_0 - s)^{j-1}s$. This implies that by Lemma 2.6,

$$\|D\beta_1\|_{\mathbb{R}/\mathbb{Z}} = \|\sum_{j=1}^d D(j\alpha_j(n + a_0 - s)^{j-1}s\|_{\mathbb{R}/\mathbb{Z}} \ll \tilde{R}^{O_d(1)}H^{-1}.$$ 

Denote $p(n) = \sum_{j=1}^d D(j\alpha_j(n + a_0 - s)^{j-1}s$. By Lemma 2.7, there is an integer $\eta$ with $0 < |\eta| \leq \tilde{R}^{O_d(1)}$ such that

$$\|\eta p(n)(\mod Z)\|_{C^\infty[N]} \ll \tilde{R}^{O_d(1)}H^{-1}.$$ 

Then by the definition of smoothness norm, $\|d!\eta Dd\alpha_d s\|_{\mathbb{R}/\mathbb{Z}} \ll \tilde{R}^{O_d(1)}H^{-1}N^{-(d-1)}$. Choose $Q = \eta Dd$. We obtain (ii) in the claim of this proposition. $\square$

Next we deduce a quantitative factorization theorem. It is fundamental in the proof of Theorem 1.14. We use Proposition 2.9 and a variant of the arguments of [HW19, Theorem 3.6] to obtain it.

Theorem 2.10. Let $f : \mathbb{Z} \to \mathbb{R}$ be a polynomial of degree $d \geq 0$ and $g(n,h) = f(n+h)$. Let $R \geq 10$. Let $s \geq 1$ be an integer. Then for $B \geq 1, N, H \in \mathbb{N}$ such that $N > (Hs)^{O(1)}$ and
$H \gg R^{O(1)}$, there is an integer $W \in [R, R^{O(B^d)}]$, a group $G' = \mathbb{R}$ or $\{0\}$, a set $\mathcal{N} \subseteq [N]$ with $|\mathcal{N}| \geq (1 - W^{-B/2})N$ and a polynomial decomposition

$$g(n, h) = \mathcal{E}(n, h) + g'(n, h) + \gamma(n, h)$$

with $\mathcal{E}, g', \gamma : \mathbb{Z}^2 \to \mathbb{R}$ satisfying

(i). $|\mathcal{E}(n, h + 1) - \mathcal{E}(n, h)| \ll_d \frac{1}{WH}$ for $h \in [Hs]$.

(ii). $g'(n, h)$ takes values in $G'$, and for any $n \in \mathcal{N}$, there are at least $(1 - W^{-B/2})s$ integers of $a \in [s]$ such that $\{g'(n, h)\Gamma_h \}_{h \in \mathcal{A}_a}$ is totally $W^{-B}$-equidistributed in $G'/\Gamma'$, where $\mathcal{A}_a = \{1 \leq h \leq Hs : h \equiv a \pmod{s}\}$ and $\Gamma' = G' \cap \mathbb{Z}$.

(iii). For some integer $q$ with $1 \leq q \leq W$, $\{\gamma(n, h)\mathcal{Z} \}_{(n, h) \in \mathbb{Z}^2}$ is $qs$-periodic both in $n$ and $h$.

**Proof.** Suppose that $f(n) = \sum_{i=0}^{d} \alpha_in^i$. Let $\bar{R} = R^{B}$. If $\{g(n, h)\mathcal{Z} \}_{h \in \mathcal{A}_a}$ is totally $\bar{R}^{-1}$-equidistributed in $\mathbb{R}/\mathbb{Z}$ for all but at most $\bar{R}^{-1}Ns$ values of $(n, a) \in [N] \times [s]$, then choose $\mathcal{E}(n, h) = \gamma(n, h) = 0, g'(n, h) = g(n, h)$ and $G' = \mathbb{R}, W = R$, we obtain the claim. Otherwise, applying Proposition 2.9 to $g(n, h)$, there is an integer $Q$ with $0 < Q \leq \bar{R}^{O(1)}$ such that

$$\|Q\alpha_d s\|_{\mathbb{Z}/Z} \leq \bar{R}^{O(1)}H^{-1}N^{-(d-1)}.$$ 

Then there is a $c \in \mathbb{Z}$ with $|Q\alpha_d s - c| \leq \bar{R}^{O(1)}H^{-1}N^{-(d-1)}$. Write $\mathcal{E}_1(n, h) = (\alpha_d - c\gamma_d)(n+h)^d$ and $\gamma_1(n, h) = \bar{Q}s\gamma(n+h)^d$. Then $g(n, h) = \mathcal{E}_1(n, h) + g_1(n, h) + \gamma_1(n, h)$, where $g_1(n, h) = \sum_{i=0}^{d-1} \alpha_i(n+h)^i$. Let $R_1 = \bar{R} \cdot \bar{R}^{O(1)} = R^{O(B)}$. We then again apply Proposition 2.9 to $g_1(n, h)$ with $\bar{R}$ replaced by $R_1 = R_1^B$. This procedure continues if Case (i) in Proposition 2.9 does not hold. So in the $m$-th step, we shall apply Proposition 2.9 with $\bar{R}$ replaced by $\bar{R}_{m-1}$, where

$\bar{R}_{m-1} = (R_{m-1})^B$, and $R_{m-1} = \bar{R}_{m-2}\bar{R}_{m-2}^{O(1)}$. Then we obtain

- $\mathcal{E}_m(n, h)$ satisfies $|\mathcal{E}_m(n, h + 1) - \mathcal{E}_m(n, h)| \ll_d (\bar{R}_{m-1})^{O(1)}/(Hs)$ for $h \in [Hs]$;
- $\{\gamma_m(n, h)\mathcal{Z} \}_{(n, h) \in \mathbb{Z}^2}$ is $q_m$-periodic for some positive integer $q_m \leq (\bar{R}_{m-1})^{O(1)}$;
- $g(n, h) = \sum_{i=1}^{m} \mathcal{E}_i(n, h) + g_m(n, h) + \sum_{i=1}^{m} \gamma_i(n, h)$, where $g_m(n, h) = \sum_{i=0}^{d-m} \alpha_i(n+h)^i$, and $\{\sum_{i=1}^{m} \gamma_i(n, h)\mathcal{Z} \}_{(n, h) \in \mathbb{Z}^2}$ is $(\prod_{i=1}^{m} q_i)$-periodic with $\prod_{i=1}^{m} q_i \leq R_{m-1}$.

Since $d$ is a given non-negative integer, the process must stop at some $m \leq d$. This implies that either $\{g_m(n, h)\Gamma_h \}_{h \in \mathcal{A}_a}$ is totally $(\bar{R}_{m-1})^{-1}$-equidistributed in $G'/\Gamma'$ for all but $(\bar{R}_{m-1})^{-1}Ns$ values of $(n, a) \in [N] \times [s]$, where $G' = \mathbb{R}$, or $g_m(n, h) = 0$ and $G' = \{0\}$. In either case, choose $W = R_m = R^{O(B^d)}$, $\mathcal{E}(n, h) = \sum_{i=1}^{m} \mathcal{E}_i(n, h)$, $g'(n, h) = g_m(n, h)$ and $\gamma(n, h) = \sum_{i=1}^{m} \gamma_i(n, h)$. So we obtain that $\{g(n, h)\Gamma_h \}_{h \in \mathcal{A}_a}$ is totally $W^{-B}$-equidistributed in $G'/\Gamma'$ for all but at most $W^{-B}Ns$ values of $(n, a) \in [N] \times [s]$. Let $\mathcal{N}$ denote the set of $n \in [N]$ which satisfies that at least $(1 - W^{-B/2})s$ choices of $a \in [s]$ has the property that $\{g(n, h)\Gamma_h \}_{h \in \mathcal{A}_a}$ is totally $W^{-B}$-equidistributed in $G'/\Gamma'$. It is not hard to check that

$$|\mathcal{N}|W^{-B/2}s \leq W^{-B}Ns.$$ 

Then $|\mathcal{N}| \geq (1 - W^{-B/2})N$. We complete the proof. \qed

In the above, the implicit constants $O(1)$ at most depends on the degree $d$. 
3. ANQIE ENTROPY OF A CERTAIN CLASS OF ARITHMETIC FUNCTIONS

In this section, we shall prove Theorems 1.5 and 1.7. Recall that \( \|x\| = \inf_{m \in \mathbb{Z}} |x - m| = \min\{|x|, 1 - |x|\} \). Then \( \| \cdot \| \) defines a metric on \( \mathbb{R}/\mathbb{Z} \) and the topology induced by it on \( \mathbb{R}/\mathbb{Z} \) is equivalent to the Euclid topology on the unit circle. The following lemma will be used in the proofs of Theorems 1.5 and 1.7 in this section.

**Lemma 3.1.** Let \( f, g \) be real-valued arithmetic functions with

\[
\lim_{n \to \infty} \| \Delta^k f(n) - g(n) \| = 0 \tag{3.1}
\]

for some \( k \in \mathbb{N} \). Then, for any \( \epsilon > 0 \) and positive integer \( m \geq 1 \), there is some \( L \in \mathbb{N} \) such that, whenever \( n > L \), the following holds for any \( j \) with \( 0 \leq j \leq m - 1 \),

\[
\|f(n + j) - Y_n(n + j)\| \leq \epsilon,
\]

where \( Y_n(n + j) \) is defined to be \( f(n + j) \) when \( 0 \leq j \leq k - 1 \) and to be the value determined by the following linear equations when \( k \leq j \leq m - 1 \),

\[
\Delta^k Y_n(n + j) = g(n + j), \quad j = 0, 1, \ldots, m - k - 1. \tag{3.2}
\]

**Proof.** When \( k = 0 \), the claim is trivial. In the following, we assume \( k \geq 1 \). We use induction on \( m \) to prove the lemma. For \( m \leq k \), since \( Y_n(n + j) = f(n + j) \) for \( j = 0, 1, \ldots, m - 1 \), choose \( L = 0 \). Then we obtain the claim in the lemma. Assume inductively that the claim holds for some \( m_0 \geq k \). In the following we shall prove the claim holds for \( m_0 + 1 \) case. By condition (3.1) and Proposition A.1,

\[
\lim_{n \to \infty} \| \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} f(n + l) - g(n) \| = 0.
\]

Then for any \( \epsilon > 0 \), there is an \( L_1 > 0 \) such that whenever \( n > L_1 \),

\[
\| \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} f(n + m_0 - k + l) - g(n + m_0 - k) \| < \epsilon/2,
\]

i.e.,

\[
\|f(n + m_0) - (g(n + m_0 - k) - \sum_{l=0}^{k-1} (-1)^{k-l} \binom{k}{l} f(n + m_0 - k + l))\| < \epsilon/2. \tag{3.3}
\]

By the induction hypothesis, there is an \( L_0 \in \mathbb{N} \), whenever \( n > L_0 \),

\[
\|f(n + j) - Y_n(n + j)\| < \epsilon/2^{k+1}, \quad j = 0, 1, \ldots, m_0 - 1. \tag{3.4}
\]

Let \( L = \max\{L_0, L_1\} \). Then by equations (3.3) and (3.4), whenever \( n > L \),

\[
\|f(n + m_0) - (g(n + m_0 - k) - \sum_{l=0}^{k-1} (-1)^{k-l} \binom{k}{l} Y_n(n + m_0 - k + l))\| < \epsilon. \tag{3.5}
\]
By the definition of $Y_n$,
\[
\Delta^k Y_n(n + m_0 - k) = \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} Y_n(n + m_0 - k + l) = g(n + m_0 - k).
\]
Then by equation (3.5),
\[
\|f(n + m_0) - Y_n(n + m_0)\| < \epsilon.
\]
Combing with equation (3.4), we obtain the claim for $m = m_0 + 1$, completing the induction. \[\square\]

**Theorem 1.5 (restated).** Let $w$ be a positive integer and $\mathbb{N} = S_1 \cup S_2 \cup \cdots \cup S_w$ be a partition of $\mathbb{N}$, where each $1_{S_v}(n)$ has zero anqie entropy for $v = 1, \ldots, w$. Suppose that $p_1(y), \ldots, p_w(y)$ are polynomials in $\mathbb{R}[y]$. Let
\[
g(n) = \sum_{v=1}^{w} 1_{S_v}(n)p_v(n).
\]
Suppose that $f(n)$ is an arithmetic function satisfying
\[
\lim_{n \to \infty} \|\Delta^k f(n) - g(n)\| = 0
\]
for some $k \in \mathbb{N}$, then $\mathbb{E}(e(f(n))) = 0$.

**Proof.** Given $N \geq 1$, by Lemma 3.1, for each integer $m \geq 1$, there is a sufficiently large $L_m \in \mathbb{N}$ with $2^m|L_m$, such that whenever $n \geq L_m$, we have
\[
\|f(n + j) - Y_n(n + j)\| \leq 1/N, \quad j = 0, \ldots, 2^m - 1,
\]
where $Y_n(n + j)$ is defined to be $f(n + j)$ when $0 \leq j \leq k - 1$ and the value determined by the following linear equations when $k \leq j \leq 2^m - 1$,
\[
\Delta^k Y_n(n + j) = g(n + j), \quad j = 0, \ldots, 2^m - k - 1.
\]
Moreover, we may further assume that the sequence $\{L_m\}_{m=0}^{\infty}$ ($L_0 = 0$) chosen above satisfies $L_{m+1} > L_m$ for each $m \geq 1$. Let $d_m = (L_{m+1} - L_m)/2^m$. Then the following is a partition of $\mathbb{N}$,
\[
\mathbb{N} = \bigcup_{m=0}^{\infty} \bigcup_{a=0}^{d_m-1} \{L_m + a2^m, L_m + a2^m + 1, \ldots, L_m + a2^m + 2^m - 1\}.
\]
We define the arithmetic function $g_N$ as follows: for $a = 0, \ldots, d_m - 1$ and $j = 0, \ldots, 2^m - 1$,
\[
g_N(L_m + a2^m + j) = \frac{t}{N}
\]
when $\{Y_{L_m+a2^m}(L_m + a2^m + j)\} \in \left[\frac{t}{N}, \frac{t+1}{N}\right]$ for some integer $t$ with $0 \leq t \leq N - 1$. By formula (3.8),
\[
\|f(L_m + a2^m + j) - g_N(L_m + a2^m + j)\| < 2/N.
\]
Then $\sup_{n \in \mathbb{N}} \|f(n) - g_N(n)\| < 2/N$. Hence $\lim_{N \to \infty} \sup_{n \in \mathbb{N}} |e(f(n)) - e(g_N(n))| = 0$. So by Proposition 1.2 and Theorem 1.3, to prove $\mathbb{E}(e(f(n))) = 0$, it suffices to prove $\mathbb{E}(g_N) = 0$ for each $N \geq 1$. 

In the remaining part of this proof, we shall prove \( \mathcal{A}(g_N) = 0 \) for any given \( N \geq 1 \). Let \( \eta \) be the function defined as \( \eta(n) = v \) if \( n \in S_v, \ v = 1, \ldots, w \). Then the anqi entropy of \( \eta \) is zero. Note that \( \eta(n) \) has finite range. By formula (2.1),

\[
\lim_{m \to \infty} \frac{\log |\mathcal{B}_{2^m}(\eta)|}{2^m} = 0,
\]

where \( \mathcal{B}_{2^m}(\eta) \) is the set of all \( 2^m \)-blocks occurring in \( \{\eta(n)\}_{n=0}^{\infty} \). For any \( 2^m \)-block \( \nu \) in \( \mathcal{B}_{2^m}(\eta) \), denote by

\[
\mathcal{A}_{\nu,m} = \{(g_N(n2^m), \ldots, g_N(n2^m + 2^m - 1)) : n \in \mathbb{N}, \ n2^m \geq L_m, \ (\eta(n2^m), \ldots, \eta(n2^m + 2^m - 1)) = \nu\}.
\]

Recall that \( \mathcal{B}_{2^m}^{x,r}(g_N) \) denotes the set of all \( 2^m \)-regularly effective blocks occurring in \( g_N \). Then

\[
\mathcal{B}_{2^m}^{x,r}(g_N) \subseteq \bigcup_{\nu \in \mathcal{B}_{2^m}(\eta)} \mathcal{A}_{\nu,m}.
\]

In the following, we estimate the cardinality of \( \mathcal{A}_{\nu,m} \) for each given \( m \) with \( 2^m \geq \max(k, 1, d) \) and \( \nu = (\nu_0, \ldots, \nu_{2^m-1}) \) in \( \mathcal{B}_{2^m}(\eta) \). Let \( d = \max(\text{deg}(p_1), \ldots, \text{deg}(p_w)) + 1 \), where we define \( \text{deg}(p_i) \) as \(-1\) when \( p_i = 0 \). Denote by \( q = k + wd \).

We first define linear functions \( F_{\nu,0}, \ldots, F_{\nu,2^m-1} \) from \( \mathbb{R}^q \) to \( \mathbb{R} \). Define \( F_{\nu,j}(x_0, x_1, \ldots, x_{q-1}) = x_j \) for \( j = 0, \ldots, k - 1 \). Assume inductively that we have defined the linear function \( F_{\nu,j_0} \) for some \( j_0 \) with \( k - 1 \leq j_0 \leq 2^m - 2 \). Then we define \( F_{\nu,j_0+1}(x_0, x_1, \ldots, x_{q-1}) \) to be the function satisfying

\[
\sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} F_{\nu,j_0+1-k+l}(x_0, x_1, \ldots, x_{q-1}) = P_{\nu,j_0+1-k}(j_0 + 1 - k), \tag{3.12}
\]

where

\[
P_{\nu}(j_0 + 1 - k) = \sum_{r=0}^{d-1} x_{k+(v-1)d+r} \prod_{0 \leq s \leq d-1, s \neq r} \frac{(j_0 + 1 - k) - s}{r - s}, \quad 1 \leq v \leq w, \tag{3.13}
\]

when \( d \geq 2 \); \( P_{\nu}(j_0 + 1 - k) = 0 \) when \( d = 0 \); \( P_{\nu}(j_0 + 1 - k) = x_{k+(v-1)d} \) when \( d = 1 \). By equation (3.12), it follows from the inductive assumption that \( F_{\nu,j_0+1}(x_0, x_1, \ldots, x_{q-1}) \) is a linear function of \( x_0, x_1, \ldots, x_{q-1} \).

Next, given \( n \) with \( (\eta(n2^m), \eta(n2^m + 1), \ldots, \eta(n2^m + 2^m - 1)) = \nu \) and \( n2^m \geq L_m \). Suppose that \( L_{m_0} \leq n2^m < L_{m_0+1} \) for some \( m_0 \geq m \) and \( n2^m = n_02^{m_0} + u \) with \( 0 \leq u \leq 2^{m_0} - 2^m \). Taking \( y_j = \{Y_{n_02^{m_0}}(n2^m + j)\} \) for \( 0 \leq j \leq k-1 \) and \( y_{k+(v-1)d+r} = \{P_{\nu}(n2^m + r)\} \) for \( 1 \leq v \leq w, \ 0 \leq r \leq d-1 \). In the following, we show that

\[
\{F_{\nu,j}(y_0, y_1, \ldots, y_{q-1})\} = \{Y_{n_02^{m_0}}(n2^m + j)\}, \ j = 0, \ldots, 2^m - 1. \tag{3.14}
\]

By the definition,

\[
F_{\nu,j}(y_0, y_1, \ldots, y_{q-1}) = \{Y_{n_02^{m_0}}(n2^m + j)\}, \ j = 0, \ldots, k - 1. \tag{3.15}
\]
Plugging \((y_0, y_1, \ldots, y_{q-1})\) into equation (3.13), we have for \(j \geq k\),

\[
P_v(j - k) = \sum_{r=0}^{d-1} \{p_v(n2^m + r)\} \prod_{0 \leq s \leq d-1, \ s \neq r} \frac{(j - k) - s}{r - s}.
\]

By Lemma A.4, \(\{P_v(j - k)\} = \{p_v(n2^m + j - k)\}\). Then by equation (3.12),

\[
\{\sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} F_{v,j-k+l}(y_0, y_1, \ldots, y_{q-1})\} = \{p_v(n2^m + j - k)\}, \ j = k, \ldots, 2^m - 1. \tag{3.16}
\]

Using the condition (3.9), we have

\[
\sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} Y_m(n2^m + j - k + l) = g(n2^m + j - k) = p_v(n2^m + j - k), \ j = k, \ldots, 2^m - 1.
\]

Comparing with equation (3.16) and by (3.15), we conclude that equation (3.14) holds.

Note that \(|j - k| \leq 2^m\). Then by equation (3.13), \(|P_v(j - k)| \leq (d + 1)2^{md}\) when \(d = 0\), or \(d \geq 1\), \(x_{k+(v-1)d+r} \in [0, 1], r = 0, \ldots, d - 1\). By equation (3.12) and Proposition A.5, we have

\[
|F_{v,j}(y_0, y_1, \ldots, y_{q-1})| \leq (k+1)j^k(d+1)^{2md} < (k+1)(d+1)^{2m(k+d)} - 1 \leq M \leq (k+1)(d+1)^{2m(k+d)}, 0 \leq t \leq N - 1, 0 \leq j \leq 2^m - 1. \tag{3.17}
\]

Based on the linear functions \(F_{v,0}, \ldots, F_{v,2^m-1}\) constructed above, we define a family of hyperplanes \(\mathcal{H} = \{H_{M,t,j} : M, t, j \in \mathbb{Z}, -(k+1)(d+1)^{2m(k+d)} - 1 \leq M \leq (k+1)(d+1)^{2m(k+d)}, 0 \leq t \leq N - 1, 0 \leq j \leq 2^m - 1\}\), where

\[
H_{M,t,j} = \{(x_0, x_1, \ldots, x_{q-1}) \in \mathbb{R}^q : F_{v,j}(x_0, x_1, \ldots, x_{q-1}) = M + \frac{t}{N}\}.
\]

Then \(|\mathcal{F}| \leq 2(k + 2)(d + 1)2^{m(k+d)+1}\). By Lemma 2.3, there are at most \(W(|\mathcal{F}|, q)\) pieces of \(\mathbb{R}^q\) cut by the hyperplanes in \(\mathcal{F}\), where

\[
W(|\mathcal{F}|, q) \leq (q + 1)2^q(2(k + 2)(d + 1)2^{m(k+d)+1})^q. \tag{3.18}
\]

Now, we are ready to estimate \(|\mathcal{A}_{v,m}|\). Let \((g_N(n2^m), \ldots, g_N(n2^m + 2^m - 1)) \in \mathcal{A}_{v,m}\) with \(n2^m \geq L_m\). Then \((\eta(n2^m), \ldots, \eta(n2^m + 2^m - 1)) = \nu\). Suppose that \(L_m \geq n2^m < L_{m+1}\) for some \(m \geq m\) and \(n2^m = n_02^{m_0} + u\) with \(0 \leq u \leq 2^{m_0} - 2^m\). Set

\[
y_j = \{Y_{n_02^{m_0}}(n2^m + j)\}
\]

for \(0 \leq j \leq k - 1\) and

\[
y_{k+(v-1)d+r} = \{p_v(n2^m + r)\}
\]

for \(1 \leq v \leq w, 0 \leq r \leq d - 1\). By formula (3.17), there are integers \(M_0, M_1, \ldots, M_{2^m-1} \in [-(k+1)(d+1)^{2m(k+d)} - 1, (k+1)(d+1)^{2m(k+d)}]\) and \(t_0, t_1, \ldots, t_{2^m-1} \in [0, N - 1]\) such that

\[
M_j + \frac{t_j}{N} \leq F_{v,j}(y_0, y_1, \ldots, y_{q-1}) < M_j + \frac{t_j + 1}{N}, \ j = 0, \ldots, 2^m - 1. \tag{3.19}
\]

Note that any two pieces of \(\mathbb{R}^q\) cut by hyperplanes in \(\mathcal{F}\) are disjoint. Let \(P\) be the unique piece containing the point \((y_0, y_1, \ldots, y_{q-1})\). Then it is not hard to check that formula (3.19) holds.
for each \((x_0, x_1, \ldots, x_{q-1}) \in P\). By equation (3.14), \(\{F_{\nu,j}(y_0, y_1, \ldots, y_{q-1})\} = \{Y_{\nu,2^m}(n2^m + j)\} \in [\frac{t_i}{N}, \frac{t_i+1}{N}]\). Then by formula (3.10), \(g_N(n2^m + j) = \frac{t_i}{N}, \ j = 0, \ldots, 2^m - 1\). Moreover, from the above analysis, we conclude that if \(n, n' \in \mathbb{N}\) with \((\eta(n2^m), \ldots, \eta(n2^m + 2^m - 1)) = (\eta(n'2^m), \ldots, \eta(n'2^m + 2^m - 1)) = \nu\), such that the corresponding points \((y_0, y_1, \ldots, y_{q-1})\) and \((y'_0, y'_1, \ldots, y'_{q-1})\) belong to the same piece of \(\mathbb{R}^k\), then \((g_N(n2^m), \ldots, g_N(n2^m + 2^m - 1)) = (g_N(n'2^m), \ldots, g_N(n'2^m + 2^m - 1))\). Hence

\[|A_{\nu,m}| \leq W(|F|, q).\]

So by equation (3.11) and formula (3.18),

\[|B^{r,c}_{2^m}(g_N)| \leq \sum_{\nu \in B_{2^m}(\eta)} |A_{\nu,m}| \leq (q+1)2^q(2(k+2)(d+1)2^m(k+d+1)N)^q|B_{2^m}(\eta)|. \quad (3.20)\]

Note that \(N, k, d, q\) are parameters independent of \(m\). By Theorem 2.1,

\[\mathcal{E}(g_N) = \lim_{m \to \infty} \frac{\log |B^{r,c}_{2^m}(g_N)|}{2^m} \leq \lim_{m \to \infty} \frac{\log |B_{2^m}(\eta)|}{2^m} = \mathcal{E}(\eta) = 0.\]

So we complete the proof of this theorem. \(\square\)

Next, we prove Theorem 1.7, which discusses about the anqie entropy of exponential functions of concatenations.

**Theorem 1.7 (restated).** Let \(g(n)\) be defined as in Theorem 1.5. Suppose that \(\{f_i(n)\}_{i=0}^\infty\) is a sequence of real-valued arithmetic functions such that

\[\lim_{n \to \infty} \sup_{i \in \mathbb{N}} \|\Delta^k f_i(n) - g(n)\| = 0 \quad (3.21)\]

holds for some \(k \in \mathbb{N}\). Let \(\{N_i\}_{i=0}^\infty\) be an increasing sequence of natural numbers with \(N_0 = 0\) and \(\lim_{i \to \infty} (N_{i+1} - N_i) = \infty\). Then for the concatenation \(f(n)\) of \(\{f_i(n)\}_{i=0}^\infty\) with respect to \(\{N_i\}_{i=0}^\infty\), \(\mathcal{E}(e(f(n))) = 0\).

**Proof.** Given \(N \geq 1\). By condition (3.21) and Lemma 3.1, for each integer \(m \geq 1\), there is a sufficiently large \(L_m \in \mathbb{N}\) with \(2^m|L_m|\), such that for any \(i \in \mathbb{N}\), whenever \(n \geq L_m - 1\), we have

\[\|f_i(n + j) - Y_{n,i}(n + j)\| \leq 1/N, \ j = 0, 1, \ldots, 2^m - 1,\]

where \(Y_{n,i}(n + j)\) equals \(f_i(n + j)\) when \(0 \leq j \leq k - 1\), and is determined by the following linear equations when \(k \leq j \leq 2^m - 1\),

\[\Delta^k Y_{n,i}(n + j) = g(n + j), \ j = 0, 1, \ldots, 2^m - k - 1. \quad (3.23)\]

Moreover, we may further assume that the sequence \(\{L_m\}_{m=0}^\infty (L_0 = 0)\) chosen above satisfies \(L_{m+1} > L_m\) for each \(m \geq 1\). Let \(d_m = (L_{m+1} - L_m)/2^m\). For any \(N \geq 1\), we define a sequence \(\{g_{N,i}\}_{i=0}^\infty\) of arithmetic functions with finite ranges as follows: for \(a = 0, 1, \ldots, d_m - 1\) and \(j = 0, 1, \ldots, 2^m - 1\), define

\[g_{N,i}(L_m + a2^m + j) = \frac{t_j}{N}\]
when \( \{Y_{m,a2^m,i}(L_m + a2^m + j)\} \in [\frac{1}{N}, \frac{t+1}{N}] \) for some integer \( t \) with \( 0 \leq t \leq N - 1 \). By formula (3.22),

\[
\|f_i(L_m + a2^m + j) - g_{N,i}(L_m + a2^m + j)\| < 2/N.
\]

Then

\[
\|f_i(n) - g_{N,i}(n)\| < 2/N, \text{ for any } n, \ i \in \mathbb{N}.
\] (3.24)

Denote by

\[
\mathcal{C}_m = \{(g_{N,i}(n^{2^m}),...,g_{N,i}(n^{2^m + 2^m - 1})) : i \in \mathbb{N}, n \in \mathbb{N}, n^{2^m} \geq L_m\}.
\]

Recall that

\[
g(n) = \sum_{v=1}^{w} 1_{S_v}(n)p_v(n),
\] (3.25)

where each \( 1_{S_v}(n) \) has zero anique entropy. Define the arithmetic function \( \eta \) by \( \eta(n) = v \) if \( n \in S_v \), \( v = 1, \ldots, w \). Then \( \mathcal{E}(\eta) = 0 \). Recall that \( \mathcal{B}_{2^m}(\eta) \) denotes the set of all \( 2^m \)-blocks occurring in \( \{\eta(n)\}_{n=0}^{\infty} \). Given \( \nu \in \mathcal{B}_{2^m}(\eta) \), denote by \( \tilde{A}_{\nu,m} \) the set

\[
\{(g_{N,i}(n^{2^m}),...,g_{N,i}(n^{2^m + 2^m - 1})) : i \in \mathbb{N}, n \in \mathbb{N}, n^{2^m} \geq L_m, (\eta(n^{2^m}),...,\eta(n^{2^m + 2^m - 1})) = \nu\}.
\]

Then

\[
\mathcal{C}_m \subseteq \bigcup_{\nu \in \mathcal{B}_{2^m}(\eta)} \tilde{A}_{\nu,m}.
\] (3.26)

It follows, from a similar argument to the proof of formula (3.20) in Theorem 1.5, that

\[
|\mathcal{C}_m| \leq \sum_{\nu \in \mathcal{B}_{2^m}(\eta)} |\tilde{A}_{\nu,m}| \leq (g + 1)2^q(2(k + 2)(d + 1)2^{m(k+d+1)}N)^q|\mathcal{B}_{2^m}(\eta)|,
\] (3.27)

where \( d = \max(deg(p_1),...,deg(p_w)) + 1 \) and \( q = k + wd \).

Suppose that \( f(n) \) is the concatenation of \( \{f_i(n)\}_{i=0}^{\infty} \) with respect to \( \{N_i\}_{i=0}^{\infty} \). Let \( g_{N}(n) \) be the concatenation of \( \{g_{N,i}(n)\}_{i=0}^{\infty} \) with respect to \( \{N_i\}_{i=0}^{\infty} \), i.e.,

\[
g_{N}(n) = g_{N,i}(n) \quad \text{if} \quad N_i \leq n < N_{i+1}.
\]

By formula (3.24),

\[
\|g_{N}(n) - f(n)\| < 2/N, \text{ for any } n \in \mathbb{N}.
\]

This implies that \( \lim_{N \to \infty} \sup_{n \in \mathbb{N}} |e(g_{N}(n)) - e(f(n))| = 0 \). So by Proposition 1.2 and Theorem 1.3, to prove \( \mathcal{E}(e(f(n))) = 0 \), it suffices to prove \( \mathcal{E}(e(g_{N})) = 0 \) for each \( N \geq 1 \).

In the following, to show \( \mathcal{E}(g_{N}) = 0 \), we estimate \( |\mathcal{B}_{2^m}^{e,r}(g_{N})| \) for any given \( m \geq 1 \), the cardinality of all regularly effective \( 2^m \)-blocks occurring in \( g_{N} \). Let \( i_m \) be large enough such that \( N_{i+1} - N_i > 2^m \) whenever \( i > i_m \). Let \( (g_{N}(n^{2^m}),g_{N}(n^{2^m+1}),...,g_{N}(n^{2^m+2^m-1})) \) be a \( 2^m \)-block in \( g_{N} \) with \( n^{2^m} > \max\{L_m,N_{i_m}\} \). It is easy to see that there are two cases about this block: one case is that there is an \( i_n \) such that \( (g_{N}(n^{2^m}),g_{N}(n^{2^m+1}),...,g_{N}(n^{2^m+2^m-1})) = (g_{N,i_n}(n^{2^m}),g_{N,i_n}(n^{2^m+1}),...,g_{N,i_n}(n^{2^m+2^m-1})) \); the other case is that there are two natural numbers \( i_n \) and \( j_n \) such that \( g_{N}(n^{2^m} + j) = g_{N,i_n}(n^{2^m} + j) \) when \( j = 0,1,...,j_n - 1 \) and \( g_{N}(n^{2^m} + j) = g_{N,i_n+1}(n^{2^m} + j) \) when \( j = j_n,j_n+1,...,2^m - 1 \). So \( |\mathcal{B}_{2^m}^{e,r}(g_{N})| \) is less than or
equal to $2^m \times |C_m|^2$. Note that $N, k, d, q$ are parameters independent of $m$. Then by formula (3.27),

$$\mathcal{A}(g_N) = \lim_{m \to \infty} \frac{\log |\mathcal{B}_{2^m}(g_N)|}{2^m} \leq \lim_{m \to \infty} \frac{\log(|\mathcal{B}_{2^m}(\eta)|^2)}{2^m} = 2\mathcal{A}(\eta) = 0.$$ 

Now, we complete the proof of the theorem. □

Using a similar idea to the proof of Theorem 1.5, we obtain the following proposition that gives many characteristic functions with zero ancie entropy.

**Proposition 3.2.** Let $p_1(y), p_2(y) \in \mathbb{R}[y]$. Suppose that

$$S = \{n \in \mathbb{N} : \{p_1(n)\} < \{p_2(n)\}\}.$$ 

Then $1_S(n)$, the characteristic function defined on $S$, has zero ancie entropy.

**Proof.** Assume that the degrees of $p_1(n)$ and $p_2(n)$ are both less than $k$. Then $\Delta^k p_1(n) = \Delta^k p_2(n) = 0$. In the following, for any given integer $J \geq k + 1$, we estimate $|\mathcal{B}_f(1_S)|$, the cardinality of the set of all $J$-blocks occurring in $1_S$.

Firstly, we define linear functions $F_0, F_1, \ldots, F_{J-1} : \mathbb{R}^k \to \mathbb{R}$. Define $F_j(x_0, x_1, \ldots, x_{k-1})$ to be $x_j$ when $j = 0, 1, \ldots, k - 1$. Assume inductively that we have defined $F_{j_0}(x_0, x_1, \ldots, x_{k-1})$ for some $j_0 \geq k - 1$. Then define $F_{j_0+1}(x_0, x_1, \ldots, x_{k-1})$ to be the linear function satisfying

$$\left(\sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} F_{j_0+1-k+l}(x_0, x_1, \ldots, x_{k-1}) = 0, \right) \quad (3.28)$$

equivalently,

$$F_{j_0+1}(x_0, x_1, \ldots, x_{k-1}) = -\sum_{l=0}^{k-1} (-1)^{k-l} \binom{k}{l} F_{j_0+1-k+l}(x_0, x_1, \ldots, x_{k-1}).$$

By Proposition A.1, it is not hard to see that for $i = 1, 2$ and any $n \in \mathbb{N}$,

$$\{F_j(\{p_i(n)\}, \{p_i(n+1)\}, \ldots, \{p_i(n+k-1)\})\} = \{p_i(n+j)\}, \quad j = 0, 1, \ldots, J - 1. \quad (3.29)$$

By Proposition A.5,

$$|F_j(\{p_i(n)\}, \{p_i(n+1)\}, \ldots, \{p_i(n+k-1)\})| < (k + 1)J^k, \quad j = 0, 1, \ldots, J - 1. \quad (3.30)$$

Secondly, we define a family of hyperplanes $\mathcal{F} = \{H_{L,j}, H_{M,j}, H_{N,j} : j, L, M, N \in \mathbb{Z}, 0 \leq j \leq J - 1, -2(k + 1)J^k - 1 \leq L \leq 2(k + 1)J^k, -(k + 1)J^k - 1 \leq M, N \leq (k + 1)J^k \}$ in $\mathbb{R}^{2k}$, where $H_{L,j} = \{(x_0, x_1, \ldots, x_{2k-1}) \in \mathbb{R}^{2k} : F_j(x_0, x_1, \ldots, x_{k-1}) - F_j(x_k, x_{k+1}, \ldots, x_{2k-1}) = L\}$, $H_{M,j} = \{(x_0, x_1, \ldots, x_{2k-1}) \in \mathbb{R}^{2k} : F_j(x_0, x_1, \ldots, x_{k-1}) = M\}$, and $H_{N,j} = \{(x_0, x_1, \ldots, x_{2k-1}) \in \mathbb{R}^{2k} : F_j(x_k, x_{k+1}, \ldots, x_{2k-1}) = N\}$.

Then

$$|\mathcal{F}| \leq 8(k + 2)J^{k+1}.$$
By Lemma 2.3, there are at most $W(|\mathcal{F}|, 2k)$ pieces of $\mathbb{R}^{2k}$ cut by the hyperplanes in $\mathcal{F}$, where

$$W(|\mathcal{F}|, 2k) \leq 8^{2k}(2k + 1)2^{2k}(k + 2)^2J^{2k(k+1)}.$$  

Let $\mathcal{P}$ be the set of all pieces of $\mathbb{R}^{2k}$ cut by the hyperplanes in $\mathcal{F}$. Recall that $\mathcal{B}_J(1_S)$ denotes the set of all $J$-blocks occurring in $\{1_S(n)\}_{n=0}^\infty$. Denote by $|\mathcal{B}_J(1_S)| = C_J$. Suppose that $\mathcal{B}_J(1_S) = \{B_1, \ldots, B_{C_J}\}$. Let $n_m = \min\{n \in \mathbb{N} : (1_S(n), 1_S(n + 1), \ldots, 1_S(n + J - 1)) = B_m\}$, for $m = 1, \ldots, C_J$. Then $B_m = (1_S(n_m), 1_S(n_m + 1), \ldots, 1_S(n_m + J - 1))$. Define the map

$$\psi: \mathcal{B}_J(1_S) \to \mathcal{P}$$

by $\psi(B_m) = P_m$, where $P_m$ is the unique piece in $\mathcal{P}$ containing the point $(p_1(n_m), \ldots, p_1(n_m + k - 1), p_2(n_m), \ldots, p_2(n_m + k - 1))$, for $m = 1, \ldots, C_J$. Since any two pieces in $\mathcal{P}$ are disjoint, $\psi$ is well-defined.

In the following, we show that $\psi$ is injective. Given $n \in \mathbb{N}$, let

$$(y_0, \ldots, y_{k-1}, y_k, \ldots, y_{2k-1}) = (p_1(n), \ldots, p_1(n + k - 1), p_2(n), \ldots, p_2(n + k - 1)).$$

Suppose that $(y_0, \ldots, y_{k-1}, y_k, \ldots, y_{2k-1}) \in P$, a piece of $\mathbb{R}^{2k}$ cut by the hyperplanes in $\mathcal{F}$. Then by (3.30), there are integers $M_0, \ldots, M_{J-1}, N_0, \ldots, N_{J-1} \in [-2(k+1)J^{k-1}, (k+1)J^k]$ and $L_0, \ldots, L_{J-1} \in [-2(k+1)J^{k-1}, (k+1)J^k]$, such that

$$L_j \leq F_j(y_0, \ldots, y_{k-1}) - F_j(y_k, \ldots, y_{2k-1}) < L_j + 1, \quad (3.31)$$

and

$$M_j \leq F_j(y_0, \ldots, y_{k-1}) < M_j + 1, \quad N_j \leq F_j(y_k, \ldots, y_{2k-1}) < N_j + 1. \quad (3.32)$$

Moreover, the above inequalities also hold for each point in $P$. From formulas (3.31) and (3.32), it is not hard to see that for any given $j$ with $0 \leq j \leq J - 1$,

$$\{F_j(x_0, x_1, \ldots, x_{k-1})\} < \{F_j(x_k, x_{k+1}, \ldots, x_{2k-1})\}$$

holds for all $(x_0, x_1, \ldots, x_{2k-1}) \in P$ (when $L_j = M_j - N_j - 1$) or

$$\{F_j(x_0, x_1, \ldots, x_{k-1})\} \geq \{F_j(x_k, x_{k+1}, \ldots, x_{2k-1})\}$$

holds for all $(x_0, x_1, \ldots, x_{2k-1}) \in P$ (when $L_j = M_j - N_j$). Then by equation (3.29), we conclude that if $(p_1(n), \ldots, p_1(n + k - 1), p_2(n), \ldots, p_2(n + k - 1)), (p_1(n'), \ldots, p_1(n' + k - 1), p_2(n'), \ldots, p_2(n' + k - 1)) \in P$, then $(1_S(n), \ldots, 1_S(n + J - 1)) = (1_S(n'), \ldots, 1_S(n' + J - 1))$. So $\psi$ is injective. Hence $|\mathcal{B}_J(1_S)| = |\psi(\mathcal{B}_J(1_S))| \leq W(|\mathcal{F}|, 2k) \leq 8^{2k}(2k + 1)2^{2k}(k + 2)^kJ^{2k(k+1)}$. Then by formula (2.1),

$$\mathcal{E}(1_S) = \lim_{J \to \infty} \frac{\log |\mathcal{B}_J(1_S)|}{J} = 0,$$

as claimed. \hfill \Box

As an application of the above proposition, we give the following example that satisfies the condition in Theorem 1.5 with $g(n) = \Delta^2 f(n) \neq 0$. 


Example 3.3. Let \( f(n) = \sqrt{3n}\{\sqrt{2}n\} \). Then
\[
\Delta^2 f(n) = \begin{cases} 
2\sqrt{3}(\sqrt{2} - 1), & n \in S_1, \\
2\sqrt{3}(\sqrt{2} - 1) + \sqrt{3}n, & n \in S_3, \\
2\sqrt{3}(\sqrt{2} - 2), & n \in S_2, \\
2\sqrt{3}(\sqrt{2} - 2) - \sqrt{3}n, & n \in S_4,
\end{cases}
\]
where \( S_1 = \{n \in \mathbb{N} : \{\sqrt{2}(n + 2)\} > \{\sqrt{2}(n + 1)\} > \{\sqrt{2}n\}\} \), \( S_2 = \{n \in \mathbb{N} : \{\sqrt{2}(n + 2)\} < \{\sqrt{2}(n + 1)\} < \{\sqrt{2}n\}\} \), \( S_3 = \{n \in \mathbb{N} : \{\sqrt{2}(n + 2)\} > \{\sqrt{2}(n + 1)\}, \{\sqrt{2}(n + 1)\} < \{\sqrt{2}n\}\} \), \( S_4 = \{n \in \mathbb{N} : \{\sqrt{2}(n + 2)\} < \{\sqrt{2}(n + 1)\}, \{\sqrt{2}(n + 1)\} > \{\sqrt{2}n\}\} \).

4. Connection with Sarnak’s conjecture

In this section, we investigate some properties of the distribution of the Möbius function under the assumption of Sarnak’s Möbius disjointness conjecture. We first deduce some results on Möbius disjointness of certain exponential functions. From Theorem 1.4, we know that SMDC holds if and only if \( \mu \) is disjoint from all arithmetic functions with zero anqie entropy. In the previous section, we showed that many exponential functions have zero anqie entropy. So as an immediate consequence of Theorem 1.5 and Proposition 1.2, we obtain proposition 1.8 which states that SMDC implies that \( \mu(n) \) is disjoint from \( e(f(n)) \) with \( f(n) \) satisfying condition (1.3).

Using Theorem 1.7, we have the following proposition.

Proposition 4.1. Let \( w \) be a positive integer and \( \mathbb{N} = S_1 \cup S_2 \cup \cdots \cup S_w \) be a partition of \( \mathbb{N} \), where each \( 1_{S_i}(n) \) has zero anqie entropy. Suppose that \( p_1(y), \ldots, p_w(y) \) are polynomials in \( \mathbb{R}[y] \). Let
\[
g(n) = \sum_{v=1}^{w} 1_{S_v}(n)p_v(n).
\]
Suppose that \( \{f_i(n)\}_{i=0}^{\infty} \) is a sequence of real-valued arithmetic functions such that
\[
\lim_{n \to \infty} \sup_{i \in \mathbb{N}} \|\Delta^k f_i(n) - g(n)\| = 0 \tag{4.1}
\]
holds for some \( k \geq 1 \). Let \( \{N_i\}_{i=0}^{\infty} \) be an increasing sequence of natural numbers with \( N_0 = 0 \) and \( \lim_{i \to \infty}(N_{i+1} - N_i) = \infty \). Then SMDC implies that, for any arithmetic function \( \eta(n) \) with zero anqie entropy,
\[
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} \mu(n)\eta(n)e(f_i(n)) \right| = 0.
\]

Proof. Let \( \{\theta_i\}_{i=0}^{\infty} \) be a sequence of numbers in \([0, 1]\) such that
\[
\left| \sum_{N_i \leq n < N_{i+1}} \mu(n)\eta(n)e(f_i(n)) \right| = \left( \sum_{N_i \leq n < N_{i+1}} \mu(n)\eta(n)e(f_i(n)) \right) e(\theta_i).
\]
Define \( \tilde{f}_i(n) = f_i(n) + \theta_i \). Then \( \{\tilde{f}_i(n)\}_{i=0}^{\infty} \) satisfies condition (4.1) since \( k \geq 1 \). Let \( \tilde{f}(n) \) be the concatenation of \( \{\tilde{f}_i(n)\}_{i=0}^{\infty} \) with respect to the sequence \( \{N_i\}_{i=0}^{\infty} \). Let \( f(n) = e(\tilde{f}(n))\eta(n) \).
By Proposition 1.2 and Theorem 1.7, the anqie entropy of $f(n)$ is zero. Hence by Theorem 1.4, SMDC implies

$$
\lim_{m \to \infty} \frac{1}{N_m} \sum_{n=0}^{N_m-1} \mu(n) f(n) = \lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} \mu(n) \eta(n) e(f_i(n)) \right| = 0.
$$

\[ \square \]

Now we prove Theorem 1.9, which tells us that Sarnak’s conjecture reveals the correlation between Möbius and certain exponential functions in almost all short intervals. Before proving it, we need the following intermediate result.

**Lemma 4.2.** Let $w(n)$ be a bounded arithmetic function and $\mathcal{D}$ be a non-empty family consisting of arithmetic functions. Then the following two conditions are equivalent.

(i) For any increasing sequence $\{N_i\}_{i=0}^\infty$ of natural numbers with $N_0 = 0$ and $\lim_{i \to \infty} (N_{i+1} - N_i) = \infty$, and any sequence $\{f_i(n)\}_{i=0}^\infty$ in $\mathcal{D}$, we have

$$
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} w(n) e(f_i(n)) \right| = 0. \quad (4.2)
$$

(ii) We have

$$
\lim_{h \to \infty} \limsup_{X \to \infty} \frac{1}{X h} \int_X^{2X} \sup_{f \in \mathcal{D}} \left| \sum_{x \leq n < x+h} w(n) e(f(n)) \right| dx = 0. \quad (4.3)
$$

**Proof.** We first prove that $(i) \Rightarrow (ii)$. Assume on the contrary that formula $(4.3)$ does not hold. Then there is a $\delta \in (0, 1)$ and a sequence $\{h_j\}_{j=0}^\infty$ of positive integers with $\delta h_j > 1$ and $\lim_{j \to \infty} h_j = \infty$, such that

$$
\limsup_{X \to \infty} \frac{1}{X} \int_X^{2X} \sup_{f \in \mathcal{D}} \left| \sum_{x \leq n < x+h} w(n) e(f(n)) \right| dx > 2\delta h_j.
$$

Given $h_j$, choose $X_j$ large enough with $\delta X_j > 16h_j$ and $X_j > 4X_{j-1}$ satisfying

$$
\int_{X_j}^{2X_j} \sup_{f \in \mathcal{D}} \left| \sum_{x \leq n < x+h_j} w(n) e(f(n)) \right| dx > 2\delta X_j h_j.
$$

By the pigeonhole principal, there is a $y_j \in [0, h_j)$, such that

$$
\sum_{l_j = \lfloor X_j/h_j \rfloor - 1}^{\lfloor 2X_j/h_j \rfloor} \left( \sup_{f \in \mathcal{D}} \left| \sum_{n = l_j h_j + y_j}^{l_j + 1 h_j + y_j - 1} w(n) e(f(n)) \right| \right) > 2\delta X_j.
$$

Furthermore, for each $l_j$ with $\lfloor X_j/h_j \rfloor - 1 \leq l_j \leq \lfloor 2X_j/h_j \rfloor + 1$, we can find $g_{l_j}(n) \in \mathcal{D}$ such that
Now we construct \{N_i\}_{i=0}^{\infty} \text{ and } \{f_i(n)\}_{i=0}^{\infty} \text{ in the following way. Choose } N_0 = 0 \text{ and } N_i = [X_{J+1}/h_{J+1}]h_{J+1} + J_1 + j_{J+1} + y_{J+1} \text{ when } i = \sum_{j=1}^{J_i} [X_j/h_j] - J + t, t = 0, 1, \ldots, [X_{J+1}/h_{J+1}] - 2, \text{ where } J = 0, 1, \ldots. \text{ Then } \lim_{i \to \infty} (N_{i+1} - N_i) = \infty. \text{ For } J = 0, 1, \ldots, \text{ we choose } f_i(n) = g_i[X_{J+1}/h_{J+1}] + t(n) \text{ when } i = \sum_{j=1}^{J_i} [X_j/h_j] - J + t, t = 0, 1, \ldots, [X_{J+1}/h_{J+1}] - 3, \text{ and } f_i(n) = g_i[X_i/h_i](n) \text{ otherwise. Then by equation } (4.2), \text{ there is an } m_0 \text{ and } J_0 \text{ with } m_0 = \sum_{j=1}^{J_0-1} [X_j/h_j] - J_0 + 2, \text{ such that}

\[
\frac{1}{N_{m_0} + [X_{J_0}/h_{J_0}] - 2} \left( \sum_{i=0}^{m_0 + [X_{J_0}/h_{J_0}] - 3} \sum_{N_i \leq n < N_{i+1} - 1} w(n) e(f_i(n)) \right) < \frac{\delta}{4}. \tag{4.5}
\]

Note that \(N_{m_0} = [X_{J_0}/h_{J_0}]h_{J_0} + y_{J_0} \text{ and } N_{m_0} + [X_{J_0}/h_{J_0}] - 2 = 2([X_{J_0}/h_{J_0}] - 1)h_{J_0} + y_{J_0}. \text{ By formula } (4.4), \text{ we have that } \sum_{i=m_0}^{m_0 + [X_{J_0}/h_{J_0}] - 3} |\sum_{N_i \leq n < N_{i+1} - 1} w(n) e(f_i(n))| > \frac{\delta X_{J_0}}{2}. \text{ Thus the left side of formula } (4.5) \text{ is greater than } (\delta/2)X_{J_0}/2X_{J_0} - h_{J_0} > \frac{\delta}{4}. \text{ This contradicts the right side of formula } (4.5). \text{ Hence equation } (4.3) \text{ holds.}

Next, we show that } (ii) \Rightarrow (i). \text{ Given } \epsilon \text{ with } 0 < \epsilon < 1. \text{ Let } h \text{ be a fixed sufficiently large positive integer. By equation } (4.3) \text{ and a dyadic subdivision, there is an } m_0 \geq 1 \text{ such that whenever } m > m_0, \text{ we have } N_m - N_{m-1} > \frac{2}{\epsilon}h \text{ and}

\[
\left( \left\lfloor \frac{N}{h} \right\rfloor + 1 \right) h - 1 \sum_{x=0}^{\left\lfloor \frac{N}{h} \right\rfloor} \sup_{f \in \mathcal{F}} \left| \sum_{x \leq h < x+h} w(n) e(f(n)) \right| \leq \epsilon N_m h. \tag{4.6}
\]

Let } S_j = \{lh + j : l = 0, 1, \ldots, \left\lfloor \frac{N}{h} \right\rfloor \}, \text{ for } 0 \leq j \leq h - 1. \text{ Then, by } (4.6), \text{ there is a } j_0 \text{ such that}

\[
\sum_{x \in S_{j_0}} \sup_{f \in \mathcal{F}} \left| \sum_{x \leq n < x+h} w(n) e(f(n)) \right| \leq \epsilon N_m. \tag{4.7}
\]

Suppose that } S_{j_0} \cap [N_i, N_{i+1}) = \{x_1^{(i)}, x_2^{(i)}, \ldots, x_{l_i}^{(i)} \}, \text{ where } x_1^{(i)} < x_2^{(i)} < \cdots < x_{l_i}^{(i)}, i = 0, 1, \ldots, m-1. \text{ Then}

\[
\left| \sum_{N_i \leq n < N_{i+1}} w(n) e(f_i(n)) \right| \leq \sum_{i=1}^{l_i-1} \sup_{f \in \mathcal{F}} \left| \sum_{x \leq x^{(i)} < x^{(i)} + h} w(n) e(f(n)) \right| + 2h.
\]

Hence

\[
\sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} w(n) e(f_i(n)) \right| \leq \sum_{x \in S_{j_0}} \sup_{f \in \mathcal{F}} \left| \sum_{x \leq n < x+h} w(n) e(f(n)) \right| + 2mh.
\]
By formula (4.7),

\[
\frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_i+1} w(n)e(f_i(n)) \right| \leq \frac{eN_m}{N_m} + \frac{2mh}{N_m} < 2\epsilon.
\]

So we obtain equation (4.2).\[ \square \]

We are now ready to prove Theorem 1.9.

**Proof of Theorem 1.9.** Choose \( w(n) = \mu(n)\eta(n) \) in Lemma 4.2, where \( \eta(n) \) is an arithmetic function with zero anqie entropy. Then as a consequence of Proposition 4.1 and Lemma 4.2, we obtain Theorem 1.9.\[ \square \]

Now, we shall prove Theorem 1.10. The major tool we use here is that any zero anqie entropy function can be approached by zero anqie entropy functions with finite ranges. This follows from [Wei18a, Theorem 1.8] stated below.

**Lemma 4.3.** Suppose that \( f \) is a bounded arithmetic function with anqie entropy \( \lambda \) (\( 0 \leq \lambda < +\infty \)). Then for any \( N \geq 1 \), there is an arithmetic function \( f_N \) with finite range, such that \( \mathcal{E}(f_N) \leq \lambda \) and \( \sup_{n \in \mathbb{N}} |f_N(n) - f(n)| \leq \frac{1}{N} \).

**Proof of Theorem 1.10.** For the sufficient part, assume that equation (1.7) holds, that is, for any \( f \) with \( \mathcal{E}(f) = 0 \),

\[
\lim_{h \to \infty} \limsup_{X \to \infty} \frac{1}{Xh} \int_X^{2X} \left| \sum_{x \leq n < x+h} \mu(n)f(n) \right| dx = 0,
\]

then we show that \( \mu(n) \) is disjoint from \( f(n) \). For any given \( \epsilon > 0 \), from the above equality, there are sufficiently large positive integers \( h \) and \( X_0 \) such that whenever \( X > X_0 \),

\[
\frac{1}{Xh} \int_X^{2X} \left| \sum_{x \leq n < x+h} \mu(n)f(n) \right| < \epsilon. \tag{4.8}
\]

By the above formula and a dyadic subdivision, for \( N > X_0 \) large enough, we have

\[
\frac{1}{N} \sum_{n=1}^{N} \mu(n)f(n) = \frac{1}{Nh} \sum_{i=0}^{h-1} \sum_{n=1}^{N} \mu(n+i)f(n+i) + O\left(\frac{h}{N}\right)
\]

\[
= \frac{1}{Nh} \sum_{n=1}^{N} \sum_{n \leq m < n+h} \mu(m)f(m) + O\left(\frac{h}{N}\right)
\]

\[
= \frac{1}{Nh} \left( \sum_{n=1}^{X_0} \sum_{n \leq m < n+h} \mu(m)f(m) + \sum_{n=X_0+1}^{N} \sum_{n \leq m < n+h} \mu(m)f(m) \right) + O\left(\frac{h}{N}\right)
\]

\[
< 2\epsilon.
\]

Hence \( \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)f(n) = 0. \)
For the necessary part, let \( f(n) \) be any bounded arithmetic function with \( \mathcal{A}(f) = 0 \). In the following, we show that \( f(n) \) can be expressed as a linear combination of functions of the form \( e(h(n)) \) for some real-valued function \( h(n) \) with zero anqie entropy. Firstly \( f(n) = f_1(n) + f_2(n) \), \( f_1(n) = \frac{f(n) + f^*(n)}{2} \) and \( f_2(n) = \frac{f(n) - f^*(n)}{2} \). So \( f_1(n), f_2(n) \) are real-valued. Next we write \( f = (f_1^+ - f_1^-) + (f_2^+ - f_2^-) \), where \( f_j^+ = \frac{f_j + |f_j|}{2}, f_j^- = \frac{|f_j| - f_j}{2}, j = 1, 2, \) are non-negative valued. By Proposition 1.2, \( f_j^+ \), \( f_j^- \), \( j = 1, 2 \), are all have zero anqie entropy. Then \( f \) can be written as a linear combination of at most four non-negative valued functions with zero anqie entropy. We may assume that these functions are bounded by 1.

Suppose that \( \Phi \) is such a function, then \( \Phi = \frac{1}{2\sqrt{2}}(\phi_1 + \phi_2) + \frac{1}{2\sqrt{2}}(\phi_1 + \phi_2) \), where \( \phi_1 = \frac{\phi + \sqrt{1 - \phi^2}}{\sqrt{2}} + \frac{\phi - \sqrt{1 - \phi^2}}{\sqrt{2}}, \) and \( \phi_2 = \frac{\phi - \sqrt{1 - \phi^2}}{\sqrt{2}} + \frac{\phi + \sqrt{1 - \phi^2}}{\sqrt{2}}. \) Note that for \( j = 1, 2, |\phi_j(n)| = 1 \) for each \( n \in \mathbb{N} \), and \( \phi_j \) has zero anqie entropy by Proposition 1.2 again. By [Wei18a, Corollary 7.7], each \( \phi_j(n) \) can be written as \( e(h(n)) \) for some real-valued function \( h(n) \) with zero anqie entropy.

According to the above analysis, the proof of this theorem is reduced to proving that for each \( e(f(n)) \) with \( \mathcal{A}(f(n)) = 0 \), we have

\[
\lim_{h \to \infty} \limsup_{X \to \infty} \frac{1}{X} \int_X^{2X} \left| \sum_{x \leq n < x + h} \mu(n)e(f(n)) \right| \, dx = 0. \tag{4.9}
\]

Assume on the contrary that formula (4.9) does not hold. Then there is a \( \delta > 0 \) and a sequence \( \{h_j\}_{j=0}^\infty \) of positive integers with \( h_j \delta > 1 \) and \( \lim_{j \to \infty} h_j = \infty \), such that

\[
\limsup_{X \to \infty} \frac{1}{X} \int_X^{2X} \left| \sum_{x \leq n < x + h_j} \mu(n)e(f(n)) \right| \, dx > 2\delta h_j. \tag{4.10}
\]

By Lemma 4.3, there is a sequence of real-valued arithmetic functions \( \{f_N(n)\}_{N=0}^\infty \) with finite ranges and zero anqie entropy, such that \( \lim_{N \to \infty} \sup_{n \in \mathbb{N}} |e(f_N(n)) - e(f(n))| = 0. \) Then by formula (4.10), we can find an appropriate \( N' > 0 \) satisfying

\[
\limsup_{X \to \infty} \frac{1}{X} \int_X^{2X} \left| \sum_{x \leq n < x + h_j} \mu(n)e(f_{N'}(n)) \right| \, dx > \delta h_j.
\]

This contradicts the fact that

\[
\lim_{h \to \infty} \limsup_{X \to \infty} \frac{1}{X} \int_X^{2X} \left| \sum_{x \leq n < x + h} \mu(n)e(f_{N'}(n)) \right| \, dx = 0
\]

by Theorem 1.9 with the settings \( k = 1, g(n) = \Delta(f_{N'}) \) and \( \eta(n) = 1. \) So we obtain equation (4.9) and complete the proof of this theorem.

\[\square\]

5. Proof of Theorem 1.14

In this section, we shall prove Theorem 1.14. Actually, we can prove a more general version of it. Before stating the result, let us recall the following distance function introduced by
Granville and Soundararajan in [GS07]. For two multiplicative functions \( f(n) \) and \( g(n) \) with \(|f(n)|, |g(n)| \leq 1 \) for all \( n \geq 1 \),

\[
\mathbb{D}_s(f(n), g(n); X) := \left( \sum_{\substack{p \leq X \atop p|s}} \frac{1 - \text{Re}(f(p)g(p))}{p} \right)^{1/2}
\]

This distance function was used in [BGS13] to measure the pretentiousness between any multiplicative function \( f(n) \) and some function for which exceptional characters do exist. Throughout define

\[
M_s(f; X, T) := \inf_{\chi \mod s} \inf_{|t| \leq T} \mathbb{D}_s(f \chi, n \mapsto n^t; X)^2,
\]

\[
M(f; X, T, Y) := \inf_{1 \leq s \leq Y} \inf_{|t| \leq T} M_s(f; X, T) = \inf_{1 \leq s \leq Y} \inf_{|t| \leq T} \mathbb{D}_s(f \chi, n \mapsto n^t; X)^2.
\]

**Theorem 5.1.** Let \( s \geq 1 \) and \( H \geq 3 \) be integers. Let \( \beta(n) \) be a multiplicative function with \(|\beta(n)| \leq 1 \) for all \( n \geq 1 \). Let \( F : \mathbb{R}/\mathbb{Z} \to \mathbb{C} \) be a Lipschitz function. For \( N \) large enough with \( Hs \leq (\log N)^{1/32} \), then

\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \left| \sum_{1 \leq h \leq Hs \atop n+h \equiv a(s)} \beta(n+h)F(f(n+h)\mathbb{Z}) \right| \ll NHs(\log \log(s+2)) \frac{(\log \log H)^3}{(\log H)^{1/2}} \tag{5.1}
\]

\[
+ NHs(\log \log H)^2 \left( \left( M(\beta; N^{1/4}, N, \log N) + 1 \right) \exp (- M(\beta; N^{1/4}, N, \log N)) \right)^{1/2}.
\]

We remark that the constant implied in formula (5.1) would be uniformly for any \( F \) from the family consisting of all Lipschitz functions whose Lipschitz norm is bounded by a given constant.

Theorem 1.14 follows directly from Theorem 5.1 through the following well-known result (see, e.g., [MRT15, (1.12)]) about the “non-pretentious” nature of the Möbius function.

**Proposition 5.2.** Let \( N \) be large enough. Then for \( \varepsilon > 0 \) sufficiently small,

\[
M(\mu; N^{1/4}, N, \log N) \geq (1/3 - \varepsilon) \log \log N + O(1).
\]

In the rest part of this section, we shall prove Theorem 5.1. We now outline the strategy in our proof. Using the quantitative factorization Theorem 2.10, we split the proof into two cases, the major and minor arc cases. This point is more clear if the Lipschitz function \( F \) has zero mean. For this situation, the major arc part corresponds to \( g' = 0 \) (\( G' = \{0\} \)), while the minor arc part corresponds to \( g' \neq 0 \) (\( G' = \mathbb{R} \)) in Theorem 2.10. In the major arc case, we use second author’s work ([Wei18b, Theorem 1.7]) on sums of multiplicative functions over almost all short arithmetic progressions. In the minor arc case we use a variant of the arguments of He-Wang [HW19, Section 6] to obtain the cancellation we need.

Recall that we use \([N]\) to denote the set \( \{1, \ldots, N\} \) and \( \alpha \mathbb{Z} \) to denote an equivalence class in \( \mathbb{R}/\mathbb{Z} \) for any \( \alpha \in \mathbb{R} \). In the following, for integers \( n, a, s \), the notation \( n \equiv a(s) \) means \( n \equiv a(\mod s) \). In the literature of this section, many implicit constants \( O(1) = O_d(1) \geq 1 \) will
appear and may not be the same in different occurrences. For convenience, we use a common constant $C_0 = O_d(1)$ that is large enough for all these purposes.

For given $H$ sufficiently large, set

$$W = \log H,$$

and

$$B = 100C_0^2, \quad B_2 = \frac{1}{10}C_0^{-2}B.$$  \hfill (5.3)

Suppose $f(n) = \sum_{i=0}^{d} \alpha_i n^i$ with each $\alpha_i \in \mathbb{R}$ and $g(n, h) = f(n + h)$. Applying Theorem 2.10 to $g(n, h)$, there is a group $G' = \mathbb{R}$ or \{0\}, a set

$$\mathcal{N} \subseteq \lfloor N \rfloor$$

with $|\mathcal{N}| \geq (1 - W^{-B/2})N$, and a decomposition

$$g(n, h) = \mathcal{E}(n, h) + g'(n, h) + \gamma(n, h)$$

into polynomials $\mathcal{E}, g', \gamma : \mathbb{Z}^2 \to \mathbb{R}$ with the following properties:

(i). $|\mathcal{E}(n, h + 1) - \mathcal{E}(n, h)| \ll d \frac{1}{sH^W}$ for $h \in \lfloor Hs \rfloor$.

(ii). $g'(n, h)$ takes values in $G'$, and for any $n \in \mathcal{N}$, there are at least $(1 - W^{-B/2})s$ integers of $a \in [s]$ such that \{$(g'(n, h)\Gamma')_{h \in A_a}$\} is totally $W^{-B}$-equidistributed in $G'/\Gamma'$, where $A_a = \{1 \leq h \leq Hs : h \equiv a(s)\}$ and $\Gamma' = G' \cap \mathbb{Z}$.

(iii). For some integer $q$ with $1 \leq q \leq W$, \{$\gamma(n, h)\mathbb{Z} \}_{(n, h) \in \mathbb{Z}^2}$ is $qs$-periodic.

In addition, we assume that $q \in \lfloor W/2, W \rfloor$, otherwise we consider a multiplier of $q$. Let $\beta : \mathbb{N} \to \mathbb{C}$ be a multiplicative function with $|\beta(n)| \leq 1$ for each $n \in \mathbb{N}$ and let $F : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ be a Lipschitz function with $\|F\|_{\text{Lip}} \leq 1$ (see equation (2.7) for definition of $\|\cdot\|_{\text{Lip}}$).

**Splitting into major and minor arc cases.** Given $s \geq 1$, for any $a = 1, \ldots, s$ and $n \in \mathbb{N}$, choose $\theta_{n,a} \in \mathbb{C}$ with $|\theta_{n,a}| = 1$ such that

$$\left| \sum_{h \leq Hs \atop n + h \equiv a(s)} \beta(n + h)F(g(n, h)\mathbb{Z}) \right| = \theta_{n,a} \sum_{h \leq Hs \atop n + h \equiv a(s)} \beta(n + h)F(g(n, h)\mathbb{Z}).$$ \hfill (5.6)

We first split the interval $[0, Hs]$ into disjoint subintervals $I_1, \ldots, I_{W^2}$, each has length $W^{-2}Hs$. To apply the periodicity of $\gamma(n, h)\mathbb{Z}$, we further split each subinterval $I_t$ into progressions $n + h \equiv js + a(qs)$ for $j = 0, \ldots, q - 1$. For any given $j = 0, \ldots, q - 1$ and $t = 1, \ldots, W^2$, define $\mathcal{E}_{a,n,j,t} = \mathcal{E}(n, h')$, where $h'$ is the smallest integer in $I_t$ with $n + h' \equiv js + a(qs)$. Then by property (i), for any $h \in I_t$,

$$|\mathcal{E}_{a,n,j,t} - \mathcal{E}(n, h)| \leq W H^{-1} W^{-2} Hs = W^{-1}.$$ \hfill (5.7)

Choose $\gamma_{a,n,j,t} \in [0, 1]$ with $\gamma_{a,n,j,t}\mathbb{Z} = \gamma(n, h)\mathbb{Z}$ for any $h \in I_t$ satisfying $n + h \equiv js + a(qs)$. Note that $\gamma(n, h)$ is $qs$-periodic, so $\gamma_{a,n,j,t}$ is independent of $h$ with $n + h \equiv js + a(qs)$.

Define the function $F_{a,n,j,t} : G'/\Gamma' \to \mathbb{C}$ as $F_{a,n,j,t}(y\Gamma') = \theta_{n,a} F((\mathcal{E}_{a,n,j,t} + \gamma_{a,n,j,t} + y)\mathbb{Z})$ for any $y \in G'$. Note that $G'/\Gamma' = \mathbb{R}/\mathbb{Z}$ or \{0\}. It is easy to see that $F_{a,n,j,t}$ is well-defined. By formulas
(5.5) and (5.7),
\[ |F_{a,n,j,t}(g'(n,h)\Gamma') - \theta_{a,n}F(g(n,h)\mathbb{Z})| \leq W^{-1} \]
holds for all \( h \in I_t \) with \( n + h \equiv js + a(qs) \). Then by the above formula and formula (5.6),
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{h \leq Hs, n+h \equiv a(s)} \beta(n+h)F(g(n,h)\mathbb{Z}) = \sum_{a=1}^{s} \sum_{n=1}^{N} W^2 q^{-1} \sum_{t=1}^{q} \sum_{j=0}^{q-1} \sum_{h \in I_t, n+h \equiv js + a(qs)} \beta(n+h)F_{a,n,j,t}(g'(n,h)\Gamma') + O(W^{-1}HNs).
\]

We write \( F_{a,n,j,t} \) as \( \tilde{F}_{a,n,j,t} + E_{a,n,j,t} \), where \( E_{a,n,j,t} = \int_{G'/\Gamma'} F_{a,n,j,t} \) is a constant and \( \tilde{F}_{a,n,j,t} \) has zero average on \( G'/\Gamma' \). In the following, we shall apply different methods to estimate the major arc part
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} W^2 q^{-1} \sum_{t=1}^{q} \sum_{j=0}^{q-1} \sum_{h \in I_t, n+h \equiv js + a(qs)} \beta(n+h)E_{a,n,j,t},
\]
and the minor arc part
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} W^2 q^{-1} \sum_{t=1}^{q} \sum_{j=0}^{q-1} \sum_{h \in I_t, n+h \equiv js + a(qs)} \beta(n+h)\tilde{F}_{a,n,j,t}(g'(n,h)\Gamma').
\]

**The estimate of major arc case.** For the major arc estimate, we shall show the following result.

**Theorem 5.3.** Let \( s \geq 1 \) and \( H \geq 3 \) be integers. Let \( N \) be large enough with \( Hs \leq (\log N)^{1/32} \). Let \( \beta(n) \) be a multiplicative function with \( |\beta(n)| \leq 1 \) for all \( n \geq 1 \). Then
\[
(5.10) \quad \ll NS(\log \log (s+2))^{1/2} (\log \log H)^{3} (\log H)^{1/2} + NS(\log \log H)^{2} \left( \frac{M(\beta; N^{1/4}, N, \log N) + 1}{\exp\left( M(\beta; N^{1/4}, N, \log N) \right)} \right)^{1/2}.
\]

Note that formula (5.10) is bounded by
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} W^2 q^{-1} \sum_{t=1}^{q} \sum_{j=0}^{q-1} \sum_{h \in I_t, n+h \equiv js + a(qs)} \beta(n+h) = \sum_{c=1}^{qs} \sum_{n=1}^{N} W^2 \sum_{t=1}^{q} \sum_{h \in I_t, n+h \equiv c(qs)} \beta(n+h).
\]

To prove Theorem 5.3, we aim to bound
\[
\sum_{c=1}^{qs} \sum_{n=1}^{N} W^2 \sum_{t=1}^{q} \sum_{h \in I_t, n+h \equiv c(qs)} \beta(n+h).
\]
Let \( \beta_1 \) denote the completely multiplicative function defined by \( \beta_1(p) = \beta(p) \) for all prime numbers \( p \). We first bound the above in terms of \( \beta_1 \) rather than \( \beta \) by a standard technique (see, e.g., [MRT15, p2177-p2178]). The Dirichlet inverse of \( \beta_1 \) is \( \mu \beta_1 \). Then we can write \( \beta = \beta_1 * \alpha \), where \(*\) is the Dirichlet convolution and \( \alpha = \beta * \mu \beta_1 \). Observe that \( \alpha(n) \) is multiplicative and for all primes \( p, h(p) = 0 \) and \( |h(p^j)| \leq 2 \) for \( j \geq 2 \). So from the Euler product, we see that 
\[
\sum_{n=1}^{\infty} |\alpha(n)| n^{-\frac{3}{2}} = O(1).
\]
Then
\[
\begin{align*}
(5.14) &= \sum_{c=1}^{qs} \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \sum_{b \in N, ab \equiv c(qs)} \beta_1(b) \left| \sum_{c=1}^{qs} \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \sum_{b \in N, ab \equiv c(qs)} \beta_1(b) \right| \\
&= \sum_{c=1}^{qs} \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \sum_{b \in N, ab \equiv c(qs)} \beta_1(b) + O \left( \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \left( \sum_{b \in N, ab \equiv c(qs)} 1 \right) \right) \\
&= \sum_{c=1}^{qs} \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \sum_{b \in N, ab \equiv c(qs)} \beta_1(b) + O \left( W^{-1/4} NSW \right). 
\end{align*}
\]
Now the estimate of formula (5.10) is reduced to estimating
\[
\sum_{c=1}^{qs} \sum_{n=1}^{N} \sum_{t=1}^{W^2} |\alpha(a)| \sum_{b \in N, ab \equiv c(qs)} \beta_1(b). 
\]
We split the summation over \( b \) according to \( \text{gcd}(a, qs) \). Then
\[
(5.16) = \sum_{t=1}^{W^2} \sum_{e \mid qs} \sum_{a \leq W/e} |\alpha(a)| \sum_{1 \leq c \leq q/e} \sum_{e \mid c} \sum_{b \in \mathbb{N}, ab \equiv c(qs/e)} \beta_1(b) \\
= \sum_{t=1}^{W^2} \sum_{e \mid qs} \sum_{a \leq W/e} |\alpha(a)| \sum_{1 \leq c \leq q/e} \sum_{e \mid c} \sum_{b \in \mathbb{N}, ab \equiv c(qs/e)} \beta_1(b) + O \left( NSW^3q \right). 
\]
To give an upper bound for (5.17), we apply the second author's work on sums of multiplicative functions in short arithmetic progressions.
Lemma 5.4. [Wei18b, Theorem 1.7] Let $X$ be large enough with $1 \leq s \leq (\log X)^{1/32}$. Let $3 \leq H \leq X/s$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \geq 1$. Then

$$
\sum_{n=1}^{s} \sum_{x=X}^{2X} \left| \sum_{n=x}^{x+Hs} f(n) \right|^2 \ll H^2 X \varphi(s) \left( \frac{s}{\varphi(s)} \frac{\log \log H}{\log H} + \frac{1}{(\log X)^{1/300}} \frac{M_s(f; s; X; 2X) + 1}{\exp(M_s(f; s; X; 2X))} \right),
$$

where $\varphi$ is the Euler totient function.

By the above result and the Cauchy-Schwarz inequality, we have

Corollary 5.5. Let $s \geq 1$ and $H \geq 3$ be integers. Let $N$ be large enough with $Hs \leq (\log N)^{1/32}$. Let $\beta_1(n)$ be a completely multiplicative function with $|\beta_1(n)| \leq 1$ for all $n \geq 1$. Then

$$
\sum_{a=1}^{s} \sum_{x=1}^{N} \left| \sum_{n=x}^{x+Hs} \beta_1(n) \right| \ll HNs \left( \frac{s}{\varphi(s)} \frac{\log \log H}{\log H} \right)^{1/2} + HNs \left( \frac{M(\beta_1; N^{1/3}, N, s) + 1}{\exp(M(\beta_1; N^{1/3}, N, s))} \right)^{1/2}.
$$

According to the previous analysis and Corollary 5.5, we now prove Theorem 5.3.

Proof of Theorem 5.3. Note that $q \leq W$ and by formula (5.2), $W = \log H$. Applying Corollary 5.5 to formula (5.17), we have that for $N$ large enough with $Hs \leq (\log N)^{1/32}$,

$$
\sum_{t=1}^{W^2} \sum_{e \leq W} \sum_{a \leq W/e} \sum_{1 \leq c \leq qs/e} \sum_{n \leq N/ae} \left| \sum_{b \in \mathbb{N}} \left( \sum_{b \in \mathbb{N}} \beta_1(b) \right) \right|
\ll \sum_{t=1}^{W^2} \sum_{e \leq W} \sum_{a \leq W/e} \left( \frac{qs}{e} \frac{N W^{-2} Hs}{aq} \right) \left( \frac{qs/e}{\varphi(qs/e)} \frac{\log \log H}{\log H} \right)^{1/2} + \left( \frac{M(\beta_1; N^{1/4}, N, qs) + 1}{\exp(M(\beta_1; N^{1/4}, N, qs))} \right)^{1/2}
\ll NHs(\log W)^2 \left( \frac{\log \log(qs)}{\log H} \right)^{1/2} + \left( \frac{M(\beta_1; N^{1/4}, N, \log N) + 1}{\exp(M(\beta_1; N^{1/4}, N, \log N))} \right)^{1/2}
\ll NHs(\log^2(s + 2))^{1/2} \left( \frac{\log \log H}{(\log H)^{1/2}} \right) + NHs(\log \log H)^2 \left( \frac{M(\beta; N^{1/4}, N, \log N) + 1}{\exp(M(\beta; N^{1/4}, N, \log N))} \right)^{1/2}.
$$

By formulas (5.13), (5.15), (5.17) and (5.20), we obtain the conclusion stated in this theorem. □

The estimate of minor arc case. For the minor arc estimate, we shall prove the following result.
\textbf{Theorem 5.6.} Let \( s \geq 1, H \geq 3 \) and \( N \geq Hs \) be integers. Let \( \beta : \mathbb{N} \to \mathbb{C} \) be a multiplicative function with \( |\beta(n)| \leq 1 \) for each \( n \in \mathbb{N} \). Then we have

\[
(5.11) \quad \sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W} \sum_{j=0}^{q-1} \sum_{h \in I_t \atop n+h \equiv js+a(qs)} \beta(n+h) \tilde{F}_{a,n,j,t}(g'(n,h)\Gamma') \ll \frac{\log \log H}{\log H} N H s \frac{s}{\varphi(s)}.
\]

Let \( P_1, Q_1 \) be parameters to be determined later with \( W < P_1 < Q_1 < H \). Given integer \( s \geq 1 \). Define \( \mathcal{P} \) as the set of primes \( p \) in \([P_1, Q_1]\) which are coprime to \( s \). Note that here our setting of \( \mathcal{P} \) has a little difference with that in \([\text{HW19, Section 6}]\) by the requirement \((p,s) = 1\).

Let \( S = \{ n : 1 \leq n \leq N, \exists p \in \mathcal{P}, p \mid n \} \). Similar to the proof of \([\text{HW19, Lemma 6.2}]\), we have

\[
\sum_{h \leq Hs \atop n+h \in S} \beta(n+h) - \sum_{p \in \mathcal{P}} \sum_{l \in \mathbb{N}} \frac{1}{1 + |\{ q \in \mathcal{P} : q \mid l \}|} \ll \frac{H}{P_1}.
\]

Then by the above inequality,

\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W} \sum_{j=0}^{q-1} \sum_{h \in I_t \atop n+h \equiv js+a(qs)} 1_S(n+h) \beta(n+h) \tilde{F}_{a,n,j,t}(g'(n,h)\Gamma') \ll \frac{H}{P_1}.
\]

is approximated by

\[
\sum_{a=1}^{s} \sum_{n \in \mathcal{N}} \sum_{t=1}^{W} \sum_{j=0}^{q-1} \sum_{p \in \mathcal{P}} \sum_{l \in \mathbb{N}} \frac{1}{1 + |\{ q \in \mathcal{P} : q \mid l \}|} \frac{1}{pl \equiv js+a(qs)} \tilde{F}_{a,n,j,t}(g'(n,pl-n)\Gamma')
\]

up to an additive error \( O(P_1^{-1}HNs + W^{-B/2}HNs) \), where \( \mathcal{N} \) is the set defined in equation (5.4).

Let

\[
P_1 = 2^{r_-}, \quad Q_1 = 2^{r_+}, \quad r_- = \lceil 2 \log \log H \rceil + 1, \quad Q_1 = \lceil 1/10 \log H \rceil + 1.
\]

Then (5.23) becomes

\[
\sum_{r \in [r_-, r_+]} \sum_{a=1}^{s} \sum_{n \in \mathcal{N}} \sum_{t=1}^{W} \sum_{j=0}^{q-1} \sum_{p \in (2^{r-1}, 2^r)} \sum_{l \in \mathbb{N}} \frac{1}{1 + |\{ q \in \mathcal{P} : q \mid l \}|} \tilde{F}_{a,n,j,t}(g'(n,pl-n)\Gamma').
\]

Given integer \( r \in [r_-, r_+] \) and prime \( p \in (2^{r-1}, 2^r] \) with \((p,s) = 1\), by the choice of \( P \) and \( W \), we see that \((p,q) = 1\). Fix an integer \( \overline{p} \) with \( \overline{pp} \equiv 1(qs) \). Give \( a = 1, \ldots, s \) and \( j = 0, \ldots, q-1 \), let \( b_a \in [1, s] \) with \( \overline{pb_a} \equiv a(s) \). Write \( \overline{pb_a} = l_a s + a \) for some integer \( l_a \). Define \( J_{p,a,j} \) as the integer in \([0,q-1]\) with

\[
J_{p,a,j} \equiv (p(j-l_a)(mod \ q)).
\]
In the following, for simplicity, we use $j$ to denote $J_{p,a,j}$. Now the expression inside the summation over $r$ in formula (5.25) becomes

$$\sum_{p \in (2^{r-1}, 2^r)} \sum_{n \in \mathcal{N}} \sum_{t=1}^{s} \sum_{a=1}^{q-1} \sum_{j=0}^{s-1} \sum_{l \equiv \frac{t+n}{a} + a(qs)}^{l \leq \frac{N+H_s}{j} \mod s} \frac{1_{p \in n + I_l} \beta(p) \beta(l)}{1 + |\{q \in \mathcal{P} : q | l\}|} \tilde{F}_{a,n,j,t}(g'(n, pl - n) \Gamma')$$

(5.27)

$$= \sum_{p \in (2^{r-1}, 2^r)} \sum_{n \in \mathcal{N}} \sum_{t=1}^{s} \sum_{a=1}^{q-1} \sum_{j=0}^{s-1} \sum_{l \equiv \frac{t+n}{a} + a(qs)}^{l \leq \frac{N+H_s}{j} \mod s} \frac{1_{p \in n + I_l} \beta(p) \beta(l)}{1 + |\{q \in \mathcal{P} : q | l\}|} \tilde{F}_{pa,n,j,t}(g'(n, pl - n) \Gamma').$$

Here we view the index $pa$ appearing in $\tilde{F}$ as $pa \mod s$. By Cauchy-Schwarz inequality, we can bound equation (5.27) by

$$\leq \left( \frac{N}{2^{r-1}} W^2 \right)^{1/2} \left( \sum_{a=1}^{s} \sum_{j=0}^{q-1} \sum_{t=1}^{W^2} \sum_{l \equiv \frac{t+n}{a} + a(qs)}^{l \leq \frac{N+H_s}{j} \mod s} \left| \sum_{n \in \mathcal{N}} \sum_{p \in (2^{r-1}, 2^r)} \beta(p) \tilde{F}_{pa,n,j,t}(g'(n, pl - n) \Gamma') \right|^2 \right)^{1/2}$$

(5.28)

$$= \left( \frac{N}{2^{r-1}} W^2 \right)^{1/2} \cdot S^{1/2},$$

where

$$S := \sum_{a=1}^{s} \sum_{j=0}^{q-1} \sum_{t=1}^{W^2} \sum_{l \equiv \frac{t+n}{a} + a(qs)}^{l \leq \frac{N+H_s}{j} \mod s} \left| \sum_{n \in \mathcal{N}} \sum_{p \in (2^{r-1}, 2^r)} \beta(p) \tilde{F}_{pa,n,j,t}(g'(n, pl - n) \Gamma') \right|^2.$$

Now the proof of Theorem 5.6 is reduced to proving

$$S \ll 2^r W^{-(B_2+2)} H^2 N s^2,$$

(5.29)

where the parameter $B_2$ is given in formula (5.3).

Proof of Theorem 5.6. (Assume formula (5.29)) By formulas (5.25), (5.27), (5.28) and (5.29), we have

$$\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W^2} \sum_{j=0}^{q-1} \sum_{p \in \mathcal{P}} \sum_{l \equiv \frac{t+n}{a} + a(qs)}^{l \leq \frac{N+H_s}{j} \mod s} \frac{1_{p \in n + I_l} \beta(p) \beta(l)}{1 + |\{q \in \mathcal{P} : q | l\}|} \tilde{F}_{a,n,j,t}(g'(n, pl - n) \Gamma') \ll r_+ W^{-\frac{B_2}{2}} H N s.$$
Then by formula (5.23),
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W^2} \sum_{j=0}^{q-1} \sum_{h \in I_t} \sum_{n+h \equiv j a(qs)} 1_{S}(n+h) \beta(n+h) \widetilde{F}_{a,n,j,t}(g'(n,h)\Gamma') \ll (r_+ W^{-\frac{B_2}{2}} + P_1^{-1} + W^{-\frac{B}{2}}) H N s.
\]

We know that \( B = 100 C_0^2, \) \( B_2 = \frac{1}{10} C_0^{-2} B \geq 10 \) by formula (5.3), \( P_1 = 2^{[2 \log \log H] + 1}, \) \( Q_1 = 2^{[1/10 \log H] + 1} \) by (5.24), and \( W = \log H \) by (5.2). Note that \( B \geq 10. \) Then
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W^2} \sum_{j=0}^{q-1} \sum_{h \in I_t} \sum_{n+h \equiv j a(qs)} 1_{S}(n+h) \beta(n+h) \widetilde{F}_{a,n,j,t}(g'(n,h)\Gamma') \ll (\log H)^{-2} N H s. \tag{5.30}
\]

We are left with estimating the contribution from
\[
\sum_{a=1}^{s} \sum_{n=1}^{N} \sum_{t=1}^{W^2} \sum_{j=0}^{q-1} \sum_{h \in I_t} \sum_{n+h \equiv j a(qs)} \beta(n+h) \widetilde{F}_{a,n,j,t}(g'(n,h)\Gamma'). \tag{5.31}
\]

By the fundamental sieve, we have
\[
\sum_{n \leq X} 1 \ll X \frac{\log P_1}{\log Q_1} \sum_{P_1 \leq p \leq Q_1} (1 - 1/p)^{-1} + O(s Q_1^2).
\]

So
\[
(5.31) \leq \sum_{n=1}^{N} \sum_{m=1}^{n} \sum_{m \in S} \sum_{m=n-Hs}^{n+Hs} 1 \ll N H s \frac{\log P_1}{\log Q_1} \frac{s \log \log H}{\varphi(s) \log H}.
\]

\( \square \)

At the end, we focus on showing formula (5.29). Denote by \( A_{n,p,t} = \{ l \in \mathbb{N} : pl - n \in I_t \} \) and \( A_{n_1,n_2,p_1,p_2,t} = A_{n_1,p_1,t} \cap A_{n_2,p_2,t} \). It is not hard to check that \( |A_{n_1,n_2,p_1,p_2,t}| \ll 2^{-t} W^{-2} H s. \) And for given \( t \in [W^2], n_1 \in \mathcal{N}, p_1, p_2 \in (2^{r-1}, 2^r], \)
\[
|\{n_2 : n_2 \in \mathcal{N}, A_{n_1,n_2,p_1,p_2,t} \neq \emptyset\}| \leq W^{-2} H s. \tag{5.32}
\]

Expanding the square term in \( S \), we have that \( S \) is bounded by
\[
\left| \sum_{a=1}^{s} \sum_{n_1,n_2 \in \mathcal{N}} \sum_{(p_1,p_2,s) = 1} \sum_{j=0}^{q-1} \sum_{t=1}^{W^2} \sum_{l \in A_{n_1,n_2,p_1,p_2,t}} \widetilde{F}_{p_1 a,n_1,j,t}(g'(n_1, p_1 l - n_1)\Gamma') \widetilde{F}_{p_2 a,n_2,j,t}(g'(n_2, p_2 l - n_2)\Gamma') \right|, \tag{5.33}
\]
where $j_1 = j_{p_1,a,j}$ and $j_2 = j_{p_2,a,j}$ by equation (5.26). To prove formula (5.29), it is enough to prove

$$ (5.33) \leq 2^r W^{-(B_2+2)} H^2 N s^2. $$

(5.34)

From Appendix B, the proof of formula (5.34) can be reduced to proving the following proposition, which is a variant version of [HW19, Proposition 6.9] in arithmetic progressions that is adapted to our situation.

**Proposition 5.7.** Let $C_0$ be a large enough constant only depending on $d$, $B_2 \geq 10$, $B \geq 10C_0^2 B_2$, $H \geq \max(W^B, 2^{10r+})$. Given $s \geq 1$, $j = 0, \ldots, q - 1$ and $t = 1, \ldots, W^2$. For $r \in [r_-, r_+]$, $n_1 \in N$ and a prime number $p_1 \in (2^{r-1}, 2^r]$, the set $\Omega_{r,n_1,p_1,B_2}$ is defined to be the set of all triples:

$$(a, n, p) \in [1, s] \times N \times \{p : p \in (2^{r-1}, 2^r], (p, s) = 1\}.$$ such that

(i) $p$ is prime;

(ii) $|A_{n_1,n,p_1,p,t}| \geq 2^{-r} W^{-(B_2+2)} H s$;

(iii) $$ \left| \sum_{l(t) \in A_{n_1,n,p_1,p,t}, t \equiv js + a(qs)} \tilde{F}_{p_1,a,n_1,j_1,t}(g'(n_1, p_1 l - n_1) \Gamma') \tilde{F}_{p,a,n,j}(g'(n, p l - n) \Gamma') \right| \geq \frac{W^{-B_2}|A_{n_1,n,p_1,p,t}|}{qs}. $$

Then

$$ |\Omega_{r,n_1,p_1,B_2}| < 2^r W^{-B_2-2} H s^2. $$

(5.35)

By the previous analysis and formulas (5.29) and (5.34), to prove Theorem 5.6, it remains to proving Proposition 5.7. Next, we explain the main issue to prove Proposition 5.7. In the proof, we mainly apply the method of [HW19, Section 7], in which the authors deal with $g'(n, p l - n)$, a polynomial sequence of variables $n, p, l$ in general nilmanifolds. Specially, $g'(n, p l - n)$ is a polynomial of variables $p, l$ in $\mathbb{R}/\mathbb{Z}$ if $g'(n, h) = f(n + h)$ for some polynomial $f : \mathbb{Z} \to \mathbb{R}$. While for our case, it is not enough if we just treat $g'(n, p(s l + a) - n)$ as a polynomial of variables $p, l$. So we treat $g'(n, p(s l + a) - n)$ as a polynomial of variables $a, p, l$. This mainly because that $a$ may vary in a wide range $[1, s]$ compared with $[1, H]$. Combining with this idea and the proof in [HW19, Proposition 6.9], under the assumption that (5.35) does not hold, we shall obtain a conclusion contradicting the property that the sequence $\{g'(n, h)\}_{h \in A_{n}}$ is totally $W^{-B/2}$-equidistribution in $\mathbb{R}/\mathbb{Z}$ for almost all $a$ in $[1, s]$ (see property (ii) below formula (5.5)). For self-containing, we give the detailed proof of Proposition 5.7 in Appendix B.

Finally, we are ready to prove Theorem 5.1.

**Proof of Theorem 5.1.** By formulas (5.9), (5.10), (5.11) and Theorems 5.3, 5.6, we obtain the statement in this theorem. \qed
6. The Möbius disjointness of \( e(f(n)) \) with the \( k \)-th difference of \( f(n) \) tending to zero

In this section, we shall study the correlation between Möbius and exponential functions of arithmetic functions with the \( k \)-th differences tending to a constant (Theorems 1.12, 1.13, 1.16, and Proposition 1.17). We shall also prove Proposition 1.15, which states that the product of \( \mu(n) \) and any polynomial phase is disjoint from any finite products of translations of \( \mu^2 \). We first show the following property that arithmetic functions with \( k \)-th differences tending to zero can be approximated by certain concatenations of polynomials of degrees less than \( k \).

**Lemma 6.1.** Suppose that \( f(n) \) is a real-valued arithmetic function such that

\[
\lim_{n \to \infty} \| \Delta^k f(n) \| = 0,
\]

for some integer \( k \geq 1 \). Then for any integer \( N \geq 1 \), there is an increasing sequence \( \{N_i\}_{i=0}^\infty \) of natural numbers with \( N_0 = 0 \) and \( \lim_{i \to \infty} (N_{i+1} - N_i) = \infty \), and a sequence \( \{p_i(y)\}_{i=0}^\infty \) in \( \mathbb{R}[y] \) of degrees less than \( k \), such that

\[
\| f(n) - g_N(n) \| \leq 1/N, \text{ for any } n \in \mathbb{N}, \tag{6.1}
\]

where \( g_N \) is the concatenation of \( \{p_i(n)\}_{i=0}^\infty \) with respect to \( \{N_i\}_{i=0}^\infty \).

**Proof.** By Lemma 3.1, for each integer \( m \geq 1 \) and \( N \geq 1 \), there is a sufficiently large \( L_m \in \mathbb{N} \) with \( 2^m|L_m \) such that, whenever \( n \geq L_m \), we have

\[
\| f(n + j) - Y_n(n + j) \| \leq 1/N, \quad j = 0, 1, \ldots, 2^m - 1, \tag{6.2}
\]

where \( Y_n(n + j) \) is defined to be \( f(n + j) \) when \( 0 \leq j \leq k - 1 \) and the value determined by the following linear equations when \( k \leq j \leq m - 1 \),

\[
\Delta^k Y_n(n + j) = 0.
\]

It is not hard to check that

\[
Y_n(n + j) = \sum_{l=0}^{k-1} f(n + l) \prod_{t=0, t \neq l}^{k-1} \frac{j-t}{l-t}, \quad j = 0, 1, \ldots, 2^m - 1. \tag{6.3}
\]

We may further assume that the sequence \( \{L_m\}_{m=0}^\infty \) \((L_0 = 0)\) chosen above satisfies \( L_{m+1} > L_m \) and \( 2^m|L_{m+1} - L_m \) for each \( m \geq 1 \). Let \( d_m = (L_{m+1} - L_m)/2^m \). Then the following is a partition of \( \mathbb{N} \).

\[
\mathbb{N} = \bigcup_{m=0}^\infty \bigcup_{q=0}^{d_m-1} \{L_m + q2^m, L_m + q2^m + 1, \ldots, L_m + q2^m + 2^m - 1\}.
\]

Choose the sequence \( \{N_i\}_{i=0}^\infty \) with \( N_0 < N_1 < N_2 < \cdots \) such that

\[
\{N_0, N_1, \ldots\} = \{L_m + q2^m : m \in \mathbb{N}, \ 0 \leq q \leq d_m - 1\}.
\]

Then define

\[
p_i(n) = \sum_{l=0}^{k-1} f(N_i + l) \prod_{t=0, t \neq l}^{k-1} \frac{n - N_i - t}{l-t}. \tag{6.4}
\]
It is a polynomial of degree less than $k$. Let
\[ g_N(n) = p_i(n), \quad \text{when } N_i \leq n \leq N_{i+1} - 1. \]

By formula (6.2) and equation (6.3), we obtain formula (6.1).

**Lemma 6.2.** Let $\{N_i\}_{i=0}^\infty$ be an increasing sequence of natural numbers with $N_0 = 0$ and $\lim_{i \to \infty}(N_{i+1} - N_i) = \infty$. Let $\{p_i(y)\}_{i=0}^\infty$ be a sequence in $\mathbb{R}[y]$ with degrees less than $k$ for some positive integer $k$. Suppose that $f(n)$ is the concatenation of $\{p_i(n)\}_{i=0}^\infty$ with respect to $\{N_i\}_{i=0}^\infty$. Let $q$ be a positive integer and $0 \leq a \leq q - 1$. Then
\[ \lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N \atop n \equiv a \pmod{q}} \mu(n) e(f(n)) = 0 \] (6.5)

if and only if
\[ \lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \sum_{N_i \leq n \leq N_{i+1} - 1 \atop n \equiv a \pmod{q}} \mu(n) e(p_i(n)) = 0. \] (6.6)

**Proof.** It is obvious that (6.5) $\Rightarrow$ (6.6). We now show (6.6) $\Rightarrow$ (6.5). In this process, we need to use a classical result (see e.g., [Hua65, Chapter 6, Theorem 10]) stated as follows,
\[ \sup_{p(y) \in \mathbb{R}[y] \atop \deg(p(y)) < k} \left| \sum_{1 \leq n \leq N \atop n \equiv a \pmod{q}} \mu(n) e(p(n)) \right| \ll N / \log N, \]

where the implied constant at most depends on $q$. By the above inequality and equation (6.6), for any given $\epsilon > 0$, there is a positive integer $M$ such that whenever $m \geq M$ and $N \geq N_M$, we have
\[ \sum_{i=0}^{m-1} \sum_{N_i \leq n \leq N_{i+1} - 1 \atop n \equiv a \pmod{q}} \mu(n) e(p_i(n)) < (\epsilon/2) N_m, \] (6.7)

and
\[ \sup_{i \in \mathbb{N}} \left| \sum_{1 \leq n \leq N \atop n \equiv a \pmod{q}} \mu(n) e(p_i(n)) \right| < (\epsilon/4) N. \] (6.8)

Let $N \geq N_M$. Choose an appropriate $l \geq M$ with $N_l \leq N \leq N_{l+1} - 1$. Then
\[ \left| \sum_{n=1}^{N} \mu(n) e(f(n)) \atop n \equiv a \pmod{q} \right| \leq \left| \sum_{i=0}^{l-1} \sum_{N_i \leq n < N_{i+1} \atop n \equiv a \pmod{q}} \mu(n) e(p_i(n)) \right| + \left| \sum_{N_l \leq n \leq N \atop n \equiv a \pmod{q}} \mu(n) e(p_l(n)) \right| \]
\[ < (\epsilon/2) N_l + (\epsilon/2) N < \epsilon N. \]

Hence we obtain (6.5). $\square$
We now prove Proposition 1.17, which gives a sufficient condition of disjointness between the Möbius function and exponential functions of arithmetic functions with the $k$-th differences tending to 0.

\textit{Proof of Proposition 1.17.} By Lemma 6.1, we known that $e(f(n))$ can be approximated by $e(g_N(n))$ uniformly with respect to $n \in \mathbb{N}$, where $g_N(n)$ is a concatenation of certain polynomials. Then, under the assumption of equation (1.9), by Lemmas 4.2 and 6.2, we obtain

$$\lim \frac{1}{N} \sum_{1 \leq n \leq N, n \equiv a \pmod{q}} \mu(n)e(g_N(n)) = 0,$$

completing the proof. \hfill \square

Next, we shall prove Theorems 1.12, 1.13. They are both consequences of Proposition 6.6 below. Before proving them, we need some preparations. The following Dirichlet’s approximation theorem is classical and well-known (see e.g., [Tit86, Section 8.2]), which is proved via the fact that if there are $m + 1$ points contained in $m$ regions, then there must be at least two points lie in the same region.

\textbf{Lemma 6.3.} Given $L$ real numbers $\theta_1, \ldots, \theta_L$ and a positive integer $q$, then we can find an integer $t \in [1, q^L]$, and integers $a_1, \ldots, a_L$, such that $|\theta_j - a_j| \leq 1/q$, $j = 1, 2, \ldots, L$.

The following “asymptotical periodicity” of the concatenation of certain linear phases will be used in the proof of Proposition 6.6.

\textbf{Lemma 6.4.} Let $\{N_i\}_{i=0}^\infty$ be an increasing sequence of integers with $N_0 = 0$ and $\lim_{i \to \infty} (N_{i+1} - N_i) = \infty$. Suppose that $\alpha_0, \alpha_1, \ldots$ are real numbers such that the sequence $\{e(\alpha_i)\}_{i=0}^\infty$ has finitely many limit points. Assume that $f(n) = e(n\alpha_i)e(\beta_i)$ when $N_i \leq n < N_{i+1}$ for $i = 0, 1, 2, \ldots$, where $\beta_0, \beta_1, \ldots$ are real numbers. Then there is a $\delta$ with $0 < \delta < 1$ and a sequence $\{n_j\}_{j=0}^\infty$ of positive integers such that

$$\lim_{j \to \infty} \sum_{l=0}^{n_j} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |f(n + \ln_j) - f(n)|^2 = 0.$$

\textit{Proof.} Suppose that the limit points of $\{e(\alpha_i)\}_{i=0}^\infty$ are $e(\theta_1), \ldots, e(\theta_L)$ for some integer $L \geq 1$. Let $q_j = (j + 6)^2 \pi^2$, $j \geq 0$. By Lemma 6.3, we can find an integer $n_j$ with $1 \leq n_j \leq q_j^L$ such that $|\theta_s - \frac{a_{s,j}}{n_j}| \leq \frac{1}{n_jq_j}$, where $a_{s,j}$ is some integer for $s = 1, \ldots, L$. Moreover, there is an $i_0$ such that when $i \geq i_0$ we can choose an $s \in \{1, 2, \ldots, L\}$ satisfying $\|\alpha_i - \theta_s\| < \frac{1}{n_jq_j}$. Then for $i \geq i_0$, $|n_j\alpha_i| < \frac{2}{q_j}$ and $|e(n_j\alpha_i) - 1| = 2|\sin(\pi n_j\alpha_i)| \ll \frac{1}{q_j}$. So, for any given $n_j$,

$$\sum_{l=0}^{n_j} \limsup_{N \to \infty} \frac{1}{N} \left( \sum_{i=0}^{m-1} \sum_{n=N_i}^{N_{i+1}-1} |f(n + \ln_j) - f(n)|^2 + \sum_{n=N_m}^{N} |f(n + \ln_j) - f(n)|^2 \right)$$
The next result on the self-correlation of \( \mu(n) e(P(n)) \) in short arithmetic progressions on average can be deduced from Theorem 1.14. This result will be used in the proof of Proposition 6.6 later.

**Lemma 6.5.** Let \( d \geq 0 \), \( s \geq 1 \) and \( h \geq 3 \) be integers. Suppose that \( P(x) \in \mathbb{R}[x] \) of degree \( d \).

\[
\limsup_{N \to \infty} \frac{1}{N} \left| \sum_{n=1}^{N} \frac{1}{h} \sum_{l=1}^{h} \mu(n+ls) e(P(n+ls)) \right|^2 \ll (\log \log(s+2)) \frac{(\log \log h)^3}{(\log h)^{1/2}},
\]

the implied constant depending on \( d \) at most.

**Proof.** Given \( s \geq 1 \) and \( h \geq 3 \). For \( X \) large enough with \( (\log X)^{1/32} > hs \),

\[
\sum_{n=X}^{2X} \left| \sum_{l=1}^{h} \mu(n+ls) e(P(n+ls)) \right| = \sum_{a=1}^{s} \sum_{m=X/s}^{2X/s} \left| \sum_{n=a(m+1)s}^{(m+h+1)s} \mu(n+ls) e(P(n+ls)) \right| + O(hs)
\]

\[
= \sum_{a=1}^{s} \sum_{m=X/s}^{2X/s} \left| \sum_{n=a(m+1)s}^{(m+h+1)s} \mu(n+ls) e(P(n+ls)) \right| + O(hs)
\]

\[
= \sum_{a=1}^{s} \sum_{x=X/a}^{x+h} \sum_{n=a(m+1)s}^{n=x} \mu(n+ls) e(P(n+ls)) \right| + O(hs)
\]

\[
= \sum_{a=1}^{s} \sum_{x=X/a}^{x+h} \sum_{n=x}^{x+h} \mu(n+ls) e(P(n+ls)) \right| + O(hs)
\]

\[
= \frac{1}{s} \sum_{a=1}^{s} \sum_{x=x/a}^{x+h} \sum_{n=x}^{x+h} \mu(n+ls) e(P(n+ls)) \right| + O(X).
\]
By the above and Theorem 1.14,
\[
\sum_{n=X}^{2X} \left| \sum_{l=1}^{h} \mu(n + ls)e(P(n + ls)) \right|^2 \leq h \sum_{n=X}^{2X} \left| \sum_{l=1}^{h} \mu(n + ls)e(P(n + ls)) \right| \\
\leq Xh^2(\log \log s) \frac{(\log \log h)^3}{(\log h)^{1/2}}.
\]

Then
\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{h} \sum_{l=1}^{h} \mu(n + ls)e(P(n + ls)) \right|^2 \ll (\log \log s) \frac{(\log \log h)^3}{(\log h)^{1/2}},
\]
as claimed. \qed

Now applying Lemmas 6.4 and 6.5, we show the following proposition which concerns the disjointness of M"obius from concatenations of certain function phases.

**Proposition 6.6.** Let \( \{N_i\}_{i=0}^{\infty} \) be an increasing sequence of natural numbers with \( N_0 = 0 \) and \( \lim_{i \to \infty}(N_{i+1} - N_i) = \infty \). Suppose that \( \alpha_0, \alpha_1, \ldots \) are real numbers such that the sequence \( \{e(\alpha_i)\}_{i=0}^{\infty} \) has finitely many limit points. Suppose \( P(x) \in \mathbb{R}[x] \). Then
\[
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} \mu(n)e(P(n))e(n\alpha_i) \right| = 0.
\]

**Proof.** Choose \( \{\beta_i\}_{i=0}^{\infty} \) as a sequence of real numbers such that
\[
\left| \sum_{N_i \leq n < N_{i+1}} \mu(n)e(P(n))e(n\alpha_i) \right| = \sum_{N_i \leq n < N_{i+1}} \mu(n)e(P(n))e(n\alpha_i)e(\beta_i).
\]
Define \( f(n) \) to be \( e(n\alpha_i)e(\beta_i) \) when \( N_i \leq n < N_{i+1}, i = 0, 1, \ldots \). So it suffices to prove that
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)e(P(n))f(n) = 0.
\] 

(6.11)

By Lemma 6.4, there is a \( \delta \) with \( 0 < \delta < 1 \) and a sequence \( \{n_j\}_{j=0}^{\infty} \) of positive integers with \( \lim_{j \to \infty} n_j = \infty \) such that
\[
\lim_{j \to \infty} \sum_{l=1}^{n_j^\delta} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |f(n + ln_j) - f(n)|^2 = 0.
\] 

(6.12)

By Lemma 6.5, for any \( n_j \geq 3, \)
\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{n_j^\delta} \sum_{l=1}^{n_j^\delta} \mu(n + ln_j)e(P(n + ln_j)) \right|^2 \ll (\log \log n_j)(\log \log(n_j^\delta)) \frac{\log \log(n_j^\delta)}{\log(n_j^\delta)}.
\] 

(6.13)
Then the right side of the above inequality tends to zero as \( j \to \infty \). Hence for any given \( \epsilon > 0 \), by formulas (6.12) and (6.13), there are positive integers \( j_0 \) and \( M_0 \) such that whenever \( N > M_0 \),

\[
\frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \frac{1}{N} \sum_{n=1}^{N} |f(n + \ln j_0) - f(n)|^2 < \epsilon^2/9
\]

and

\[
\frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0)) \right|^2 \leq \epsilon^2/9.
\]

Then by the Cauchy-Schwarz inequality,

\[
\frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \frac{1}{N} \sum_{n=1}^{N} |f(n + \ln j_0) - f(n)| < \epsilon/3,
\]

and

\[
\frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0)) \right| < \epsilon/3.
\]

Observe that

\[
\frac{1}{N} \sum_{n=1}^{N} \mu(n)e(P(n))f(n) = \frac{1}{N} \sum_{n=1}^{N} \mu(n + \ln j_0)e(P(n + \ln j_0))f(n + \ln j_0) + O\left(\frac{\ln j_0}{N}\right).
\]

Then there is a positive integer \( M_1 \) such that whenever \( N > M_1 \),

\[
\left| \frac{1}{N} \sum_{n=1}^{N} \mu(n)e(P(n))f(n) - \frac{1}{N} \sum_{n=1}^{N} \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0))f(n + \ln j_0) \right| < \epsilon/3, \quad (6.16)
\]

Let \( M_2 = \max\{M_0, M_1\} \). Then by formulas (6.14), (6.15) and (6.16), for \( N > M_2 \), we have

\[
\left| \frac{1}{N} \sum_{n=1}^{N} \mu(n)e(P(n))f(n) \right| < \left| \frac{1}{N} \sum_{n=1}^{N} \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0))f(n + \ln j_0) \right| + \epsilon/3
\]

\[
\leq \left| \frac{1}{N} \sum_{n=1}^{N} \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0))(f(n + \ln j_0) - f(n)) \right|
\]

\[
+ \left| \frac{1}{N} \sum_{n=1}^{N} \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} \mu(n + \ln j_0)e(P(n + \ln j_0))f(n) \right| + \epsilon/3
\]

\[
\leq \frac{1}{N} \sum_{n=1}^{N} \frac{1}{n^j_{j_0}} \sum_{l=1}^{n^j_{j_0}} |f(n + \ln j_0) - f(n)|
\]
Then we obtain equation (6.11).

Actually, summarizing the proof of equation (6.11) in the above proposition, we conclude the following general result which concerns the disjointness of the product of \( \mu \) and any polynomial phase from any bounded arithmetic function which has “asymptotical periodicity”.

**Proposition 6.7.** Let \( f(n) \) be a bounded arithmetic function satisfying that there is a positive number \( \delta \) and a sequence \( \{ n_j \} \) of positive integers such that

\[
\lim_{j \to \infty} \frac{1}{n_j^\delta} \sum_{l=0}^{n_j^\delta} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |f(n + \ln_j) - f(n)|^2 = 0.
\]  

(6.17)

Suppose that \( P(x) \in \mathbb{R}[x] \). Then

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)e(P(n))f(n) = 0.
\]

As applications of Proposition 6.6, we now prove Theorems 1.12 and 1.13.

**Proof of Theorem 1.12.** Let \( f_0(n) = cn \) for \( n \in \mathbb{N} \). Then \( \Delta f_0(n) = c \) and \( \lim_{n \to \infty} \| \Delta (f - f_0)(n) \| = 0 \). Then for any given \( \epsilon > 0 \), by Lemma 6.1, there is an increasing sequence \( \{ N_i \} \) with \( N_0 = 0 \) and \( \lim_{i \to \infty} (N_{i+1} - N_i) = \infty \), and a function \( g_\epsilon(n) \) defined by \( g_\epsilon(n) = c_i \) when \( N_i \leq n < N_{i+1} \) for \( i = 0, 1, \ldots \), where \( c_0, c_1, \ldots \in \mathbb{R} \), such that

\[
\sup_{n \in \mathbb{N}} |e(f(n)) - e(cn + g_\epsilon(n))| < \epsilon.
\]  

(6.18)

By Proposition 6.6,

\[
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(cn + g_\epsilon(n)) = 0.
\]

By Lemma 6.2,

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(cn + g_\epsilon(n)) = 0.
\]

So by formula (6.18), we have, for \( N \) large enough,

\[
\frac{1}{N} \left| \sum_{1 \leq n \leq N} \mu(n)e(f(n)) \right| < 2\epsilon,
\]

as claimed. \( \square \)
Proof of Theorem 1.13. Let \( f_0(n) = \frac{c}{2}n^2 + c_1n + c_0 \) for \( n \in \mathbb{N} \). Then \( \Delta^2 f_0(n) = c \) and \( \lim_{n \to \infty} \| \Delta^2(f - f_0)(n) \| = 0 \). For any given \( \epsilon > 0 \), by Lemma 6.1 and equation (6.4), there is an increasing sequence \( \{ N_i \}_{i=0}^\infty \) with \( N_0 = 0 \) and \( \lim_{i \to \infty} (N_{i+1} - N_i) = \infty \), and a function \( g_i(n) \) defined by \( g_i(n) = n(f(N_i + 1) - f(N_i)) + (N_i + 1)f(N_i) - f(N_i + 1)N_i \) when \( N_i \leq n < N_{i+1} \) for \( i = 0, 1, \ldots \), such that

\[
\sup_{n \in \mathbb{N}} |e(f(n)) - e(f_0(n) + g_i(n))| < \epsilon. \tag{6.19}
\]

Since the set \( \{ e(f(N_i + 1) - f(N_i)) : i = 0, 1, \ldots \} \) has finitely many limit points,

\[
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(f_0(n))e(g_i(n)) = 0
\]

by Proposition 6.6. Using Lemma 6.2,

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f_0(n) + g_i(n)) = 0.
\]

So by formula (6.19), we have, for \( N \) large enough,

\[
\left| \frac{1}{N} \sum_{1 \leq n \leq N} \mu(n)e(f(n)) \right| < 2\epsilon
\]

as claimed. \( \square \)

In the remaining part of this section, we shall prove Theorem 1.16. The major ingredient of our proof is Matomäki-Radziwiłl-Tao-Teräväinen-Ziegler’s recent work [MRTTZ20] on averages of the correlation between multiplicative functions and polynomial phases in short intervals. To state this estimate, we first recall some notation appearing in Section 5. Recall that

\[
\mathbb{D}_1(g(n), n \mapsto n^it; X) := \left( \sum_{p \leq x} \frac{1 - \text{Re}(g(p)p^{-it})}{p} \right)^\frac{1}{2}
\]

the distance function used to measure the pretentiousness between \( n \mapsto n^it \) and \( g(n) \), a multiplicative function with \( |g(n)| \leq 1 \) for all \( n \in \mathbb{N} \). Denote by

\[
M(g; X, Q) := \inf_{\chi \equiv q \pmod{Q}} \mathbb{D}_1(g(n)\chi(n), n \mapsto n^it; X)^2 = \inf_{\chi \equiv q \pmod{Q}} \sum_{1 \leq \eta \leq Q} \sum_{p \leq X} \frac{1 - \text{Re}(g(p)\chi(p)p^{-it})}{p}
\]

Lemma 6.8. ([MRTTZ20, Theorems 1.3 and 1.8]) Let \( k \) be a given positive integer, and let \( 5/8 < \tau < 1 \) and \( 0 < \theta < 1 \) be fixed. Denote by \( \mathcal{D}_k \) the set of all polynomials in \( \mathbb{R}[y] \) of degrees less than \( k \). Let \( f : \mathbb{N} \to \mathbb{C} \) be a multiplicative function with \( |f(n)| \leq 1 \) for any \( n \in \mathbb{N} \). Suppose that \( X \geq 1 \), \( X^\theta \geq H \geq \exp((\log X)^\tau) \), and \( \eta > 0 \) are such that

\[
\int_X^{2X} \sup_{p(y) \in \mathcal{D}_k} \left| \sum_{x \leq n < x + H} f(n)e(p(n)) \right| \, dx \geq \eta HX.
\]
Then
\[ M(f; AX^k/H^{k-1/2}, Q) \ll_{k,\eta,\theta,\tau} 1 \]
for some positive numbers \( A, Q \ll_{k,\eta,\theta,\tau} 1 \).

By Proposition 5.2 about the “non-pretentious” of \( \mu \) and the above lemma, we have the following result which states that \( \mu(n) \) does not correlate with polynomial phases in short intervals on average.

**Lemma 6.9.** Let \( k \) be a given positive integer, and let \( 5/8 < \tau < 1 \) and \( 0 < \theta < 1 \) be fixed. Suppose that \( X \geq 1 \) and \( X^\theta \geq H \geq \exp((\log X)^\tau) \). Denote by \( D_k \) the set of all polynomials in \( \mathbb{R}[y] \) of degrees less than \( k \). Then for any \( \eta > 0 \), there is an \( X_0 > 0 \) (at most depends on \( k, \eta, \theta, \tau \)) such that whenever \( X > X_0 \),
\[
\int_X^{2X} \sup_{p(y) \in D_k} \left| \sum_{x \leq n < x + H} \mu(n)e(p(n)) \right| dx < \eta HX.
\]

Now applying Lemma 6.9, we show the following result.

**Theorem 6.10.** Let \( \tau \in (5/8, 1) \) be given. Let \( \{N_i\}_{i=0}^\infty \) be an increasing sequence of natural numbers with \( N_0 = 0 \) and \( N_{i+1} - N_i \geq \exp((\log i)^\tau) \) for \( i \) large enough, and let \( \{p_i(y)\}_{i=0}^\infty \) be a sequence in \( \mathbb{R}[y] \) of degrees less than \( k \) for some positive integer \( k \). Then
\[
\lim_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \left| \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_i(n)) \right| = 0. \tag{6.20}
\]

**Proof.** Given \( \epsilon \) sufficiently small with \( (1 - 2\epsilon)\tau > 5/8 \). Choose \( h_m = \exp((\log N_m)^{(1-2\epsilon)\tau}) \). By Lemma 6.9 and a dyadic subdivision, for \( N_m \) large enough,
\[
\sum_{x=0}^{\lfloor \frac{N_m}{h_m} \rfloor + 1} \sup_{p(y) \in D_k} \left| \sum_{x \leq n < x + h_m} \mu(n)e(p(n)) \right| \leq \epsilon N_m h_m, \tag{6.21}
\]
where \( D_k \) is the set of polynomials in \( \mathbb{R}[y] \) with degrees less than \( k \).

Let \( S_j = \{lh_m + j : l = 0, 1, \ldots, \lfloor \frac{N_m}{h_m} \rfloor \} \), for \( 0 \leq j \leq h_m - 1 \). Then, by (6.21), there is a \( j_0 \) such that
\[
\sum_{x \in S_{j_0}} \sup_{p(y) \in D_k} \left| \sum_{x \leq n < x + h_m} \mu(n)e(p(n)) \right| \leq \epsilon N_m. \tag{6.22}
\]

Suppose \( S_{j_0} \cap [N_i, N_{i+1}) = \{x_1^{(i)}, x_2^{(i)}, \ldots, x_l^{(i)}\} \), where \( x_1^{(i)} < x_2^{(i)} < \cdots < x_l^{(i)}, i = 0, 1, \ldots, m - 1 \). Then
\[
\left| \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_i(n)) \right| \leq \sum_{l=1}^{l-1} \sup_{p(y) \in D_k} \left| \sum_{x_t^{(i)} \leq n < x_{t+1}^{(i)} + h_m} \mu(n)e(p(n)) \right| + 2h_m.
\]
So
\[
\sum_{i=0}^{m-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_i(n)) \leq \sum_{x \in S_{2i}} \sup_{p(y) \in \mathcal{P}_k} \left| \sum_{n \leq x + h_m} \mu(n)e(p(n)) \right| + 2mh_m.
\]

By formula (6.22),
\[
\frac{1}{N_m} \sum_{i=0}^{m-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_i(n)) \leq \frac{\varepsilon N_m}{N_m} + \frac{2mh_m}{N_m}.
\]

Since \(N_{i+1} - N_i \geq \exp((\log i)^\tau)\) for \(i\) large enough, we have \(m \leq N_m/(\exp((\log N_m)^{(1-\epsilon)\tau}))\) for \(N_m\) sufficiently large. Inserting this into the above inequality and letting \(m \to \infty\), we obtain
\[
\limsup_{m \to \infty} \frac{1}{N_m} \sum_{i=0}^{m-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_i(n)) \leq \varepsilon.
\]

Letting \(\varepsilon \to 0\), we obtain equation (6.20). \(\square\)

Now we are ready to prove Theorem 1.16, which state that the Möbius disjointness of \(e(f(n))\) with the \(k\)-th difference of \(f(n)\) tending to zero as in formula (1.12).

**Proof of Theorem 1.16.** Firstly, by Proposition A.6, for \(j \geq k\), there are integers \(a_k, ..., a_j\) such that for each \(n \in \mathbb{N}\),
\[
\sum_{s=k}^{j} a_s \cdot \Delta^k f(n + s - k) = f(n + j) - \sum_{l=0}^{k-1} f(n + l) \prod_{t=0, t \neq l}^{k-1} \frac{j - t}{l - t}, \tag{6.23}
\]
where \(0 \leq a_s \leq j^{k-1}\) for \(s = k, ..., j\). Note that \(\sum_{l=0}^{k-1} f(n + l) \prod_{t=0, t \neq l}^{k-1} \frac{j - t}{l - t} = f(n + j)\) for \(j = 0, ..., k - 1\). Then by condition (1.12), for any \(j \in \mathbb{N}\),
\[
\left\| f(n + j) - \sum_{l=0}^{k-1} f(n + l) \prod_{t=0, t \neq l}^{k-1} \frac{j - t}{l - t} \right\| \leq \frac{Cj^k}{\exp((\log n)^\tau)}.
\]

For any given \(M \geq 1\), choose \(L_0 = 0\) and \(L_m = 2^m[\exp(MC_2^{2m}) + 1]\) for \(m = 1, 2, ...\). Then by the above inequality, we have for \(n \geq L_m\) and \(j = 0, 1, ..., 2^m - 1\),
\[
\left\| f(n + j) - \sum_{l=0}^{k-1} f(n + l) \prod_{t=0, t \neq l}^{k-1} \frac{j - t}{l - t} \right\| \leq \frac{1}{M}. \tag{6.24}
\]

Let \(d_m = (L_{m+1} - L_m)/2^m\). Setting \(0 = N_0 < N_1 < N_2 < \cdots\) with \(\{N_0, N_1, \ldots\}\) being the set of \(\{L_m + t2^m : m \in \mathbb{N}, 0 \leq t \leq d_m - 1\}\). Assume \(L_m = N_{k_m}\). Then \(k_{m+1} = k_m + d_m\). Note that \(k_0 = 0\). Then
\[
\exp(MC_2^{2m}) \ll k_m \ll m \exp(MC_2^{2m}).
\]
So we can find an appropriate $\epsilon > 0$ with $\tau - \epsilon > 5/8$ such that for $m$ large enough, $2^m \geq \exp((\log k_{m+1})^{\tau-\epsilon})$. By the choice of $N_i$, this leads to $N_{i+1} - N_i > \exp((\log i)^{\tau-\epsilon})$ for $i$ large enough. Define

$$p_M(n) = \sum_{t=0}^{k-1} f(N_i + l) \prod_{t=0, t \neq l}^{k-1} \frac{n - N_i - t}{l - t}$$

when $N_i \leq n < N_{i+1}$, $i = 0, 1, \ldots$. Hence by Theorem 6.10,

$$\lim_{s \to \infty} \frac{1}{N_s} \sum_{i=0}^{s-1} \sum_{N_i \leq n < N_{i+1}} \mu(n)e(p_M(n)) = 0,$$

and further by Lemma 6.2,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(n)e(p_M(n)) = 0.$$  \hfill (6.25)

Since $\sup_{n \in \mathbb{N}} \| f(n) - p_M(n) \| \leq \frac{1}{M}$ by equation (6.24), $\lim_{M \to \infty} \sup_{n \in \mathbb{N}} |e(f(n)) - e(p_M(n))| = 0$. Hence it follows from equation (6.25) that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(n)e(f(n)) = 0$$

as claimed. $\square$

At the end of this section, we prove Proposition 1.15. To prove it, by Proposition 6.7, it suffices to show that any finite products of translations of $\mu_r$ has the “asymptotical periodicity” (equation (6.17)). In the following, for simplicity, we write $Af(n) := f(n+1)$ as the translation of $f$ by adding 1 for $f$ an arithmetic function.

**Lemma 6.11.** Let $r \geq 2$ and $\mu_r(n) = 1$ if $n$ is $r$-th power-free and zero otherwise. For any $s \geq 1$ and $m_1, \ldots, m_s \in \mathbb{N}$, there is a strictly increasing sequence $\{n_j\}_{j=1}^\infty$ of positive integers satisfying that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left( \prod_{i=1}^{s} A^{m_i+l_{n_i}} \mu_r(n) - \prod_{i=1}^{s} A^{m_i} \mu_r(n) \right)^2$$

tends to 0 uniformly with respect to $l \in \mathbb{N}$ as $j \to \infty$.

**Proof.** We first show that there is a strictly increasing sequence $\{n_j\}_{j=1}^\infty$ of positive integers such that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu_r(n) - A^{ln_j} \mu_r(n)|^2$$

converges to 0 uniformly with respect to all $l \in \mathbb{N}$ as $j \to \infty$. \hfill (6.26)
Let \( p_j \) be the \( j \)-th prime, define the sequence \( \{n_j\}_{j=1}^{\infty} \) by \( n_j = p_1^r p_2^r \cdots p_j^r \). By [Mir49], for any positive integer \( m \), we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu_r(n) A^m \mu_r(n) = \prod_{p} \left( 1 - \frac{2}{p^r} \right) \prod_{p > p_j} \left( 1 + \frac{1}{p^r - 2} \right).
\]

So for any positive integer \( l \), we see that

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu_r(n) A^{n_j} \mu_r(n) \leq \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu_r(n) A^{n_j} \mu_r(n).
\]

Moreover, by expanding the square term,

\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu_r(n) - A^{n_j} \mu_r(n)|^2 \leq 2 \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu_r^2(n) - 2 \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu_r(n) A^{n_j} \mu_r(n)
\]

\[
= 2 \sum_{n=1}^{\infty} \frac{\mu(n)}{n^r} - 2 \prod_{p} \left( 1 - \frac{1}{p^r} \right) \prod_{p > p_j} \left( 1 + \frac{1}{p^r - 2} \right)^{-1}
\]

\[
= 2 \prod_{p} \left( 1 - \frac{1}{p^r} \right) \prod_{p > p_j} \left( 1 + \frac{1}{p^r - 2} \right)^{-1}
\]

tends to 0 when \( j \to \infty \). Hence we obtain the claim about (6.26).

Note that (6.26) is invariant under the translation \( A \), that is for any \( m \in \mathbb{N} \),

\[
\limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu_r(n) - A^{n_j} \mu_r(n)|^2 = \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} |A^m \mu_r(n) - A^{m+n_j} \mu_r(n)|^2.
\]

Then by the standard plus and minus a common term technique, for such a sequence \( \{n_j\}_{j=1}^{\infty} \), we obtain the claim in this lemma.

**Proof of Proposition 1.15.** By Lemma 6.11, it is not hard to check that \( \prod_{i=1}^{s} A^{m_i} \mu_r \) satisfies equation (6.17). Then it follows from Proposition 6.7 that

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n)e(p(n)) \mu_r(n + m_1) \cdots \mu_r(n + m_s) = 0
\]

as claimed.

**Appendix A. Some properties of the difference operator**

Recall that the difference operator \( \Delta \) is defined as \( (\Delta f)(n) = f(n+1) - f(n) \) for any arithmetic function \( f \). In this section, we give some basic properties of the operator \( \Delta^k \) for \( k \in \mathbb{N} \), which are used in this paper. The following one can be easily deduced by induction.
Proposition A.1. For any $k \in \mathbb{N}$ and any arithmetic function $f$, we have

$$\Delta^k f(n) = \sum_{l=0}^k (-1)^{k-l} \binom{k}{l} f(n + l). \quad (A.1)$$

It is known that $\Delta^k f(n) \equiv 0$ if and only if $f(n)$ is a polynomial with degree $k - 1$. Moreover, we have the following proposition, which is known as interpolation polynomial in the Lagrange form.

Proposition A.2. Given integers $J,k$ with $J > k \geq 0$. Suppose that $f(0), \ldots, f(J - 1)$ satisfy

$$\sum_{l=0}^k (-1)^{k-l} \binom{k}{l} f(n + l) = 0 \quad \text{for} \quad n = 0, \ldots, J - 1 - k.$$

Then

$$f(n) = \sum_{j=1}^{k-1} f(j) \prod_{0 \leq i \leq k, j \neq j} \frac{n - i}{j - i}$$

for $n = 0, \ldots, J - 1$.

Proof. Suppose $g(n) = \sum_{j=0}^{k-1} f(j) \prod_{0 \leq i \leq k, j \neq j} \frac{n - i}{j - i}$. Then $g(n)$ is a polynomial of degree at most $k - 1$ and satisfies $g(0) = f(0), \ldots, g(k - 1) = f(k - 1)$. By Proposition A.1,

$$\Delta^k g(n) = \sum_{l=0}^k (-1)^{k-l} \binom{k}{l} g(n + l) = 0, \quad n \geq 0.$$

For any given $x_0, \ldots, x_{k-1}$, the solution $(x_k, \ldots, x_{J-1})$ that satisfies $\sum_{l=0}^k (-1)^{k-l} \binom{k}{l} x_{n+l} = 0$ for $n = 0, \ldots, J - k - 1$ is unique. Then $(f(k), \ldots, f(J-1)) = (g(k), \ldots, g(J-1))$. Hence $f(n)$ is of the form as claimed in this proposition. \hfill \square

The following is a simple and useful fact that will be used in the later proofs in this paper.

Proposition A.3. For any $n \in \mathbb{N}$, $k \geq 1$ and $0 \leq j \leq k - 1$, $\prod_{0 \leq i \leq k-1, i \neq j} \frac{n - i}{j - i}$ is an integer.

Proof. For $0 \leq n \leq k - 1$ and $n \neq j$, $\prod_{0 \leq i \leq k-1, i \neq j} \frac{n - i}{j - i} = 0$; for $n = j$, $\prod_{0 \leq i \leq k-1, i \neq j} \frac{n - i}{j - i} = 1$; for $n \geq k$,

$$\prod_{0 \leq i \leq k-1, i \neq j} \frac{n - i}{j - i} = \prod_{0 \leq i \leq j} \frac{n - i}{j - i} \prod_{j+1 \leq i \leq k-1} \frac{n - i}{j - i}$$

$$= (-1)^{k-j-1} \frac{n!}{(n-k)!j!(k-1-j)!(n-j)}$$

$$= (-1)^{k-j-1} \binom{n-j}{n-k} \binom{n}{n-j}$$

is an integer. \hfill \square

The next one gives a variant version of Proposition A.2.
Lemma A.4. Given integers $J, k$ with $J > k \geq 0$. Suppose that $x_0, \ldots, x_{J-1} \in \mathbb{R}$. Then the following two statements are equivalent.

(i) For $n = 0, \ldots, J - 1 - k$,
\[
\left\{ \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} \{x_{n+l}\} \right\} = 0. \tag{A.2}
\]

(ii) For $n = 0, \ldots, J - 1$,
\[
\{x_n\} = \left\{ \sum_{j=0}^{k-1} \{x_j\} \prod_{0 \leq i < k-1 \atop i \neq j} \frac{n-i}{j-i} \right\}, \tag{A.3}
\]

where $\{\cdot\}$ denotes the fractional part function.

Proof. (i) $\Rightarrow$ (ii). We first show that there are integers $c_0, c_1, \ldots, c_{J-1}$ such that $x_0 + c_0, x_1 + c_1, \ldots, x_{J-1} + c_{J-1}$ satisfy the following linear equations
\[
\sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} (x_{n+l} + c_{n+l}) = 0, \quad n = 0, \ldots, J - 1 - k. \tag{A.4}
\]

Let $c_0 = \cdots = c_{k-1} = 0$. Letting $n = 0$ in equation (A.4), by equation (A.2) we have that the solution $c_k$ is an integer. Repeating the above process with $n = 1, \ldots, J - 1 - k$, we obtain successively solutions $c_{k+1}, \ldots, c_{J-1}$, which are all integers. By equation (A.4) and Proposition A.2,
\[
x_n + c_n = \sum_{j=0}^{k-1} \{x_j + c_j\} \prod_{0 \leq i < k-1 \atop i \neq j} \frac{n-i}{j-i}, \quad n = 0, \ldots, J - 1.
\]

By Proposition A.3, we have equation (A.3).

(ii) $\Rightarrow$ (i). Assume that $x_0, x_1, \ldots, x_{J-1}$ satisfy equation (A.3). Let
\[
g(n) = \sum_{j=0}^{k-1} \{x_j\} \prod_{0 \leq i < k-1 \atop i \neq j} \frac{n-i}{j-i}, \quad n = 0, \ldots, J - 1.
\]

Then $g(n)$ is a polynomial of degree at most $k-1$. By Proposition A.1,
\[
\Delta^k g(n) = \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} g(n + l) = 0, \quad n \geq 0.
\]

By (ii), $\{g(0)\} = \{x_0\}, \{g(1)\} = \{x_1\}, \ldots, \{g(J-1)\} = \{x_{J-1}\}$. Then
\[
\left\{ \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} \{x_{n+l}\} \right\} = \left\{ \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} \{g(n + l)\} \right\} = 0, \quad n = 0, \ldots, J - 1 - k.
\]

$\square$
The following gives an estimate of \( f(n) \) through the initial values and the upper bound of \( \Delta^k f(n) \).

**Proposition A.5.** Given \( n \in \mathbb{N} \) and a real number \( c > 0 \). Suppose integers \( J > k \geq 0 \). If \( f(n), f(n+1), \ldots, f(n+J-1) \) satisfy:

(a) \( |\Delta^k f(n+j)| \leq c, \ j = 0, 1, \ldots, J-1-k; \)

(b) \( f(n+j) \in [0,c], \ j = 0, 1, \ldots, k-1 \).

Then we have

\[
|f(n+j)| \leq (k+1)j^kc, \ j = k, k+1, \ldots, J-1.
\]

**Proof.** By (b), for \( j = 0, 1, \ldots, k-2 \),

\[
|\Delta f(n+j)| \leq c,
\]

and by induction,

\[
|\Delta^m f(n+j)| \leq 2^{m-1}c, \ j = 0, 1, \ldots, k-m-1, \ 1 \leq m \leq k-1.
\]

We first claim that, when \( 1 \leq m \leq k \) and \( k-m \leq j \leq J-1-m \),

\[
|\Delta^m f(n+j)| \leq \sum_{i=0}^{k-m-1} \left( \frac{j-(k-m)+i}{j-(k-m)} \right) 2^{m-1+i}c + \left( \frac{j}{k-m} \right) c,
\]

where \( \binom{0}{0} = 1 \). In the following we shall prove formula (A.6). When \( m = k \), the inequality (A.6) holds by (a). Assume inductively that formula (A.6) holds when \( m = m_0 + 1 \leq k \). We shall prove that formula (A.6) holds when \( m = m_0 \) and \( k-m_0 \leq j \leq J-1-m_0 \). For \( j = k-m_0 \), we have

\[
|\Delta^{m_0} f(n+k-m_0)| \leq |\Delta^{m_0+1} f(n+k-m_0-1)| + |\Delta^{m_0} f(n+k-m_0-1)|.
\]

Then by the inductive hypothesis on the \( m_0 + 1 \) case and formula (A.5),

\[
|\Delta^{m_0} f(n+k-m_0)| \leq 2^{m_0-1}c + \sum_{i=0}^{k-m_0-2} \left( \frac{k-m_0-1-(k-m_0-1)+i}{k-m_0-1-(k-m_0-1)} \right) 2^{m_0+i}c + \left( \frac{k-m_0}{k-m_0-1} \right) c
\]

\[
= \sum_{i=0}^{k-m_0-1} \left( \frac{k-m_0-(k-m_0)+i}{k-m_0-(k-m_0)} \right) 2^{m_0-1+i}c + \left( \frac{k-m_0}{k-m_0} \right) c.
\]

Note that in the above formula the coefficients before \( 2^{m_0+i}c \) and \( 2^{m_0-1+i}c \) are all 1. So formula (A.6) holds for \( m = m_0 \) and \( j = k-m_0 \). Now assume inductively that formula (A.6) holds for \( m = m_0 \) and some \( j = j_0 \geq k-m_0 \). When \( j = j_0 + 1 \), we have

\[
|\Delta^{m_0} f(n+j_0 + 1)| \leq |\Delta^{m_0} f(n+j_0)| + |\Delta^{m_0+1} f(n+j_0)|
\]

\[
\leq \sum_{i=0}^{k-m_0-1} \left( \frac{j_0-(k-m_0)+i}{j_0-(k-m_0)} \right) 2^{m_0-1+i}c + \left( \frac{j_0}{k-m_0} \right) c
\]
\[
+ \sum_{i=0}^{k-m_0-2} \left( j_0 - (k - m_0 - 1) + i \right) 2^{m_0+i}c + \left( \frac{j_0}{j_0 - (k - m_0 - 1)} \right) c \\
= \sum_{i=0}^{k-m_0-1} \left( j_0 - (k - m_0) + i \right) 2^{m_0+i}c + \left( \frac{j_0}{j_0 - (k - m_0)} \right) c \\
+ \sum_{i=1}^{k-m_0-1} \left( j_0 - (k - m_0) + i \right) 2^{m_0+i}c + \left( \frac{j_0}{j_0 - (k - m_0) + 1} \right) c \\
= \sum_{i=1}^{k-m_0-1} \left( j_0 - (k - m_0) + i \right) 2^{m_0+i}c + 2^{m_0}c + \left( \frac{j_0 + 1}{j_0 - (k - m_0)} \right) c \\
= \sum_{i=0}^{k-m_0-1} \left( j_0 + 1 - (k - m_0) + i \right) 2^{m_0+i}c + \left( \frac{j_0 + 1}{j_0 + 1 - (k - m_0)} \right) c.
\]

Then formula (A.6) holds by the above induction process. In particular, taking \( m = 1 \) in formula (A.6), for \( k - 1 \leq j \leq J - 2, \)

\[
|\triangle f(n + j)| \leq \sum_{i=0}^{k-2} \left( j - k + i + 1 \right) 2^i c + \left( \frac{j}{k-1} \right) c.
\]

Hence for \( k \leq j \leq J - 1, \) by the above inequality,

\[
|f(n + j)| \leq |f(n + k - 1)| + \sum_{l=k-1}^{j-1} |\triangle f(n + l)|
\]

\[
\leq c + \sum_{l=k-1}^{j-1} \left( \sum_{i=0}^{k-2} \left( \begin{array}{c} l - k + i + 1 \\ l - k + 1 \end{array} \right) 2^i \right) + \left( \frac{l}{k-1} \right) c
\]

\[
= c + \sum_{i=0}^{k-2} \left( j - 1 - k + i + 2 \right) 2^i + \left( \frac{j}{k} \right) c.
\]

Since \( \binom{j}{k} \leq j^k \) and \( \binom{j-k+i+1}{j-k} 2^i \leq j^{i+1} \) when \( 0 \leq i \leq k - 2, \)

\[
|f(n + j)| \leq (k + 1)j^{k}c, \ k \leq j \leq J - 1.
\]

This completes the proof. \( \square \)

The following proposition shows that \( f \) can be approached by polynomials if the \( k \)-th difference of \( f \) is small. This proposition will be used in the proof of Theorem 1.16 in Section 5.
Proposition A.6. Given \( j \geq k \geq 1 \). There are constants \( a_k, ..., a_j \) such that for each arithmetic function \( f \) and each \( n \in \mathbb{N} \),

\[
\sum_{l=k}^{j} a_l \cdot \Delta^k f(n + l - k) = f(n + j) - \sum_{m=0}^{k-1} f(n + m) \prod_{i=0, i \neq m}^{k-1} \frac{j - i}{m - i}. \tag{A.7}
\]

Proof. Let

\[
g(l) = \binom{j - l + k - 1}{k - 1} = \frac{(j - l + 1)(j - l + 2) \cdots (j - l + k - 1)}{(k - 1)!}
\]

be a polynomial of \( l \) with degree \( k - 1 \). Choose \( a_l = g(l) \) when \( l \leq j \leq j + k - 1 \). So \( a_l = 0 \) when \( l = j + 1, ..., j + k - 1 \). By equation (A.1), the left side of equation (A.7) is

\[
\sum_{l=k}^{j} a_l \sum_{t=0}^{k} (-1)^{k-t} \binom{k}{t} f(n + l - k + t). \tag{A.8}
\]

In the following, we consider the coefficients of \( f(n + m) \) in the above formula for \( m = 0, ..., j \). When \( m = j \), the coefficient of \( f(n + m) \) in (A.8) is \( a_j = 1 \). For the case \( j \geq 2k \), when \( k \leq m \leq j - k \), the coefficient of \( f(n + m) \) in (A.8) is

\[
\sum_{t=0}^{k} a_{m-t+k} (-1)^{k-t} \binom{k}{t} = (-1)^{k} \sum_{t=0}^{k} g(m + t)(-1)^{k-t} \binom{k}{t} = (-1)^{k} \Delta^k g(m) = 0.
\]

Notice that \( a_l = 0 \) for \( l = j + 1, ..., j + k - 1 \). When \( j - k < m \leq j - 1 \), the coefficient of \( f(n + m) \) in (A.8) is

\[
\sum_{t=m+k-j}^{k} a_{m-t+k} (-1)^{k-t} \binom{k}{t} = \sum_{t=0}^{k} a_{m-t+k} (-1)^{k-t} \binom{k}{t} = (-1)^{k} \Delta^k g(m) = 0. \tag{A.9}
\]

For the case \( j \leq 2k - 1 \), when \( k \leq m \leq j - 1 \), by a similar argument to equation (A.9), we have that the coefficient of \( f(n + m) \) in (A.8) is 0. Hence there are constants \( c_0, ..., c_{k-1} \) such that

\[
\sum_{l=k}^{j} a_l \cdot \Delta^k f(n + l - k) = f(n + j) - \sum_{m=0}^{k-1} c_m f(n + m) \tag{A.10}
\]

holds for each \( n \in \mathbb{N} \). To compute \( c_m \), let \( f_m \) be the polynomial of degree \( k - 1 \) such that \( f_m(m) = 1 \), \( f_m(t) = 0 \) for \( 0 \leq t \leq k - 1 \) and \( t \neq m \). So

\[
f_m(t) = \prod_{i=0, i \neq m}^{k-1} \frac{t - i}{m - i}.
\]
Note that $\Delta^k f_m = 0$ and equation (A.10) holds for any arithmetic function $f$. Let $f = f_m$ and $n = 0$ in equation (A.10), then

$$c_m = f_m(j) = \prod_{i=0, i \neq m}^{k-1} \frac{j - i}{m - i}.$$ 

Hence equation (A.7) holds with $a_l = g(l)$ for $l = k, \ldots, j$. □

**APPENDIX B. PROOF OF PROPOSITION 5.7**

In this section, we first show that the proof of formula (5.29) can be reduced to proving Proposition 5.7. Then we give a proof of Proposition 5.7. These results are used in Section 5.

**Proposition 5.7 (restated).** Let $C_0$ be a large enough constant only depending on $d$, $B_2 \geq 10$, $B \geq 10C_0^2B_2$, $H \geq \max(W^B, 2^{10r^2})$. Given $s \geq 1$, $j = 0, \ldots, q - 1$ and $t = 1, \ldots, W^2$. For $r \in [r_-, r_+]$, $n_1 \in N$ and a prime number $p_1 \in (2^{r-1}, 2^r)$, the set $\Omega_{r, n_1, p_1, B_2}$ is defined to be the set of all triples:

$$(a, n, p) \in [1, s] \times N \times \{p : p \in (2^{r-1}, 2^r), (p, s) = 1\}.$$

such that

(i) $p$ is prime;

(ii) $|A_{n_1, n, p_1, p, t}| \geq 2^{-r}W^{-(B_2+2)}Hs$;

(iii)

$$\left| \sum_{l \in A_{n_1, n, p_1, p, t}} \tilde{F}_{p_1, a_1, n, p_1, t}(g'(n_1, p_1 l - n_1)\Gamma')\tilde{F}_{p_1, a_1, n, p_1, t}(g'(n, p l - n)\Gamma') \right| \geq \frac{W^{-B_2}|A_{n_1, n, p_1, p, t}|}{q s}.$$ 

Then

$$|\Omega_{r, n_1, p_1, B_2}| < 2^{r}W^{-B_2-2}Hs^2.$$ 

We first show that Proposition 5.7 implies formula (5.29).

**Proof of inequality (5.29) (Assume Proposition 5.7).** We first decompose (5.33) into three parts $S_1 + S_2 + S_3$ defined below. Then we estimate each term.

By formula (5.32),

$$S_1 := \sum_{j=0}^{q-1} \sum_{l=1}^{W^2} \sum_{a=1}^{s} \sum_{p_1, p_2 \in (2^{r-1}, 2^r)} \sum_{n_1, n_2 \in N} \sum_{(p_1, p_2, a) = 1} |A_{n_1, n_2, p_1, p_2, t}| \sum_{l \in A_{n_1, n_2, p_1, p_2, t}} \tilde{F}_{p_1 a_1, n_1, p_1, t}(g'(n_1, p_1 l - n_1)\Gamma')\tilde{F}_{p_2 a_1, n_2, p_2, t}(g'(n_2, p_2 l - n_2)\Gamma')$$
\[ \ll sqW^{2} \sum_{n_{1} \in \mathcal{N}} \sum_{n_{2} \in \mathcal{N}} \sum_{(p_{1}, p_{2}, s) = 1} (|A_{n_{1}, n_{2}, p_{1}, p_{2}, t}|/qs) \]
\[ \ll 2^{r}W^{-(B_{2}+2)}H^{2}N_{s}^{2}. \]

By Proposition 5.7,
\[ S_{2} := \sum_{j=0}^{q-1} W^{2} \sum_{t=1}^{l} \sum_{n_{1} \in \mathcal{N}} \sum_{(p_{1}, s) = 1} \sum_{(a, n_{2}, p_{2}) \in \Omega_{r, n_{1}, p_{1}, B_{2}}} \sum_{|A_{n_{1}, n_{2}, p_{1}, p_{2}, t}| \geq 2^{-r}W^{-(B_{2}+2)}Hs} \]
\[ \ll qW^{2}(2^{-r}W^{-2}Hs/qs)N^{2}(2^{r}W^{-B_{2}-2}Hs^{2}) \ll 2^{r}NH^{2}s^{2}W^{-(B_{2}+2)}. \]

By the definition of \( \Omega_{r, n_{1}, p_{1}, B_{2}} \) and formula (5.32),
\[ S_{3} := \sum_{j=0}^{q-1} W^{2} \sum_{t=1}^{l} \sum_{n_{1} \in \mathcal{N}} \sum_{(p_{1}, s) = 1} \sum_{(a, n_{2}, p_{2}) \in \Omega_{r, n_{1}, p_{1}, B_{2}}} \sum_{|A_{n_{1}, n_{2}, p_{1}, p_{2}, t}| \geq 2^{-r}W^{-(B_{2}+2)}Hs} \]
\[ \ll \sum_{j=0}^{q-1} \sum_{t=1}^{l} s \sum_{n_{1} \in \mathcal{N}} \sum_{(p_{1}, p_{2}, s) = 1} \sum_{n_{2} \in \mathcal{N}} \sum_{|A_{n_{1}, n_{2}, p_{1}, p_{2}, t}| \neq 0} \frac{W-B_{2}|A_{n_{1}, n_{1}, p_{1}, p_{1}, t}|}{qs} \ll NH^{2}s^{2}2^{r}W^{-B_{2}-2}. \]

Then we have
\[ S \leq (5.33) = S_{1} + S_{2} + S_{3} \ll 2^{r}NH^{2}s^{2}W^{-(B_{2}+2)}. \]
as claimed in formula (5.29). \( \square \)

We now prove Proposition 5.7.

**Proof of Proposition 5.7.** For given \( j, t, p_{1}, n_{1} \), assume on the contrary that
\[ |\Omega_{r, n_{1}, p_{1}, B_{2}}| \geq 2^{r}W^{-B_{2}-2}Hs^{2}. \]

Write \( \{ l : l \in A_{n_{1}, p_{1}, l \equiv js+a(qs)} \} \) as \( \{ qs \ell + c_{0} + a : \ell \in [L] \} \) for some \( c_{0} \) with \( s|c_{0} \) (independent of \( a, n, p \)), where \( L \leq 4 \cdot 2^{-r}W^{-3}H \). Let \( (a, n, p) \in \Omega_{r, n_{1}, p_{1}, B_{2}} \), by conditions (ii) and (iii), we
have
\[
\sum_{l \in \mathcal{A}_{n,p}'} \overline{F}_{p_1a,n_1,j_1,l}(g'(n_1, p_1(qsl + c_0 + a) - n_1)\Gamma') \overline{F}_{p_2a,n_2,j_2,l}(g'(n, p(qsl + c_0 + a) - n)\Gamma') \geq W^{-B_2} |\mathcal{A}_{n,p}'|,
\]
where \( \mathcal{A}_{n,p}' \) is a subinterval of \([L]\) with length \(|\frac{\mathcal{A}_{n_1,p_1,q_1}}{qs}\)| or \(|\frac{\mathcal{A}_{n_2,p_2,q_2}}{qs}\)| + 1. In the following, we apply the method used in [HW19, Section 7] to our situation with appropriate modifications.

Denote by \( g_{a,n_1,p_1}(u) = g'(n_1, qsu + p_1(c_0 + a) - n_1) \) and \( g_{a,n,p}(u) = g'(n, qsu + p(c_0 + a) - n) \). By formula (B.1), the sequence \( \{(g_{a,n_1,p_1}(p_1l)\Gamma', g_{a,n,p}(pl)\Gamma')\}_{l \in \mathcal{A}_{n,p}'} \) is not \( 2^{-2}W^{-B_2} \)-equidistributed in \( G'/\Gamma' \times G'/\Gamma' \). By [HW19, Lemma 2.10], we can find a short length \( L_{n,p}' \geq 2^{-5}W^{-B_2}L \) such that the sequence \( \{(g_{a,n_1,p_1}(p_1l)\Gamma', g_{a,n,p}(pl)\Gamma')\}_{l \in [L_{n,p}']} \) is not \( 2^{-5}W^{-2B_2} \)-equidistributed in \( G'/\Gamma' \times G'/\Gamma' \). Then by Lemma 2.8, there is a character \( \eta_{a,n,p} \in \mathbb{Z}^2 \) of \( G'/\Gamma' \times G'/\Gamma' \) with \( 0 < |\eta_{a,n,p}| < W^{O(B_2)} \) such that
\[
\|\eta_{a,n,p} \circ (g_{a,n_1,p_1}(p_1l)\Gamma', g_{a,n,p}(pl)\Gamma')\|_{C^\infty[L_{n,p}']} \ll W^{O(B_2)}.
\]
As \( L_{n,p}' \gg W^{-B_2}L \), this implies that
\[
\|\eta_{a,n,p} \circ (g_{a,n_1,p_1}(p_1l)\Gamma', g_{a,n,p}(pl)\Gamma')\|_{C^\infty[L]} \ll W^{O(B_2)}.
\]
By the pigeonhole principle, we can find a non-zero additive character \( \eta \in \mathbb{Z}^2 \) of \( G'/\Gamma' \times G'/\Gamma' \) such that for a set \( \Omega_{r,n_1,p_1,B_2}^* \subseteq \Omega_{r,n_1,p_1,B_2} \) with \( |\Omega_{r,n_1,p_1,B_2}^*| \geq 2^rW^{-O(B_2)}Hs^2 \), formula (B.3) holds and \( \eta_{a,n,p} = \eta \).

Let \( P_{r,n_1,p_1} = \{p : 2^{r-1} \leq p \leq 2^r, (p, s) = 1\} \), there are at least \( W^{-O(B_2)}Hs^2 \) choices of \((a,n)\) such that \((a,n,p) \in \Omega_{r,n_1,p_1,B_2}^* \). Since
\[
\Omega_{r,n_1,p_1,B_2}^* \leq |P_{r,n_1,p_1}| \cdot s \cdot |\{n : n \in \mathcal{N}, A_{n_1,n_1,p_1,t} \neq \emptyset\}| + 2^{r-1}W^{-O(B_2)}Hs^2,
\]
we have
\[
|P_{r,n_1,p_1}| \gg 2^rW^{-O(B_2)}.
\]
Furthermore, for any \( p \in P_{r,n_1,p_1} \), there are at least \( W^{-O(B_2)}s \) numbers of \( a \in [s] \) such that \((a,n,p) \in \Omega_{r,n_1,p_1,B_2}^* \) for some \( n \). Let \( g'(n, h) = \sum_{i=0}^{d'} \alpha_i(n + h)^i \). Write \( \eta = \eta_1 \oplus \eta_2 \). Suppose
\[
\eta_1 \circ (g_{a,n_1,p_1}(p_1l)\Gamma') = \sum_{i_1,i_2 \geq 0} \gamma_{i_1,i_2} a^{i_1} l_{i_2}.
\]
and
\[
\eta_2 \circ (g_{a,n,p}(u)\Gamma') = \sum_{i_1,i_2,i'_1 \geq 0} \beta_{i_1,i_2,i'_1} p^{i_1} a^{i_2} l_{i'_1}.
\]
Then
\[
\eta \circ (g_{a,n_1,p_1}(p_1l)\Gamma', g_{a,n,p}(pl)\Gamma') = \sum_{i_1,i_2 \geq 0} \gamma_{i_1,i_2} a^{i_1} l_{i_2} + \sum_{i=0}^{d'} \sum_{w=0}^{d'-i} \sum_{w'=0}^w \beta_{w,i,w} p^{i+w} l_{i+w'}.
\]
By the previous argument, for any \( p \in \mathcal{P}_{r,n_1,p_1} \), there is a set \( \mathcal{C}_{r,n_1,p_1} \subseteq [s] \) with
\[
|\mathcal{C}_{r,n_1,p_1}| \gg W^{-O(B_2)}s,
\]
such that the following holds
\[
\| \eta \circ (g_{a,n_1,p_1}(p_1 l), g_{a,n_2,p}(p l)) \|_{C^\infty [L]} \ll W^{O(B_2)},
\]
where we treat formula (B.6) as a polynomial of \( l \). Then by Lemma 2.6, there is a positive integer \( Z_1 \ll O(1) \) such that for all \( 1 \leq i \leq d' \),
\[
\left\| Z_1 \left( \sum_{w'=0}^{d'-i} \gamma_{w',i} a^{w'} + \sum_{w'=0}^{d'-i} \left( \sum_{w' \geq w'} \beta_{w,i,w'} p^{w+i} \right) a^{w'} \right) \right\|_{\mathbb{R}/\mathbb{Z}} \ll W^{O(B)} L^{-i} \ll 2^{i} W^{O(B_2)} H^{-i}. \tag{B.8}
\]
Using the pigehole principle, one can make \( Z_1 \) independent of \( a \) after substituting \( \mathcal{C}_{r,n_1,p_1} \) with a smaller subset whose cardinality still satisfies formula (B.7). We now view
\[
Z_1 \left( \sum_{w'=0}^{d'-i} \gamma_{w',i} a^{w'} + \sum_{w'=0}^{d'-i} \left( \sum_{w' \geq w'} \beta_{w,i,w'} p^{w+i} \right) a^{w'} \right)
\]
as a polynomial of \( a \). Then applying Lemma 2.7 (with \( \epsilon = 2^{i} W^{O(B_2)} H^{-i} \) and \( \delta = W^{-O(B_2)} \)) to formula (B.8), there is a positive integer \( Z_2 \) with \( Z_2 \ll W^{O(B_2)} \) such that for any \( i = 1, \ldots, d' \),
\[
\left\| Z_2 Z_1 \left( \sum_{w'=0}^{d'-i} \gamma_{w',i} a^{w'} + \sum_{w'=0}^{d'-i} \left( \sum_{w' \geq w'} \beta_{w,i,w'} p^{w+i} \right) a^{w'} \right) \right\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{i} W^{O(B_2)} H^{-i}. \tag{B.9}
\]
By Lemma 2.6 again, there is a positive integer \( Z_3 \) with \( Z_3 \ll O(1) \) such that for \( i = 1, \ldots, d' \) and \( 1 \leq w' \leq d' - i \),
\[
\left\| Z_3 Z_2 Z_1 \left( \sum_{u \geq w} \beta_{w,i,w} p^{w+i} \right) \right\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{i} W^{O(B_2)} H^{-i} s^{-w'}. \tag{B.10}
\]
By formulas (B.8) and (B.9), and the triangle inequality, we have that for \( i = 1, \ldots, d' \) and \( w' = 0 \),
\[
\left\| Z_3 Z_2 Z_1 \left( \sum_{u \geq 0} \beta_{w,i,0} p^{w+i} \right) \right\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{i} W^{O(B_2)} H^{-i}. \tag{B.10}
\]
Applying Lemmas 2.7 and 2.6 again to the following polynomials
\[
Z_3 Z_2 Z_1 \left( \gamma_{w',i} + \sum_{u \geq w'} \beta_{w,i,u} p^{w+i} \right)
\]
of \( p \in [2^r] \), then there are positive integers \( Z_4, Z_5 \) with \( Z_4 \ll W^{O(B_2)} \) and \( Z_5 \ll O(1) \), such that for any \( i = 1, \ldots, d' \), \( 0 \leq w' \leq d' - i \) and \( w' \leq w \leq d' - i \),
\[
\left\| Z_5 Z_4 Z_3 Z_2 Z_1 \beta_{w,i,u} \right\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{-w} W^{O(B_2)} H^{-i} s^{-w'}. \tag{B.11}
\]
Let $Z = Z_5Z_4Z_3Z_2Z_1$. Then $Z \ll W^{O(B_2)}$ and the character $Z\eta_2$ satisfies

$$|Z\eta_2| \ll |Z| |\eta| \ll W^{O(B_2)}.$$  \hfill (B.12)

Now we choose a sufficiently large positive number $C_0 = O(1)$ which serves as the implicit constant in the exponent $O(B_2)$ appearing in the above. We write formula (B.11) as

$$\|Z\beta_{l_1,l_2,l_1'}\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{-l_1}rW^{C_0B_2}H^{-l_2}s^{-l_1'},$$

which holds for any $l_1 \geq 0, l_2 \geq 1, l_1 + l_2 \leq d'$ and $l_1' \leq l_1$.

By formula (5.32), the number of pairs $(n, p)$ with $(n, p) \in \Omega^*_{r,n_1,p_1,B_2}$ is at most $2^{r}Hs$. Since $|\Omega^*_{r,n_1,p_1,B_2}| \geq 2^{r}W^{-C_0B_2}Hs^{2}$, there is a pair $(n, p) \in \mathcal{N} \times \{ p : p \in (2^{-1},2^{r}], (p, s) = 1 \}$ such that at least $W^{-C_0B_2}s$ choices of $a \in [s]$ have the property $(a, n, p) \in \Omega^*_{r,n_1,p_1,B_2}$. For such $(a, n, p)$, similar to the process of [HW19, Lemma 7.1], we fix any interval in form of $\mathcal{U}_a' = \{ u \in \mathbb{Z} : qsu + p(c_0 + a) - n \in [Hs] \}$ with length $\left\lceil \frac{2^{r}}{q} \right\rceil$. Note that for $u \in \mathcal{U}_a', |u| \leq 2H/q$. For any $u_1, u_2 \in \mathcal{U}_a'$, by formulas (B.5) and (B.13), we have the following estimate.

$$\|Z\eta_2 \circ g'(n, qsu_1 + p(c_0 + a) - n)\Gamma' - Z\eta_2 \circ g'(n, qsu_2 + p(c_0 + a) - n)\Gamma'\|_{\mathbb{R}/\mathbb{Z}}$$

$$= \|Z \sum_{l_1 \geq 0, l_2 \geq 1, l_1' \geq 0 \atop l_1 + l_2 \leq d', l_1' \leq l_1} \beta_{l_1,l_2,l_1'} p^{l_1}d'^{l_1'}(u_1 - u_2) \sum_{h=0}^{l_2-1} u_1^h u_2^{l_2-1-h} \|_{\mathbb{R}/\mathbb{Z}}$$

$$\ll \sum_{l_1 \geq 0, l_2 \geq 1, l_1' \geq 0 \atop l_1 + l_2 \leq d', l_1' \leq l_1} W^{-C_0B_2-3}q^{-l_2} \leq C_0W^{-C_0B_2}.$$  \hfill (B.14)

Set $F(x) = e(Z\eta_2(x))$, a Lipschitz function from $G'/\Gamma'$ to $\mathbb{C}$ with $\|F\|_{Lip} \ll W^{C_0B_2}$ by formula (B.12). Let $\mathcal{B}_a' = \{ qsu + p(c_0 + a) - n : u \in \mathcal{U}_a' \}$. Then by formula (B.14),

$$|E_{h \in \mathcal{B}_a'} F(g'(n, h)\Gamma')| > 1 - C_0W^{-C_0B_2} \geq 1/2.$$  

Note that $\mathcal{B}_a'$ is an arithmetic progression in $\mathcal{B}_a = \{ h \in [Hs] : n + h \equiv pa(s) \}$ with length greater than $W^{-2C_0B_2-4}H$. By the above, it follows that the sequence $\{ g'(n, h)\Gamma' \}_{h \in \mathcal{B}_a}$ is not totally $\min(W^{-2C_0B_2-4} \frac{1}{2}W^{-C_0B_2})$-equidistributed in $G'/\Gamma'$. By the assumption in the proposition, $\min(W^{-2C_0B_2-4} \frac{1}{2}W^{-C_0B_2}) \geq W^{-C_0B_2-4}$. So there is an $n \in \mathcal{N}$ such that for at least $W^{-C_0B_2}s$ choices of $a \in [s]$, $\{ g'(n, h)\Gamma' \}_{h \in \mathcal{B}_a}$ is not totally $W^{-C_0B_2}$-equidistributed in $G'/\Gamma'$. This contradicts the construction of $\mathcal{N}$ in formula (5.4). We complete the proof.  \hfill \square

References

[BGS13] A. Balog, A. Granville and K. Soundararajan, *Multiplicative functions in arithmetic progressions*, Annales mathématiques du Québec (1) 27, 3-30, 2013.

[Cho65] S. Chowla, *The Riemann Hypothesis and Hilbert’s tenth problem*, Mathematics and Its Applications, vol. 4 (Gordon and Breach Science Publishers, New York, NY, 1965).

[Dav37] H. Davenport, *On some infinite series involving arithmetical functions II*, Quart. J. Math. 8, 313-350, 1937.
[Wei18a] F. Wei, Angie entropy and arithmetic compactification of natural numbers, arXiv:1811.11000v3, 2018.
[Wei18b] F. Wei, Disjointness of Möbius from asymptotically periodic functions, arXiv:1810.07360v6, 2018.
[Wei16] F. Wei, Entropy of arithmetic functions and Sarnak’s Möbius disjointness conjecture, Ph.D. Thesis, The University of Chinese Academy of Sciences, 2016.
[Zhan91] T. Zhan. On the representation of large odd integer as a sum of three almost equal primes. Acta Math. Sinica (3) 7, 259-272, 1991.

Department of Mathematics, University of New Hampshire, Durham, NH 03824, USA –and– Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

Email address: guweichen14@mails.ucas.ac.cn

Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China

Email address: weif@mail.tsinghua.edu.cn