Additive functions in short intervals, gaps and a conjecture of Erdős

Alexander P. Mangerel[1,2]

Abstract

With the aim of treating the local behaviour of additive functions, we develop analogues of the Matomäki–Radziwiłł theorem that allow us to approximate the average of a general additive function over a typical short interval in terms of a corresponding long average. As part of this treatment, we use a variant of the Matomäki–Radziwiłł theorem for divisor-bounded multiplicative functions recently proven in Mangerel (Divisor-bounded multiplicative functions in short intervals. arXiv: 2108.11401). We consider two sets of applications of these methods. Our first application shows that for an additive function $g:\mathbb{N} \to \mathbb{C}$ any non-trivial savings in the size of the average gap $|g(n) - g(n-1)|$ implies that g must have a small first centred moment i.e. the discrepancy of $g(n)$ from its mean is small on average. We also obtain a variant of such a result for the second moment of the gaps. This complements results of Elliott and of Hildebrand. As a second application, we make partial progress on an old question of Erdős relating to characterizing constant multiples of $\log n$ as the only *almost everywhere* increasing additive functions. We show that if an additive function is almost everywhere non-decreasing then it is almost everywhere *well approximated* by a constant times a logarithm. We also show that if the set $\{n \in \mathbb{N} : g(n) < g(n-1)\}$ is sufficiently sparse, and if g is not extremely large too often on the primes (in a precise sense), then g is identically equal to a constant times a logarithm.

✉ Alexander P. Mangerel
smangerel@gmail.com

1   Centre de Recherches Mathématiques, Université de Montréal, 2920 Chemin de la Tour, Montréal, QC H3T 1J4, Canada

2   Department of Mathematical Sciences, Durham University, Upper Mountjoy Campus, Stockton Road, Durham DH1 3LE, UK

1 Introduction

An arithmetic function $g : \mathbb{N} \to \mathbb{C}$ is called *additive* if, whenever $n, m \in \mathbb{N}$ are coprime, $g(nm) = g(n) + g(m)$; it is said to be *completely additive* if the coprimality condition on n, m can be ignored. Additive functions are objects of classical study in analytic and probabilistic number theory, their study being enriched by a close relationship with the probabilistic theory of random walks.

Much is understood about the *global* behaviour of general additive functions. For instance, the orders of magnitude of all of the centred moments

$$\frac{1}{X} \sum_{n \leq X} \left| g(n) - \frac{1}{X} \sum_{n \leq X} g(n) \right|^k, \quad k > 0,$$

have been computed by Hildebrand [1]. When $k = 2$, the slightly weaker but generally sharp Turán–Kubilius inequality (see Lemma 3.2) gives an upper bound, uniform in g, of the form

$$\frac{1}{X} \sum_{n \leq X} \left| g(n) - \frac{1}{X} \sum_{n \leq X} g(n) \right|^2 \ll B_g(X)^2, \tag{1}$$

where we have denoted by $B_g(X)^2$ the approximate variance defined via

$$B_g(X) := \left(\sum_{p^k \leq X} \frac{|g(p^k)|^2}{p^k} \right)^{1/2}.$$

When g is real-valued one can determine necessary and sufficient conditions according to which the distribution functions $F_X(z) := \frac{1}{X} |\{n \leq X : g(n) \leq z\}|$ converge to a distribution function F as $X \to \infty$; this is the content of the Erdős–Wintner theorem [2]. Under certain conditions the corresponding distribution functions (with suitable normalizations) converge to a Gaussian, a fundamental result of Erdős and Kac [3].

Much less is understood regarding the *local* behaviour of additive functions i.e. the simultaneous behaviour of g at neighbouring integers. Questions of interest from this perspective include

  (i) the distribution of $\{g(n)\}_n$ in *typical* short intervals $[x, x + H]$, where $x \in [X, 2X]$ and $H = H(X)$ grows slowly,
  (ii) the distribution of the sequence of gaps $|g(n) - g(n - 1)|$ between consecutive values and
  (iii) the distribution of tuples $(g(n + 1), \ldots, g(n + k))$, for $k \geq 2$.

Pervasive within this scope are questions surrounding the characterization of those additive functions g whose local behaviour is rigid in some sense, such questions are discussed in Sect. 1.2.

The purpose of this paper is to consider questions of a local nature about general additive functions.

1.1 Matomäki–Radziwiłł type theorems for additive functions

The study of additive functions is intimately connected with that of multiplicative functions i.e. arithmetic functions $f : \mathbb{N} \to \mathbb{C}$ such that $f(nm) = f(n)f(m)$ whenever $(n, m) = 1$. The mean-value theory of bounded multiplicative functions, which provides tools for the analysis of the global behaviour of multiplicative functions, was developed in the '60s and '70s in the seminal works of Wirsing [4] and Halász [5]. In contrast, the study of the local behaviour of multiplicative functions has long been the source of intractable problems. An important example of this is Chowla's conjecture [6]. This conjecture states, among other things, that for any $k \geq 2$ and any tuple $\epsilon \in \{-1, +1\}^k$, the set

$$\{n \leq X : \lambda(n + 1) = \epsilon_1, \ldots, \lambda(n + k) = \epsilon_k\}$$

has $(2^{-k} + o(1))X$ elements, where λ is the Liouville[1] function. In other terms, the sequence of tuples $(\lambda(n + 1), \ldots, \lambda(n + k))$ equidistributes among the tuples of signs in $\{-1, +1\}^k$. The depth of this conjecture is revealed upon observing that when $k = 1$, this corresponds to the statement that $\lambda(n)$ takes value $+1$ and -1 with asymptotically equal probability $1/2$. This was shown by Landau [7] to be equivalent to the prime number theorem.

Problems of this type have recently garnered significant interest, thanks to the celebrated theorems of Matomäki and Radziwiłł [8]. Broadly speaking, their results show that averages of a bounded multiplicative function in typical short intervals are well approximated by a corresponding long average. In a strong sense, this suggests that the local behaviour of many multiplicative functions is determined by their global behaviour. The simplest version of their theorems to state is as follows.

Theorem (Matomäki–Radziwiłł [8]) *Let* $f : \mathbb{N} \to [-1, 1]$ *be multiplicative. Let* $10 \leq h \leq X/100$. *Then*

$$\frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} f(m) - \frac{2}{X} \sum_{X/2 < m \leq X} f(m) \right|^2 \ll \frac{\log \log h}{\log h} + (\log X)^{-1/50}.$$

This result, its natural extensions to complex-valued functions [9], and further improvements, extensions and variants (e.g. [10]) have had profound impacts not only in analytic number theory, but equally in combinatorics and dynamics. For instance, Tao [11] used this result to develop technology in order to obtain estimates for the logarithmically-averaged binary correlation sums

---

[1] The Liouville function is the multiplicative function defined as $\lambda(n) := (-1)^{\Omega(n)}$, where $\Omega(n)$ is the number of prime factors of n, counted with multiplicity.

$$\frac{1}{\log X} \sum_{n \leq X} \frac{f(n) f(n+h)}{n}, \text{ for multiplicative functions } f : \mathbb{N} \to \mathbb{C}, |f(n)| \leq 1.$$

This was essential in his proof of the Erdős discrepancy problem [12], and also enabled him to obtain a logarithmic density analogue of the case $k = 2$ of Chowla's conjecture. It has also been pivotal in the various developments towards Sarnak's conjecture on the disjointness of the Liouville function from zero entropy dynamical systems (see [13] for a survey).

Our first main result establishes an ℓ^1-averaged comparison theorem for short and long averages of additive functions, inspired by the theorem of Matomäki and Radziwiłł.

Theorem 1.1 *Let* $g : \mathbb{N} \to \mathbb{C}$ *be an additive function. Let* $10 \leq h \leq X/100$ *be an integer.[2] Then*

$$\frac{2}{X} \sum_{\frac{X}{2} < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} g(m) - \frac{2}{X} \sum_{\frac{X}{2} < m \leq X} g(m) \right|$$

$$\ll \left(\sqrt{\frac{\log \log h}{\log h}} + (\log X)^{-\frac{1}{800}} \right) B_g(X).$$

Remark 1.2 Theorem 1.1 should be compared to the "trivial bound" arising from applying the triangle inequality, the Cauchy–Schwarz inequality and (1) (which is valid for dyadic long averages as well) to obtain

$$\frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} g(m) - \frac{2}{X} \sum_{X/2 < m \leq X} g(m) \right|$$

$$\leq \frac{2}{X} \sum_{X/2-h < m \leq X} \left| g(m) - \frac{2}{X} \sum_{X/2 < n \leq X} g(n) \right|$$

$$\ll \left(\frac{1}{X} \sum_{m \leq X} \left| g(m) - \frac{2}{X} \sum_{X/2 < n \leq X} g(n) \right|^2 \right)^{\frac{1}{2}}$$

$$\ll B_g(X).$$

In contrast, Theorem 1.1 gives the non-trivial bound $o(B_g(X))$ whenever $h = h(X) \to \infty$ as $X \to \infty$.

To get a more precise additive function analogue of the Matomäki–Radziwiłł theorem, one would hope to obtain a mean square (or ℓ^2) version of Theorem 1.1. We

---

[2] The requirement that h be an integer is possibly unnecessary, but assuming it allows us to avoid certain pathologies associated with functions g taking very large values.

are limited in this matter by the possibility of very large values of g. Specifically, if $|g(p)|/B_g(X)$ can get very large for many primes $p \leq X$, it is possible for the ℓ^2 average to be dominated by a sparse set (i.e. the multiples of these p), wherein the discrepancy between the long and short sums is not small. We will thus work with a specific collection of additive functions in order to preclude such pathological behaviour.

To describe this collection we introduce the following notations. Given $\varepsilon > 0$ and an additive function g, we define[3]

$$F_g(\varepsilon) := \limsup_{X \to \infty} \frac{1}{B_g(X)^2} \sum_{\substack{p \leq X \\ |g(p)| > \varepsilon^{-1} B_g(X)}} \frac{|g(p)|^2}{p}.$$

Roughly speaking, $F_g(\varepsilon)$ measures the contribution to $B_g(X)^2$ from prime values $g(p)$ of very large absolute value.

Clearly, $0 \leq F_g(\varepsilon) \leq 1$ for all $\varepsilon > 0$ and additive functions g. We will concern ourselves with functions g such that $F_g(\varepsilon) \to 0$ as $\varepsilon \to 0^+$, a condition that is satisfied by many additive functions. When g is bounded on the primes e.g. when $g(n) = \Omega(n)$, the number of prime factors of n counted with multiplicity, it is clear that $F_g(\varepsilon) = 0$ whenever ε is sufficiently small. For a different example, taking $g = c \log$ for some $c \in \mathbb{C}$ we find $B_g(X) \sim \frac{|c|}{\sqrt{2}} \log X$, so that $|g(p)| \leq (\sqrt{2} + o(1)) B_g(X)$ for all primes p and hence $F_g(\varepsilon) = 0$ for all $\varepsilon < 1/2$, say.

Definition 1.3 We define the collection $\mathcal{A}$ to be the set of those additive functions $g : \mathbb{N} \to \mathbb{C}$ such that

(a) $B_g(X) \to \infty$, and
(b) $B_g(X)$ is dominated by the prime values $|g(p)|$, in the sense that

$$\limsup_{X \to \infty} \frac{1}{B_g(X)^2} \sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|g(p^k)|^2}{p^k} = 0.$$

We shall see below (see Lemma 3.6a)) that $\mathcal{A}$ contains all completely additive and all strongly additive[4] functions g with $B_g(X) \to \infty$. Within $\mathcal{A}$ we define

$$\mathcal{A}_s := \{g \in \mathcal{A} : \lim_{\varepsilon \to 0^+} F_g(\varepsilon) = 0\}. \tag{2}$$

---

[3] By Chebyshev's inequality, $\sum_{\substack{p \leq X \\ |g(p)| > \varepsilon^{-1} B_g(X)}} p^{-1} \ll \varepsilon^2$, and thus the proportion of integers divisible by a prime p with $|g(p)| > \varepsilon^{-1} B_g(X)$ is sparse, namely of size $O(\varepsilon^2 X)$. Nevertheless, if $F_g(\varepsilon) \gg 1$ for all $\varepsilon > 0$ the values $g(n)^2$ at multiples of such primes can have an outsized influence on the second moment.

[4] By a *strongly additive* function we mean an additive function g such that $g(p^k) = g(p)$ for all primes p and all $k \geq 1$.

Thus, among other examples, $\Omega(n)$, $\omega(n) := \sum_{p|n} 1$ and, for any $c \in \mathbb{C}$, $c \log$ all belong to $\mathcal{A}_s$. We show in general that whenever $g \in \mathcal{A}_s$, we may obtain an ℓ^2 analogue of Theorem 1.1.

Theorem 1.4 *Let* $g : \mathbb{N} \to \mathbb{C}$ *be an additive function in* $\mathcal{A}_s$. *Let* $10 \leq h \leq X/100$ *be an integer with* $h = h(X) \to \infty$. *Then*

$$\frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} g(m) - \frac{2}{X} \sum_{X/2 < m \leq X} g(m) \right|^2 = o(B_g(X)^2).$$

Our proof of Theorem 1.4 relies on a variant of the Matomäki–Radziwiłł theorem that applies to a large collection of *divisor-bounded* multiplicative functions, proven in the recent paper [14]. See Theorem 5.3 for a statement relevant to the current circumstances.

Remark 1.5 The rate of decay in this result depends implicitly on the rate at which $F_g(\varepsilon) \to 0$ as $\varepsilon \to 0^+$, and on the size of the contribution to $B_g(X)$ from the prime power values of g. We have therefore chosen to state the theorem in this qualitative form for the sake of simplicity.

It deserves mention that the application of the Matomäki–Radziwiłł method, which will be used in this paper, to the study of specific additive functions is not entirely new. Goudout [15, 16] applied this technique to derive distributional information about $\omega(n)$ in typical short intervals; for example, he proved in [15] that the Erdős–Kac theorem holds in short intervals $(x - h, x]$ for almost all $x \in [X/2, X]$, as long as $h = h(X) \to \infty$. The specific novelty of Theorems 1.1 and 1.4 lie in their generality, and it is this aspect which will be used in the applications to follow.

1.2 Applications: gaps and rigidity problems for additive functions

Given $c \in \mathbb{C}$, the arithmetic function $n \mapsto c \log n$ is completely additive. In contrast to a typical additive function g, whose values $g(n)$ depend on the prime factorization of n which might vary wildly from one integer to the next, $c \log$ varies slowly and smoothly, with very small gaps

$$c \log(n + 1) - c \log n = O(1/n) \quad \text{for all } n \in \mathbb{N}.$$

In the seminal paper [17], Erdős studied various characterization problems for real- and complex-valued additive functions relating to their local behaviour, and in so doing found several characterizations of the logarithm as an additive function. Among a number of results, he showed that if either

(a) $g(n + 1) \geq g(n)$ for all $n \in \mathbb{N}$, or
(b) $g(n + 1) - g(n) = o(1)$ as $n \to \infty$,

then there exists $c \in \mathbb{R}$ such that $g(n) = c \log n$ for all $n \geq 1$.

Moreover, Erdős and later authors posited that these hypotheses could be relaxed. Kátai [18] and independently Wirsing [19] weakened assumption (b), and proved the above result under the averaged assumption

$$\lim_{X \to \infty} \frac{1}{X} \sum_{n \le X} |g(n+1) - g(n)| = 0.$$

Hildebrand [20] showed the stronger conjecture of Erdős that if $g(n_k+1) - g(n_k) \to 0$ on a set $\{n_k\}_k$ of density[5] 1 then $g = c \log$; this, of course, is an *almost sure* version of (b).

In a different direction, Wirsing [21] showed that for completely additive functions g, (b) may be weakened to $g(n+1) - g(n) = o(\log n)$ as $n \to \infty$, and this is best possible.

A number of these results were strengthened and generalized by Elliott [22, Ch. 11], in particular to handle functions g with small gaps $|g(an+b) - g(An+B)|$, for independent linear forms $n \mapsto an+b$ and $n \mapsto An+B$ (i.e. such that $aB - Ab \neq 0$). Characterization problems of these kinds for both additive and multiplicative functions have continued to garner interest more recently. In [23], Klurman proved a long-standing conjecture of Kátai, showing that if a unimodular multiplicative function $f : \mathbb{N} \to S^1$ has gaps satisfying $|f(n+1) - f(n)| \to 0$ on average then there is a $t \in \mathbb{R}$ such that $f(n) = n^{it}$ for all n. In a later work, Klurman and the author [24] proved a conjecture of Chudakov from the '50s characterizing completely multiplicative functions having uniformly bounded partial sums. See Kátai's survey paper [25] for numerous prior works in this direction for both additive and multiplicative functions.

While these multiplicative results have consequences for additive functions, they are typically limited by the fact that if g is a real-valued additive function then the multiplicative function $e^{2\pi i g}$ is only sensitive to the values $g(n) \pmod 1$. In particular, considerations about e.g. the monotone behaviour of g cannot be directly addressed by appealing to corresponding results for multiplicative functions.

1.2.1 Erdős' conjecture for almost everywhere monotone additive functions

One still open problem stated in [17] concerns the *almost sure* variant of problem (a) above. For convenience, given an additive function $g : \mathbb{N} \to \mathbb{R}$ we set $g(0) := 0$ and define the set of decrease of g:

$$\mathcal{B} := \{n \in \mathbb{N} : g(n) < g(n-1)\}, \qquad \mathcal{B}(X) := \mathcal{B} \cap [1, X].$$

Conjecture 1.6 [17] *Let* $g : \mathbb{N} \to \mathbb{R}$ *be an additive function, such that*

$$|\mathcal{B}(X)| = o(X) \text{ as } X \to \infty. \tag{3}$$

Then there exists $c \in \mathbb{R}$ *such that* $g(n) = c \log n$ *for all* $n \in \mathbb{N}$.

---

[5] Given a sequence $\mathcal{C} \subset \mathbb{N}$, the *(natural) density* of $\mathcal{C}$, if it exists, is the limit $d\mathcal{C} := \lim_{X \to \infty} \frac{|\mathcal{C} \cap [1,X]|}{X}$.

Thus, if g is non-decreasing except on a set of integers of natural density 0 then it is conjectured that g must be a constant times a logarithm.

Condition (3) is necessary, as for any $\varepsilon > 0$ one can construct a function g, not a constant multiple of $\log n$, which is monotone except on a set of density at most ε. Indeed, picking a prime $p_0 > 1/\varepsilon$ and defining $g = g_{p_0}$ to be the completely additive function defined at primes by

$$g_{p_0}(p) := \begin{cases} \log p & : \ p \neq p_0 \\ p_0 & : \ p = p_0, \end{cases}$$

one finds that $g_{p_0}(n) = \log n$ if and only if $p_0 \nmid n$, and that $\mathcal{B} = \{mp_0 + 1 : m \in \mathbb{N}\}$. It is easily checked that the density $d\mathcal{B}$ of $\mathcal{B}$ satisfies $0 < d\mathcal{B} = 1/p_0 < \varepsilon$.

As a consequence of our results on short interval averages of additive functions, we will prove the following partial result towards Erdős' conjecture.

Corollary 1.7 *Let* $g : \mathbb{N} \to \mathbb{R}$ *be a completely additive function that satisfies*

$$\lim_{\varepsilon \to 0^+} F_g(\varepsilon) = \lim_{\varepsilon \to 0^+} \limsup_{X \to \infty} \frac{1}{B_g(X)^2} \sum_{\substack{p \leq X \\ |g(p)| > \varepsilon^{-1} B_g(X)}} \frac{g(p)^2}{p} = 0. \tag{4}$$

Assume furthermore that there is a $\delta > 0$ *such that*

$$|\mathcal{B}(X)| \ll X/(\log X)^{2+\delta}.$$

Then there is a constant $c \in \mathbb{R}$ *such that* $g(n) = c \log n$ *for all* $n \in \mathbb{N}$.

The above corollary reflects the fact that the main difficulties involved in fully resolving Conjecture 1.6 are

(i) the possible lack of sparseness of $\mathcal{B}$ beyond $|\mathcal{B}(X)| = o(X)$, and
(ii) the possibility of very large values $|g(p)|$.

More generally, we show that any function $g \in \mathcal{A}_s$ that satisfies $|\mathcal{B}(X)| = o(X)$ is *close* to a constant multiple of a logarithm at prime powers.

Theorem 1.8 *Let* $g : \mathbb{N} \to \mathbb{R}$ *be an additive function belonging to* $\mathcal{A}_s$, *and suppose* $|\mathcal{B}(X)| = o(X)$. *Let* $X \geq 10$ *be large. Then there is* $\lambda = \lambda(X)$ *with* $|\lambda(X)| \ll B_g(X)/\log X$ *such that*

$$\sum_{p^k \leq X} \frac{|g(p^k) - \lambda(X) \log p^k|^2}{p^k} = o\left(\sum_{p^k \leq X} \frac{g(p^k)^2}{p^k} \right) \quad \text{as } X \to \infty.$$

Moreover, λ *is slowly varying as a function of* X *in the sense that for every fixed* $0 < u \leq 1$,

$$\lambda(X^u) = \lambda(X) + o\left(\frac{B_g(X)}{\log X} \right).$$

Finally, using a result of Elliott [26], we will prove the following *approximate* version of Erdős' conjecture under weaker conditions than in Corollary 1.7.

Theorem 1.9 *Let $g : \mathbb{N} \to \mathbb{R}$ be an additive function, such that $|\mathcal{B}(X)| = o(X)$. Then there are parameters $\lambda = \lambda(X)$ and $\eta = \eta(X)$ such that for all but $o(X)$ integers $n \le X$,*

$$g(n) = \lambda \log n - \eta + o(B_g(X)). \tag{5}$$

The functions λ, η are slowly varying in the sense that for any $u \in (0, 1)$ fixed,

$$\lambda(X^u) = \lambda(X) + o\left(\frac{B_g(X)}{\log X}\right), \qquad \eta(X^u) = \eta(X) + o(B_g(X)).$$

Remark 1.10 Note that if we knew (5) held for all three of $n, m, nm \in [1, X]$ then we could deduce that

$$\lambda \log(nm) - 2\eta + o(B_g(X)) = g(n) + g(m) = g(nm)$$
$$= \lambda \log(nm) - \eta + o(B_g(X)),$$

and thus that $\eta = o(B_g(X))$. As such, (5) would be valid with $\eta \equiv 0$. Unfortunately, we are not able to confirm this unconditionally.

1.2.2 On Elliott's property of gaps

Gap statistics provide an important example of local properties of a sequence. Obviously, an additive function g whose values $g(n)$ are globally close to g's mean value must have small gaps $|g(n+1) - g(n)|$. Conversely, it was observed by Elliott that the growth of the gaps between consecutive values of g also control the typical discrepancy of $g(n)$ from its mean.
More precisely, given an additive function $g : \mathbb{N} \to \mathbb{C}$ and $X \ge 2$, define

$$A_g(X) := \sum_{p^k \le X} \frac{g(p^k)}{p^k}\left(1 - \frac{1}{p}\right). \tag{6}$$

It is well known (see e.g. Lemma 3.1) that as $X \to \infty$, $A_g(X)$ is the asymptotic mean value of $\{g(n)\}_{n \le X}$. Elliott showed the following estimate relating the average deviations $|g(n) - A_g(X)|$ to the average gaps $|g(n) - g(n-1)|$.

Theorem [22, Thm. 10.1] *There is an absolute constant[6] $c > 0$ such that for any additive function $g : \mathbb{N} \to \mathbb{C}$ one has*

$$\frac{1}{X}\sum_{n \le X}|g(n) - A_g(X)|^2 \ll \sup_{X \le y \le X^c}\frac{1}{y}\sum_{n \le y}|g(n) - g(n-1)|^2.$$

---

[6] Hildebrand [27] showed that any $c > 4$ is admissible.

Elliott's result shows that if g has exceedingly small gaps on average, even at scales that *grow polynomially in X*, then g must globally be very close to its mean.

The drawback of this result is that it is in principle possible for the upper bound to be trivial even if the gaps $|g(n) - g(n-1)|$, $n \leq X$, are $o(B_g(X))$ on average, as long as the average savings over $n \leq X^c$ is not large enough to offset the difference in size between $B_g(X)$ and $B_g(X^c)$.

In Sect. 6, we obtain two results that complement Elliott's. The first shows that for any additive function g, *any* savings in the ℓ^1-averaged moment of $|g(n) - g(n-1)|$ provides a savings over the trivial bound for the first centred moment. The second, which holds whenever $g \in \mathcal{A}_s$, gives the same type of information as the first but in an ℓ^2 sense.

Theorem 1.11 *Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function.*

(a) The following are equivalent:

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| = o(B_g(X)), \quad \frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)| = o(B_g(X)).$$

(b) Assume furthermore that $g \in \mathcal{A}_s$. Then the following are equivalent:

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 = o(B_g(X)^2),$$

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 = o(B_g(X)^2).$$

See Proposition 6.1, where an explicit dependence between the rates of decay of the gap average and the first centred moment in Theorem 1.11(a) is given as a consequence of Theorem 1.1.

As a corollary of Theorem 1.11(b) and a second moment estimate of Ruzsa (see Lemma 3.3), we will deduce the following.

Corollary 1.12 *Let $g \in \mathcal{A}_s$ be an additive function. Assume that*

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 = o(B_g(X)^2).$$

Then there is a function $\lambda = \lambda(X)$ such that as $X \to \infty$,

$$\sum_{p^k \leq X} \frac{|g(p^k) - \lambda \log p^k|^2}{p^k} = o(B_g(X)^2).$$

Remark 1.13 Even in the weak sense of Theorem 1.11 and even when g takes bounded values at primes, it can be seen that having small gaps on average is a very special

property. As a simple example, $g = \omega$, for which $B_\omega(X)^2 \sim \log \log X$, satisfies

$$\frac{1}{X} \sum_{n \leq X} |\omega(n) - \omega(n-1)|^2 \gg \log \log X,$$

since by a bivariate version of the Erdős–Kac theorem (see e.g. [28]) one can find a positive proportion of integers $n \in [X/2, X]$ such that, simultaneously,

$$\frac{\omega(n) - \log \log X}{\sqrt{\log \log X}} \geq 2, \qquad \frac{\omega(n-1) - \log \log X}{\sqrt{\log \log X}} \leq 1.$$

In fact, as Corollary 1.12 shows, if $g \in \mathcal{A}_s$ has a small ℓ^2 average gap then g must behave like $\lambda(X) \log$ on average over prime powers $p^k \leq X$.

2 Proof ideas

In this section, we will explain the principal ideas that inform the proofs of our main theorems.

2.1 On the Matomäki–Radziwiłł type theorems

In Theorems 1.1 and 1.4 , our objective is to estimate the averaged deviations

$$\frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} g(m) - \frac{2}{X} \sum_{X/2 < m \leq X} g(m) \right|^k, \qquad (7)$$

where $k \in \{1, 2\}$, and $10 \leq h \leq X/10$ with $h \in \mathbb{Z}$. Though our result applies to any complex-valued additive function g, by considering first $\mathrm{Re}(g)$ and $\mathrm{Im}(g)$ separately it is always possible to restrict to $g(n) \in \mathbb{R}$ for all n, which we shall assume henceforth. The key idea underlying the results for both $k = 1, 2$ involves the fact that for $n \in \mathbb{N}$ and $z \in \mathbb{C}\backslash\{0\}$ the function[7] $n \mapsto z^{g(n)}$ is multiplicative in the n-aspect and analytic in the z-aspect. In the case of Theorem 1.1, for $t \in \mathbb{R}$ the corresponding function $G_t(n) := e^{2\pi i t g(n)}$ takes values on the unit circle S^1. Moreover, by replacing $G_t(n)$ by its constant (in n) multiple $\tilde{G}_t(n) := e^{2\pi i t(g(n) - A_g(X))}$ (see (6) for the definition of A_g), we see that for $r = 1, 2$,

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r} \left(\frac{1}{h} \sum_{n-h < m \leq n} \tilde{G}_t(m) \right) \Bigg|_{t=0} = (2\pi i)^r \frac{1}{h} \sum_{n-h < m \leq n} (g(m) - A_g(X))^r.$$

---

[7] We may always choose a suitable branch of logarithm to ensure that this is well defined.

Taylor expanding $\tilde{G}_t(m) = G_t(m)e^{-2\pi i t A_g(X)}$ to second order around $t = 0$ for each m leads to

$$\frac{1}{h} \sum_{n-h<m\leq n} G_t(m)$$

$$= e^{2\pi i t A_g(X)} \left(1 + \frac{2\pi i t}{h} \sum_{n-h<m\leq n} (g(m) - A_g(X)) \right)$$

$$+ (2\pi i)^2 e^{2\pi i t A_g(X)} \int_0^t \left(\frac{1}{h} \sum_{n-h<m\leq n} (g(m) - A_g(X))^2 \tilde{G}_u(m) \right) u \, du. \quad (8)$$

As the sum of $g(m) - A_g(X)$ over a medium-length interval $(n - h^*, n]$, where $h^* = X/(\log X)^c$ for a small constant $c > 0$, is well approximated by the sum over $[X/2, X]$ (see Lemma 4.1), it suffices to compare the short averages over $(n - h, n]$ to those over $(n - h^*, n]$. Using the Turán–Kubilius inequality to treat the integral error term in (8), the above allows us to approximate, for t close to 0, the average in (7) with $k = 1$ by the corresponding average

$$\frac{2}{X} \sum_{X/2<n\leq X} \left| \frac{1}{h} \sum_{n-h<m\leq n} G_t(m) - \frac{1}{h^*} \sum_{n-h^*<m\leq n} G_t(m) \right|,$$

where now our summands are, crucially, values of a bounded multiplicative function. After passing to the mean square by the Cauchy–Schwarz inequality, we may estimate these averages using the work of Matomäki and Radziwiłł [10] (and their joint work with Tao [8]), along with some additional ideas from pretentious number theory relating to the possible correlations of $G_t(n)$ with the so-called Archimedean characters $n^{i\lambda}$ for $\lambda \in \mathbb{R}$.

The above strategy fails to work in the case $k = 2$ for the important reason that the integral error term in (8), when squared and then averaged over n, cannot be controlled by an ℓ^2 moment of $g(n) - A_g(X)$, but rather only by an ℓ^4 moment. This can be far larger than $B_g(X)^4$, especially if g takes irregularly large values on prime powers.

In place of the Taylor approximation argument given above, we instead use Cauchy's integral formula to obtain an expression for short averages of g *without an error term*, namely[8] for $\rho \in (0, 1)$,

$$\frac{1}{h} \sum_{n-h<m\leq n} g(m) = \frac{d}{dz} \left(\frac{1}{h} \sum_{n-h<m\leq n} z^{g(m)} \right) \Bigg|_{z=1}$$

$$= \frac{1}{2\pi i} \int_{|z-1|=\rho} \left(\frac{1}{h} \sum_{n-h<m\leq n} z^{g(m)} \right) \frac{dz}{(z-1)^2}.$$

---

[8] Strictly speaking, we actually work with $z^{g(n)/B_g(X)}$, but for the convenience of exposition we omit the normalization by $B_g(X)$ in the exponent here.

Though this manoeuvre has eliminated the problematic error term and still introduced multiplicative functions into the game, it has also introduced a different issue in that the path of integration[9] intersects the region $|z| > 1$. Any point in that region yields a function $n \mapsto z^{g(n)}$ that is unbounded whenever g takes unbounded, positive values, say.

While this issue prevents us from obtaining an ℓ^2 result for arbitrary additive functions g, we may still succeed if we impose restrictions on the growth of g. Indeed, as shown in [14], the work of Matomäki and Radziwiłł can be generalized to cover certain collections of unbounded multiplicative functions of controlled growth. This includes most natural multiplicative functions f that are uniformly bounded on the primes and not too large on average at prime powers. Assuming the hypothesis $g \in \mathcal{A}_s$ and modifying g on a small set of prime powers, it can be shown that the resulting multiplicative function $z^{g(n)/B_g(X)}$ satisfies the necessary hypotheses for the generalization of the Matomäki–Radziwiłł theorem in [14] to be applicable, which is crucial to the proof of Theorem 1.4.

2.2 On gaps between consecutive values of additive functions

Theorem 1.11 establishes that for suitable additive functions g, having a small kth moment of gaps is equivalent to having a small kth centred moment, for $k \in \{1, 2\}$. Since the proof follows similar lines in each of the cases $k = 1, 2$, we will confine ourselves mainly to explaining the case $k = 1$ here.

By the triangle inequality,

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| \leq \frac{1}{X} \sum_{n \leq X} (|g(n) - A_g(X)| + |g(n-1) - A_g(X)|),$$

which implies that if the first centred moment is $o(B_g(X))$ then the average gap is also $o(B_g(X))$.

The converse is more delicate. The main idea here is to note that if $h = h(X)$ is slowly growing then as $h < n \leq X$ varies the average gap $|g(n) - g(n-1)|$ controls the size of typical differences between $g(n)$ and its length h averages:

$$g(n) - \frac{1}{h} \sum_{n-h < m \leq n} g(m) = \sum_{0 \leq j \leq h-2} \left(1 - \frac{j+1}{h}\right) (g(n-j) - g(n-j-1)). \tag{9}$$

Thus, if we assume that g has gaps $|g(n) - g(n-1)|$ of size $o(B_g(X))$ on average, then (by selecting h growing sufficiently slowly) the left-hand side of (9) will also typically be small. Now, Theorem 1.1 allows us to conclude that for *almost all* $n \in [X, 2X]$,

---

[9] This would still be true if we replaced the circle $|z - 1| = \rho$ by any other path containing 1 in its interior component.

$$\left| g(n) - \frac{1}{h} \sum_{n-h<m\leq n} g(m) \right| = \left| g(n) - \frac{2}{X} \sum_{X/2<n\leq X} g(n) \right| + o(B_g(X))$$

$$= |g(n) - A_g(X)| + o(B_g(X)),$$

and in this way we deduce that $|g(n) - A_g(X)|$ is also $o(B_g(X))$ on average. The corresponding result comparing the 2nd moments is analogous, but relies on our Theorem 1.4 instead of Theorem 1.1. For this reason, we must assume that $g \in \mathcal{A}_s$ in Theorem 1.11(b).

2.3 On the Erdős monotonicity problem

Our application to Erdős' problem, Conjecture 1.6, was the original motivation for this paper. The connection between our short interval average results and this conjecture arose from the observation that if g is a real-valued additive function that is non-decreasing outside of a set $\mathcal{B}$ of density 0, then the average of the gaps of $\{g(n)\}_n$ is nearly a telescoping sum, that is

$$\frac{1}{X} \sum_{n\leq X} |g(n) - g(n-1)| = \frac{1}{X} \sum_{n\leq X} (g(n) - g(n-1)) + \frac{2}{X} \sum_{n\in\mathcal{B}(X)} (g(n-1) - g(n))$$

$$= \frac{g(\lfloor X \rfloor)}{X} + \frac{2}{X} \sum_{n\in\mathcal{B}(X)} (g(n-1) - g(n)), \qquad (10)$$

since $g(0) = 0$ by definition. It can be shown (see Lemma 3.5) that $|g(\lfloor X \rfloor)|/X = o(B_g(X))$; via the Cauchy–Schwarz inequality, the sparseness of $\mathcal{B}$ results in the second expression also being $o(B_g(X))$. By Theorem 1.11(a), which, as just discussed, is a consequence of Theorem 1.1, the *first* centred moment thus also satisfies $\frac{1}{X} \sum_{n\leq X} |g(n) - A_g(X)| = o(B_g(X))$.

A classical second moment estimate of Ruzsa (see Lemma 3.3) shows that if, instead, we could obtain savings over $O(B_g(X)^2)$ for the *second* centred moment $\frac{1}{X} \sum_{n\leq X} |g(n) - A_g(X)|^2$, then we could conclude the existence of a *slowly-varying* function $\lambda = \lambda(X)$ such that $g_\lambda = g - \lambda \log$ takes smaller values on average over prime powers than g does. That is, $\lambda \log n$ *approximates* $g(n)$ in a precise sense. Achieving such savings in the second centred moment is the objective of the proof of Theorem 1.8.

In analogy to the treatment of the first moment of the gaps in (10), the bulk of the work towards Theorem 1.8 involves obtaining savings over $B_g(X)^2$ on sparsely-supported ℓ^2 sums of the shape

$$\frac{1}{X} \sum_{n\in\mathcal{S}(X)} |g(n) - A_g(X)|^2,$$

where $\mathcal{S} \subset \mathbb{N}$ and $\mathcal{S}(X) := \mathcal{S} \cap [1, X]$ satisfies $|\mathcal{S}(X)| = o(X)$, as $X \to \infty$. Having no recourse to Hölder's inequality for savings in ℓ^2, we instead use the large sieve

(see Proposition 7.1), together with some ideas due to Elliott, to show that either this sparse average is $o(B_g(X)^2)$, or else $\mathcal{S}$ contains many multiples of a sparse set of primes p where $|g(p)|$ is extremely large (in a precise sense). As $g \in \mathcal{A}_s$, this latter set is provably empty, and consequently we obtain the required savings. It would be interesting to understand whether a similar conclusion could be obtained under weaker conditions on g.

The slow variation of λ i.e. $\lambda(X^u) = (1 + o(1))\lambda(X)$ for fixed $0 < u \leq 1$ is a key property that we exploit in the proof of Corollary 1.7. Though we do not need to directly invoke the general theory of slowly-varying functions due to Karamata (see e.g. [29, Ch. 1]), his representation theorem informs our proof that λ is slowly growing in X i.e. $\lambda(X) \in [(\log X)^{-\varepsilon}, (\log X)^{\varepsilon}]$ for any $\varepsilon > 0$ (see Lemma 8.6). Given that, provably, $B_g(X) \asymp \lambda(X) \log X$ here, we find that $B_g(X) = (\log X)^{1+o(1)}$. For reference, as noted above we have $B_g(X) \sim \frac{|c|}{\sqrt{2}} \log X$ whenever $g = c \log$.

Corollary 1.7 follows readily from this conclusion, since if $|\mathcal{B}(X)| \ll X/(\log X)^{2+\delta}$ for some $\delta > 0$, then by Cauchy–Schwarz we have

$$\frac{1}{X} \sum_{n \in \mathcal{B}(X)} |g(n) - g(n-1)| \ll \left(\frac{|\mathcal{B}(X)|}{X} \right)^{1/2} B_g(X) = o(1).$$

Since $g(\lfloor X \rfloor)/X = o(B_g(X)/(\log X)^2) = o(1)$, the right-hand side in (10) is thus $o(1)$, and so the Kátai–Wirsing theorem mentioned in the introduction (see also Theorem 3.7 for a statement) implies that $g = c \log$ exactly, for some $c \in \mathbb{R}$. Without this additional sparseness assumption on $\mathcal{B}$, however, it is not clear how to proceed further. It would be interesting to obtain the bound $\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| = o(1)$, even assuming $g \in \mathcal{A}_s$, under weaker hypotheses on the rate of decay of $|\mathcal{B}(X)|/X$, or perhaps assuming to begin with that $B_g(X) = (\log X)^{1+o(1)}$.

3 Auxiliary lemmas

In this section, we record several results that will be used repeatedly in the sequel. For the convenience of the reader, we recall that for an additive function $g : \mathbb{N} \to \mathbb{C}$ and $X \geq 2$,

$$A_g(X) := \sum_{p^k \leq X} \frac{g(p^k)}{p^k} \left(1 - \frac{1}{p} \right), \quad B_g(X)^2 := \sum_{p^k \leq X} \frac{|g(p^k)|^2}{p^k}.$$

Lemma 3.1 *Let $g : \mathbb{N} \to \mathbb{C}$ be additive. Then for any $Y \geq 3$,*

$$\frac{1}{Y} \sum_{n \leq Y} g(n) = A_g(Y) + O \left(\frac{B_g(Y)}{\sqrt{\log Y}} \right).$$

Proof As $g(n) = \sum_{p^k \| n} g(p^k)$, we have

$$
\begin{aligned}
\frac{1}{Y} \sum_{n \le Y} g(n) - A_g(Y) &= \frac{1}{Y} \sum_{p^k \le Y} g(p^k) \sum_{\substack{n \le Y \\ p^k \| n}} 1 - \sum_{p^k \le Y} \frac{g(p^k)}{p^k} \left(1 - \frac{1}{p}\right) \\
&= \sum_{p^k \le Y} g(p^k) \left(Y^{-1} \left(\left\lfloor \frac{Y}{p^k} \right\rfloor - \left\lfloor \frac{Y}{p^{k+1}} \right\rfloor \right) - \frac{1}{p^k} (1 - 1/p) \right) \\
&\ll \frac{1}{Y} \sum_{p^k \le Y} |g(p^k)| \le Y^{-1/2} \sum_{p^k \le Y} |g(p^k)| p^{-k/2} \\
&\ll B_g(Y) (\pi(Y)/Y)^{1/2} \ll \frac{B_g(Y)}{\sqrt{\log Y}},
\end{aligned}
$$

using the Cauchy–Schwarz inequality and Chebyshev's estimate $\pi(Y) \ll Y/\log Y$ in the last two steps. □

Lemma 3.2 (Turán–Kubilius Inequality) *Let $X \ge 3$. Uniformly over all additive functions $g : \mathbb{N} \to \mathbb{C}$,*

$$
\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)|^2 \ll B_g(X)^2.
$$

Proof This is e.g. [22, Lem. 1.5] (taking $\sigma = 0$). □

The following estimate due to Ruzsa, which sharpens the Turán–Kubilius inequality, gives an order of magnitude estimate for the second centred moment of a general additive function.

Lemma 3.3 [30] *Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function. Then*

$$
\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)|^2 \asymp \min_{\lambda \in \mathbb{R}} (B_{g_\lambda}(X)^2 + |\lambda|^2) \asymp B_{g_{\lambda_0}}(X)^2 + |\lambda_0|^2,
$$

where for $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$ we set $g_\lambda(n) := g(n) - \lambda \log n$, and $\lambda_0 = \lambda_0(X)$ is given by

$$
\lambda_0(X) := \frac{2}{(\log X)^2} \sum_{p \le X} \frac{g(p) \log p}{p}.
$$

Lemma 3.4 *Let $g : \mathbb{N} \to \mathbb{C}$ be additive, and let $z \ge y \ge 2$. Then*

$$
A_g(y) = A_g(z) + O\left(B_g(z) \sqrt{\log\left(\tfrac{\log z}{\log y}\right)} \right).
$$

In particular, if $y \in (z/2, z]$ then

$$A_g(y) = A_g(z) + O\left(\frac{B_g(z)}{\sqrt{\log z}}\right).$$

Proof By Mertens' theorem,

$$|A_g(z) - A_g(y)| \leq \sum_{y < p^k \leq z} \frac{|g(p^k)|}{p^k} \leq B_g(z) \left(\sum_{y < p^k \leq z} \frac{1}{p^k}\right)^{1/2} \ll B_g(z)\sqrt{\log(\tfrac{\log z}{\log y})}.$$

The second claim follows immediately from this. □

Lemma 3.5 *Let $X \geq 3$ and let $n \in (X/2, X]$. Then $\frac{|g(n)|}{n} \ll \frac{B_g(X) \log X}{\sqrt{X}}$.*

Proof Observe that whenever $p^k \leq X$ we have $|g(p^k)|/p^{k/2} \leq B_g(X)$. It follows from the triangle inequality and the bound $\omega(n) \ll \log n$ for all $n \geq 2$ that

$$\frac{|g(n)|}{n} \leq \frac{\omega(n)}{n} \max_{p^k \| n} |g(p^k)| \leq \frac{\omega(n)}{\sqrt{n}} \max_{p^k \| n} \frac{|g(p^k)|}{p^{k/2}} \ll \frac{\log X}{\sqrt{X}} B_g(X),$$

as claimed. □

Working within the collection $\mathcal{A}$ (see Definition 1.3), the following properties will be useful.

Lemma 3.6 *(a) Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function satisfying $B_g(X) \to \infty$. If g is either completely or strongly additive then $g \in \mathcal{A}$.*
(b) Let $g \in \mathcal{A}$. Then there is a strongly additive function g^* such that $g(p) = g^*(p)$ for all primes p, and $B_{g-g^*}(X) = o(B_g(X))$ as $X \to \infty$.

Proof (a) Let g be either strongly or completely additive. We put $\theta_g := 1$ if g is completely additive, and $\theta_g := 0$ otherwise. Then $g(p^k) = k^{\theta_g} g(p)$ for any prime power p^k, and thus

$$\sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|g(p^k)|^2}{p^k} \leq \sum_{p \leq X} \frac{|g(p)|^2}{p} \sum_{k \geq 2} \frac{k^{2\theta_g}}{p^{k-1}} \ll \sum_{p \leq X} \frac{|g(p)|^2}{p^2}.$$

Since $B_g(X) \to \infty$, choosing $M = M(X)$ tending to infinity arbitrarily slowly we see that

$$\sum_{p \leq X} \frac{|g(p)|^2}{p^2} \leq \sum_{p \leq M} \frac{|g(p)|^2}{p} + \frac{1}{M} \sum_{M < p \leq X} \frac{|g(p)|^2}{p}$$

$$\leq B_g(M)^2 + \frac{B_g(X)^2}{M} = o(B_g(X)^2).$$

It follows that $g \in \mathcal{A}$, as required.

(b) We define g^* to be an additive function defined by $g^*(p^k) := g(p)$ for all primes p and $k \geq 1$. Thus, g^* is strongly additive. Moreover, if $(g - g^*)(p^k) \neq 0$ then $k \geq 2$, for any p. By assumption and part (a), $g, g^* \in \mathcal{A}$, and thus

$$\sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|h(p^k)|^2}{p^k} \ll \sum_{p \leq X} \frac{|h(p^k)|^2}{p^2} = o(B_g(X)^2)$$

for both $h = g$ and $h = g^*$. By the Cauchy–Schwarz inequality,

$$B_{g-g^*}(X)^2 \ll \sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|g(p^k)|^2}{p^k} + \sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|g^*(p^k)|^2}{p^k} = o(B_g(X)^2),$$

as required. □

Finally, we record the characterization result of Kátai and Wirsing, mentioned in the introduction.

Theorem 3.7 [18, 21] *Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function such that as $X \to \infty$,*

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| = o(1).$$

Then there is $c \in \mathbb{C}$ such that $g(n) = c \log n$ for all $n \in \mathbb{N}$.

4 The Matomäki–Radziwiłł theorem for additive functions: ℓ^1 variant

In this section, we prove Theorem 1.1.

We begin with the following simple observation, amounting to the fact that the mean value of an additive function changes little when passing from a long interval of length $\asymp X$ to a medium-sized one of length $X/(\log X)^c$, for $c > 0$ sufficiently small.

Lemma 4.1 *Let $g : \mathbb{N} \to \mathbb{C}$ be additive and let X be large. Let $X/2 < x \leq X$, and let $X/(\log X)^{1/3} \leq h \leq X/3$. Then*

$$\frac{2}{X} \sum_{X/2 < n \leq X} g(n) = \frac{1}{h} \sum_{x-h \leq n \leq x} g(n) + O\left(\frac{B_g(X)}{(\log X)^{1/6}}\right).$$

Proof Applying Lemma 3.1 with $Y = X/2$, X, $x - h$ and x, we obtain

$$\frac{2}{X} \sum_{X/2 < n \leq X} g(n) = 2A_g(X) - A_g(X/2) + O\left(\frac{B_g(X)}{\sqrt{\log X}}\right)$$

$$\frac{1}{h} \sum_{x-h<n\leq x} g(n) = \frac{x}{h} A_g(x) - \left(\frac{x}{h} - 1\right) A_g(x-h) + O\left(\frac{X B_g(X)}{h\sqrt{\log X}}\right).$$

Since $h \geq X/(\log X)^{1/3}$, the error term in the second line is $\ll \frac{B_g(X)}{(\log X)^{1/6}}$. By Lemma 3.4,

$$2A_g(X) - A_g(X/2) = A_g(X) + O(|A_g(X) - A_g(X/2)|) = A_g(X) + O\left(\frac{B_g(X)}{\sqrt{\log X}}\right)$$

for the main term in the first equation, and also

$$\frac{x}{h}|A_g(x) - A_g(x-h)| \ll (\log X)^{1/3} \cdot \frac{B_g(x)}{\sqrt{\log x}} \ll \frac{B_g(X)}{(\log X)^{1/6}},$$

so that by a second application of Lemma 3.4,

$$\frac{x}{h} A_g(x) - \left(\frac{x}{h} - 1\right) A_g(x-h) = A_g(x) + O\left(\frac{x}{h}|A_g(x) - A_g(x-h)|\right)$$
$$= A_g(X) + O\left(\frac{B_g(X)}{(\log X)^{1/6}}\right).$$

Combining these estimates, we may conclude that

$$\left| \frac{2}{X} \sum_{X/2<n\leq X} g(n) - \frac{1}{h} \sum_{x-h<n\leq x} g(n) \right| \ll \frac{B_g(X)}{(\log X)^{1/6}},$$

as claimed. $\qquad\square$

In light of the above lemma, to prove Theorem 1.1 it suffices to prove the following: if $h' = X/(\log X)^{1/3}$ and $10 \leq h \leq h'$ then

$$\frac{2}{X} \sum_{X/2<m\leq X} \left| \frac{1}{h} \sum_{m-h<n\leq m} g(n) - \frac{1}{h'} \sum_{m-h'<n\leq m} g(n) \right|$$
$$\ll B_g(X) \left(\sqrt{\frac{\log\log h}{\log h}} + (\log X)^{-1/800} \right).$$

Splitting $g = \operatorname{Re}(g) + i\operatorname{Im}(g)$, and noting that both $\operatorname{Re}(g)$ and $\operatorname{Im}(g)$ are real-valued additive functions, we may assume that g is itself real-valued, after which the general case will follow by the triangle inequality.

Let $10 \leq h \leq X/3$, with X large. Following [8], fix $\eta \in (0, 1/12)$, parameters $Q_1 = h$, $P_1 = (\log h)^{40/\eta}$, and define further parameters P_j, Q_j by

$$P_j := \exp\left(j^{4j} (\log Q_1)^{j-1} \log P_1 \right), \quad Q_j := \exp\left(j^{4j+2} (\log Q_1)^j \right),$$

for all $j \leq J$, where J is chosen maximally subject to $Q_J \leq \exp(\sqrt{\log X})$. We then define

$$\mathcal{S} = \mathcal{S}_{X,P_1,Q_1} := \{n \leq X : \omega_{[P_j,Q_j]}(n) \geq 1 \text{ for all } 1 \leq j \leq J\},$$

where for any set $S \subset \mathbb{N}$ we write $\omega_S(n) := \sum_{p|n} 1_S(p)$.

The following key step in the proof of Theorem 1.1 allows us to pass from comparing averages of the additive g to averages of a corresponding multiplicative function, supported on $\mathcal{S}$.

Lemma 4.2 *Let $g : \mathbb{N} \to \mathbb{R}$ be an additive function. Let $10 \leq h \leq h'$, where $h' := \frac{X}{(\log X)^{1/3}}$ and $h \in \mathbb{Z}$, and $(\log X)^{-1/6} < t < 1$. Then*

$$\frac{2}{X} \sum_{X/2 < m \leq X} \left| \frac{1}{h'} \sum_{m-h' < n \leq m} g(n) - \frac{1}{h} \sum_{m-h < n \leq m} g(n) \right|$$

$$\ll \frac{B_g(X)}{t} \cdot \frac{1}{X} \int_{X/2}^{X} \left| \frac{1}{h'} \sum_{\substack{x-h' < n \leq x \\ n \in \mathcal{S}}} e(t\tilde{g}(n; X)) - \frac{1}{h} \sum_{\substack{x-h < n \leq x \\ n \in \mathcal{S}}} e(t\tilde{g}(n; X)) \right| dx$$

$$+ B_g(X) \left(t + \frac{\log \log h}{t \log h} \right),$$

where $\tilde{g}(n; X) := B_g(X)^{-1}(g(n) - A_g(X))$ for all $n \in \mathbb{N}$.

Proof In view of Lemma 3.5, at the cost of an error term of size $\max_{X/2 < n \leq X} |g(n)|/h' \ll B_g(X)X^{-1/4}$, we may assume that both $h, h' \in \mathbb{Z}$ (else replace h' by $\lfloor h' \rfloor$). Given $u \in [0, 1]$, $x \in [X/2, X] \cap \mathbb{Z}$ and an integer $1 \leq H \leq h'$, define

$$S_H(u; x) := \frac{1}{H} \sum_{x-H < n \leq x} e(u\tilde{g}(n; X)),$$

which is clearly an analytic function of u. Fix $x \in [X/2, X] \cap \mathbb{Z}$, and observe that $S_{h'}(0; x) = 1 = S_h(0; x)$. By Taylor expansion in t,

$$S_{h'}(t; x) - S_h(t; x) = t(S'_{h'}(0; x) - S'_h(0; x)) + \int_0^t (S''_{h'}(u; x) - S''_h(u; x))u \, du,$$

$$(11)$$

wherein we have

$$S'_{h'}(0; x) - S'_h(0; x) \tag{12}$$

$$= \frac{2\pi i}{B_g(X)} \left(\frac{1}{h'} \sum_{x-h' < n \leq x} (g(n) - A_g(X)) - \frac{1}{h} \sum_{x-h < n \leq x} (g(n) - A_g(X)) \right)$$

$$= \frac{2\pi i}{B_g(X)} \left(\frac{1}{h'} \sum_{x-h'<n\leq x} g(n) - \frac{1}{h} \sum_{x-h<n\leq x} g(n) \right). \tag{13}$$

By inserting the expression (13) into (11), rearranging the latter and then taking absolute values and averaging over $x \in [X/2, X] \cap \mathbb{Z}$, we find

$$\frac{2}{X} \sum_{X/2<m\leq X} \left| \frac{1}{h'} \sum_{m-h'<n\leq m} g(n) - \frac{1}{h} \sum_{m-h<n\leq m} g(n) \right|$$

$$\ll B_g(X)t^{-1} \cdot \frac{1}{X} \sum_{X/2<m\leq X} \left| \frac{1}{h'} \sum_{m-h'<n\leq m} e(t\tilde{g}(n;X)) - \frac{1}{h} \sum_{m-h<n\leq m} e(t\tilde{g}(n;X)) \right|$$

$$+ B_g(X)t^{-1} \cdot \frac{t^2}{X} \sum_{X/2<m\leq X} \max_{0\leq u\leq t} |\frac{1}{h'} \sum_{m-h'<n\leq m} \tilde{g}(n;X)^2 e(u\tilde{g}(n;X))$$

$$- \frac{1}{h} \sum_{m-h<n\leq m} \tilde{g}(n;X)^2 e(u\tilde{g}(n;X))|.$$

Since g is real-valued by assumption, $|e(u\tilde{g}(n;X))| = 1$ for all n. Thus, applying the triangle inequality and Lemma 3.2, we may bound the last expression above by

$$\ll t B_g(X) \frac{1}{X} \sum_{X/2<m\leq X} \sum_{H\in\{h,h'\}} \frac{1}{H} \sum_{m-H<n\leq m} \left(\frac{g(n) - A_g(X)}{B_g(X)} \right)^2$$

$$\ll t B_g(X) \cdot \frac{1}{X} \sum_{X/2-h'<n\leq X} \left(\frac{g(n) - A_g(X)}{B_g(X)} \right)^2 \cdot \sum_{H\in\{h,h'\}} \frac{1}{H} \sum_{X/2<m\leq X} 1_{[n,n+H]}(m)$$

$$\ll t B_g(X).$$

We now split

$$S_H(t;x) = \frac{1}{H} \sum_{\substack{x-H<n\leq x \\ n\in\mathcal{S}}} e(t\tilde{g}(n;X)) + \frac{1}{H} \sum_{\substack{x-H<n\leq x \\ n\notin\mathcal{S}}} e(t\tilde{g}(n;X))$$

$$=: S_H^{(\mathcal{S})}(t;x) + S_H^{(\mathcal{S}^c)}(t;x),$$

with $H \in \{h, h'\}$. By the triangle inequality, we have

$$\frac{2}{X} \sum_{X/2<x\leq X} |S_H^{(\mathcal{S}^c)}(t;x)| \leq \frac{1}{HX} \sum_{X/2<x\leq X} |\mathcal{S}^c \cap (x-H,x]| \leq \frac{1}{X} \sum_{X/3<n\leq X} 1_{\mathcal{S}^c}(n).$$

Since $P_J \leq \exp(\sqrt{\log X})$, the union bound and the fundamental lemma of the sieve (see [31, Remark after Lem. 6.3]) yield

$$|\mathcal{S}^c \cap [X/3, X]| \ll X \sum_{1 \leq j \leq J} \prod_{P_j \leq p \leq Q_j} \left(1 - \frac{1}{p}\right) \ll X \sum_{1 \leq j \leq J} \frac{\log P_j}{\log Q_j}$$

$$= X \frac{\log P_1}{\log Q_1} \sum_{1 \leq j \leq J} \frac{1}{j^2} \ll X \frac{\log \log h}{\log h}. \tag{14}$$

We thus find by the triangle inequality that

$$\frac{2}{X} \sum_{X/2 < n \leq X} |S_h(t; n) - S_{h'}(t; n)| \ll \frac{2}{X} \sum_{X/2 < n \leq X} \left| S_h^{(\mathcal{S})}(t; n) - S_{h'}^{(\mathcal{S})}(t; n) \right| + \frac{\log \log h}{\log h}.$$

Finally, if $n \in [X/2, X] \cap \mathbb{Z}$ and $x \in [n, n+1)$ then $S_H^{(\mathcal{S})}(t; x) = S_H^{(\mathcal{S})}(t; n) + O(1/H)$, and thus

$$\frac{2}{X} \sum_{X/2 < n \leq X} |S_h^{(\mathcal{S})}(t; n) - S_{h'}^{(\mathcal{S})}(t; n)| \leq \frac{2}{X} \int_{X/2}^X |S_h^{(\mathcal{S})}(t; x) - S_{h'}^{(\mathcal{S})}(t; x)| dx + O(1/h).$$

Combined with the preceding estimates, we obtain

$$\frac{2}{X} \sum_{X/2 < m \leq X} \left| \frac{1}{h'} \sum_{m-h' < n \leq m} g(n) - \frac{1}{h} \sum_{m-h < n \leq m} g(n) \right|$$

$$\ll t^{-1} B_g(X) \left(\frac{2}{X} \int_{X/2}^X \left| S_h^{(\mathcal{S})}(t; x) - S_{h'}^{(\mathcal{S})}(t; x) \right| dx + \frac{\log \log h}{\log h} \right) + t B_g(X),$$

which implies the claim. □

Define the multiplicative function

$$G_{t,X}(n) := e(tg(n)/B_g(X)) = e(t\tilde{g}(n; X))e(tA_g(X)/B_g(X)).$$

In light of Lemma 4.2, the proof of Theorem 1.1 essentially boils down to the following comparison result for short- and medium-length interval averages of $G_{t,X}$.

Lemma 4.3 *Let $g : \mathbb{N} \to \mathbb{R}$ be an additive function. Let $X \geq 3$ be large, $(\log X)^{-1/6} < t \leq 1/100$ be small and let $10 \leq h_1 \leq h_2$ where $h_2 = X/(\log X)^{1/3}$. Then*

$$\frac{2}{X} \int_{X/2}^X \left| \frac{1}{h_1} \sum_{\substack{x-h_1 < n \leq x \\ n \in \mathcal{S}}} G_{t,X}(n) - \frac{1}{h_2} \sum_{\substack{x-h_2 < n \leq x \\ n \in \mathcal{S}}} G_{t,X}(n) \right| dx \tag{15}$$

$$\ll B_g(X) \left(\frac{\log \log h_1}{\log h_1} + (\log X)^{-1/400} \right). \tag{16}$$

To prove Lemma 4.3 we will appeal to some ideas from pretentious analytic number theory. Let $\mathbb{U} := \{z \in \mathbb{C} : |z| \leq 1\}$. In what follows, given multiplicative functions $f, g : \mathbb{N} \to \mathbb{U}$ and parameters $1 \leq T \leq X$, we introduce the pretentious distance of Granville and Soundararajan:

$$\mathbb{D}(f, g; X)^2 := \sum_{p \leq X} \frac{1 - \operatorname{Re}(f(p)\overline{g}(p))}{p},$$

$$M_f(X; T) := \min_{|\lambda| \leq T} \mathbb{D}(f, n^{i\lambda}; X)^2.$$

For multiplicative functions f, g, h taking values in $\mathbb{U}$, it is well known (see e.g. [32, Lem 3.1]) that $\mathbb{D}$ satisfies the triangle inequality

$$\mathbb{D}(f, h; X) \leq \mathbb{D}(f, g; X) + \mathbb{D}(g, h; X). \tag{17}$$

For each $t \in [0, 1]$, select $\lambda_{t,X} \in [-X, X]$ such that $M_{G_{t,X}}(X; X) = \mathbb{D}(G_{t,X}, n^{i\lambda_{t,X}}; X)^2$ (if there are multiple such minimizers, pick any one of them).

Lemma 4.4 *Let $0 < t \leq 1/100$ be sufficiently small. Then either*

(i) $M_{G_{t,X}}(X; X) \geq \frac{1}{25} \log\log X$, *or else*
(ii) $|\lambda_{t,X}| = O(1)$.

Proof Assume (i) fails. Then by assumption, $\mathbb{D}(G_{t,X}, n^{i\lambda_{t,X}}; X)^2 \leq \frac{1}{25} \log\log X$. We claim that there is also $\tilde{\lambda}_{t,X} = O(1)$ such that

$$\mathbb{D}(G_{t,X}, n^{i\tilde{\lambda}_{t,X}}; X) \ll 1. \tag{18}$$

To see that this is sufficient to prove (ii), we apply (17) to obtain

$$\mathbb{D}(n^{i\lambda_{t,X}}, n^{i\tilde{\lambda}_{t,X}}; X) \leq O(1) + \frac{1}{5}\sqrt{\log\log X} \leq \sqrt{0.3 \log\log X}$$

for large enough X. Now, if $|\lambda_{t,X} - \tilde{\lambda}_{t,X}| \geq 100$ then as $|\lambda_{t,X}|, |\tilde{\lambda}_{t,X}| \leq X$ the Vinogradov–Korobov zero-free region for ζ (see e.g. [9, (1.12)]) gives

$$\mathbb{D}(n^{i\lambda_{t,X}}, n^{i\tilde{\lambda}_{t,X}}; X)^2 = \log\log X - \log|\zeta(1 + 1/\log X + i(\lambda_{t,X} - \tilde{\lambda}_{t,X}))| + O(1)$$
$$\geq 0.33 \log\log X,$$

which is a contradiction. It follows that

$$|\lambda_{t,X}| \leq |\tilde{\lambda}_{t,X}| + 100 = O(1),$$

as required.

It thus remains to prove that (18) holds. By Lemma 3.2, we obtain

$$|\{n \leq X : |\tilde{g}(n; X)| > t^{-1/2}\}| \leq t \sum_{n \leq X} \tilde{g}(n; X)^2 \ll tX.$$

It follows from Taylor expansion that

$$\sum_{n \leq X} e(t\tilde{g}(n; X)) = \sum_{n \leq X} \left(1 + O(\sqrt{t})\right) + O(tX) = (1 + O(\sqrt{t}))X.$$

On the other hand, by Halász' theorem in the form of Granville and Soundararajan [33, Thm. 1],

$$\left| \sum_{n \leq X} e(t\tilde{g}(n; X)) \right| = \left| \sum_{n \leq X} G_{t,X}(n) \right| \ll M_{G_{t,X}}(X; U)e^{-M_{G_{t,X}}(X;U)}X + \frac{X}{U},$$

where $1 \leq U \leq \log X$ is a parameter of our choice. If U is a suitably large absolute constant and t is sufficiently small in an absolute sense, we obtain $M_{G_{t,X}}(X; U) \ll 1$, and therefore there is a $\tilde{\lambda}_{t,X} \in [-U, U]$ (thus of size $O(1)$) such that

$$\mathbb{D}(G_{t,X}, n^{i\tilde{\lambda}_{t,X}}; X) \ll 1,$$

as claimed.   □

Proof of Lemma 4.3 Set $\varepsilon = (\log X)^{-1/100}$. If $M_{G_{t,X}}(X; X) \geq 4\log(1/\varepsilon)$ then by the triangle inequality, Cauchy–Schwarz and [9, Theorem A.2], the LHS of (15) is

$$\ll \sum_{j=1,2} \left(\frac{1}{X} \int_{X/2}^{X} \left| \frac{1}{h_j} \sum_{\substack{x-h_j < n \leq x \\ n \in \mathcal{S}}} G_{t,X}(n) \right|^2 dx \right)^{1/2}$$

$$\ll \exp\left(-\frac{1}{2} M_{G_{t,X}}(X; X) \right) M_{G_{t,X}}(X; X)^{1/2} + \frac{(\log h_1)^{1/6}}{P_1^{1/12-\eta/2}} + (\log X)^{-1/200}$$

$$\ll \varepsilon^2 (\log(1/\varepsilon))^{1/2} + (\log h_1)^{-1} + (\log X)^{-1/200}$$

$$\ll (\log h_1)^{-1} + (\log X)^{-1/200}.$$

Next, assume that $M_{G_{t,X}}(X; X) < 4\log(1/\varepsilon)$. For $\lambda \in \mathbb{R}$ and $h \geq 1$ define

$$I(x; \lambda, h) := h^{-1} \int_{x-h}^{x} u^{i\lambda} du.$$

By Lemma 4.4 we have $\lambda_{t,X} = O(1)$, so that with $h \in \{h_1, h_2\}$,

$$I(x; \lambda_{t,X}, h) = x^{i\lambda_{t,X}} \left(\frac{1 - (1 - h/x)^{1+i\lambda_{t,X}}}{(1 + i\lambda_{t,X})h/x} \right) = x^{i\lambda_{t,X}} \left(1 + O\left(|\lambda_{t,X}| \frac{h}{X} \right) \right)$$
$$= x^{i\lambda_{t,X}} \left(1 + O\left(\frac{h}{X} \right) \right),$$

and thus for each $x \in [X/2, X]$ and $j = 1, 2$,

$$x^{i\lambda_{t,X}} \frac{2}{X} \sum_{\substack{X/2 < n \leq X \\ n \in \mathcal{S}}} G_{t,X}(n) n^{-i\lambda_{t,X}}$$
$$= I(x; \lambda_{t,X}, h_j) \cdot \frac{2}{X} \sum_{\substack{X/2 < n \leq X \\ n \in \mathcal{S}}} G_{t,X}(n) n^{-i\lambda_{t,X}} + O\left(\frac{h_j}{X} \right). \tag{19}$$

Reinstating the $n \notin \mathcal{S}$ and using the arguments surrounding (14), we also note that

$$\frac{2}{X} \int_{X/2}^{X} \left| \frac{1}{h_j} \sum_{\substack{x-h_j < n \leq x \\ n \in \mathcal{S}}} G_{t,X}(n) - x^{i\lambda_{t,X}} \frac{2}{X} \sum_{\substack{X/2 < n \leq X \\ n \in \mathcal{S}}} G_{t,X}(n) n^{-i\lambda_{t,X}} \right| dx$$
$$\ll \frac{2}{X} \int_{X/2}^{X} \left| \frac{1}{h_j} \sum_{x-h_j < n \leq x} G_{t,X}(n) - x^{i\lambda_{t,X}} \frac{2}{X} \sum_{X/2 < n \leq X} G_{t,X}(n) n^{-i\lambda_{t,X}} \right| dx$$
$$+ \frac{\log\log h_1}{\log h_1}.$$

Adding and subtracting the expression on the LHS of (19) inside the absolute values bars in (15), we obtain the upper bound

$$\ll \mathcal{T}_1 + \mathcal{T}_2 + \frac{h_2}{X},$$

where we have set

$$\mathcal{T}_j := \frac{2}{X} \int_{X/2}^{X} \left| I(x; \lambda_{t,X}, h_j) \frac{2}{X} \sum_{\substack{X/2 < n \leq X \\ n \in \mathcal{S}}} G_{t,X}(n) n^{-i\lambda_{t,X}} - \frac{1}{h_j} \sum_{\substack{x-h_j < n \leq x \\ n \in \mathcal{S}}} G_{t,X}(n) \right|.$$

If $j = 2$ then as just noted we also have

$$\mathcal{T}_2 \ll \frac{1}{X} \int_{X/2}^{X} \left| \frac{1}{h_2} \sum_{x-h_2 < n \leq x} G_{t,X}(n) - I(x; \lambda_{t,X}, h_2) \frac{2}{X} \sum_{X/2 < n \leq X} G_{t,X}(n) n^{-\lambda_{t,X}} \right| dx$$
$$+ \frac{\log\log h_1}{\log h_1},$$

and so by Cauchy–Schwarz and [34, Theorem 1.6] (taking $Q = 1$ and $\varepsilon = (\log X)^{-1/200}$ there), we have

$$
\mathcal{T}_2 \ll \left(\frac{2}{X} \int_{X/2}^{X} \left| I(x; \lambda_{t,X}, h_2) \frac{2}{X} \sum_{X/2 < n \leq X} G_{t,X}(n) n^{-i\lambda_{t,X}} - \frac{1}{h_2} \sum_{x-h_2 < n \leq x} G_{t,X}(n) \right|^2 dx \right)^{\frac{1}{2}}
$$
$$
+ \frac{\log \log h_1}{\log h_1}
$$
$$
\ll (\log X)^{-1/400} + \frac{\log \log h_1}{\log h_1}.
$$

In the same way, when $h_1 > \sqrt{X}$ we obtain the same bound $\mathcal{T}_1 \ll (\log X)^{-1/400} + \frac{\log \log h_1}{\log h_1}$ as well.

Thus, assume that $10 \leq h_1 \leq \sqrt{X}$. Combining Cauchy–Schwarz with [10, Theorem 9.2(ii)] (taking $\delta = (\log h_1)^{1/3} P_1^{-1/6+\eta}$, $\nu_1 = 1/20$ and $\nu_2 = 1/12$, there), we then get

$$
\mathcal{T}_1 \ll \left(\frac{(\log h_1)^{4/3}}{P_1^{1/12-\eta}} + (\log X)^{-1/200} \right)^{1/2} \ll (\log h_1)^{-1} + (\log X)^{-1/400}.
$$

Combining these estimates, we obtain that the LHS of (15) is

$$
\ll (\log h_1)^{-1} + (\log X)^{-1/400} + \frac{\log \log h_1}{\log h_1} + \frac{h_2}{X} \ll (\log X)^{-1/400} + \frac{\log \log h_1}{\log h_1},
$$

as claimed.   □

Proof of Theorem 1.1 Set $h_1 := h$ and $h_2 := X/(\log X)^{1/3}$. As mentioned, we may assume that g is real-valued (otherwise the result follows for complex-valued g by applying the theorem to $\mathrm{Re}(g)$ and $\mathrm{Im}(g)$ and applying the triangle inequality). By Lemma 4.1, we have

$$
\frac{2}{X} \sum_{X/2 < m \leq X} \left| \frac{1}{h_1} \sum_{m-h_1 < n \leq m} g(n) - \frac{2}{X} \sum_{X/2 < n \leq X} g(n) \right|
$$
$$
\ll \frac{2}{X} \sum_{X/2 < m \leq X} \left| \frac{1}{h_1} \sum_{m-h_1 < n \leq m} g(n) - \frac{1}{h_2} \sum_{m-h_2 < n \leq m} g(n) \right| + \frac{B_g(X)}{(\log X)^{1/6}}. \quad (20)
$$

By Lemma 4.2, the latter is

$$\ll \frac{B_g(X)}{t} \cdot \frac{1}{X} \int_{X/2}^{X} |S_{h_1}^{(\mathcal{S})}(t;x) - S_{h_2}^{(\mathcal{S})}(t;x)| \mathrm{d}x$$

$$+ B_g(X) \left(t + \frac{\log \log h}{t \log h} + (\log X)^{-\frac{1}{6}} \right). \tag{21}$$

Observe next that for any $x \in [X/2, X]$ and $t \in (0,1)$,

$$S_{h_1}^{(\mathcal{S})}(t;x) - S_{h_2}^{(\mathcal{S})}(t;x) = \frac{1}{h_1} \sum_{\substack{x-h_1 < n \le x \\ n \in \mathcal{S}}} e(t\tilde{g}(n;X)) - \frac{1}{h_2} \sum_{\substack{x-h_2 < n \le x \\ n \in \mathcal{S}}} e(t\tilde{g}(n;X))$$

$$= e(-t\tfrac{A_g(X)}{B_g(X)}) \cdot \left(\frac{1}{h_1} \sum_{\substack{x-h_1 < n \le x \\ n \in \mathcal{S}}} G_{t,X}(n) - \frac{1}{h_2} \sum_{\substack{x-h_2 < n \le x \\ n \in \mathcal{S}}} G_{t,X}(n) \right).$$

Taking $t := \max\left\{ \sqrt{\frac{\log \log h_1}{\log h_1}}, (\log X)^{-1/800} \right\}$, Theorem 1.1 now follows on combining this last expression with Lemma 4.3 and inserting the resulting bound into (21) . $\square$

5 The Matomäki–Radziwiłł theorem for additive functions: ℓ^2 variant

In this section, we will prove Theorem 1.4.
Let $g \in \mathcal{A}_s$, so that $B_g(X) \to \infty$, and the conditions

$$\lim_{\delta \to 0^+} F_g(\delta) = \lim_{\delta \to 0^+} \limsup_{X \to \infty} \frac{1}{B_g(X)^2} \sum_{\substack{p \le X \\ |g(p)| > \delta^{-1} B_g(X)}} \frac{|g(p)|^2}{p} = 0, \tag{22}$$

$$\limsup_{X \to \infty} \frac{1}{B_g(X)^2} \sum_{\substack{p^k \le X \\ k \ge 2}} \frac{|g(p^k)|^2}{p^k} = 0. \tag{23}$$

both hold. We seek to show that

$$\Delta_g(X,h) := \frac{2}{X} \sum_{X/2 < n \le X} \left| \frac{1}{h} \sum_{n-h < m \le n} g(m) - \frac{2}{X} \sum_{X/2 < m \le X} g(m) \right|^2 = o(B_g(X)^2), \tag{24}$$

whenever $10 \le h \le X/10$ is an integer that satisfies $h = h(X) \to \infty$ as $X \to \infty$. We begin by making the following convenient reduction.

Lemma 5.1 *Suppose that Theorem 1.4 holds for any non-negative, strongly additive function $g \in \mathcal{A}_s$. Then Theorem 1.4 holds for any $g \in \mathcal{A}_s$.*

Proof By splitting $g = \mathrm{Re}(g) + i\,\mathrm{Im}(g)$, and separately decomposing

$$\mathrm{Re}(g) = \mathrm{Re}(g)^+ - \mathrm{Re}(g)^-, \qquad \mathrm{Im}(g) = \mathrm{Im}(g)^+ - \mathrm{Im}(g)^-,$$

where, for an additive function h we define the non-negative additive functions $h^\pm$ on prime powers via

$$h^+(p^k) := \max\{h(p^k), 0\}, \quad h^-(p^k) := \max\{0, -h(p^k)\},$$

the Cauchy–Schwarz inequality implies that if (24) holds for non-negative g satisfying (22) and (23) then it holds for all additive g satisfying those conditions.

Therefore, we may assume that g is non-negative. Now, by Lemma 3.6, we can find a strongly additive function g^*, satisfying $g(p) = g^*(p)$ for all p, such that upon setting $G := g - g^*$ we have $B_G(X) = o(B_g(X))$. If we write, for a non-negative additive function h,

$$\tilde{F}_h(\delta; X) := \frac{1}{B_h(X)^2} \sum_{\substack{p \le X \\ h(p) > \delta^{-1} B_h(X)}} \frac{h(p)^2}{p}$$

then we see that when X is large enough,

$$\tilde{F}_{g^*}(\delta/2; X) \le 2\tilde{F}_g(\delta; X) \le 4\tilde{F}_{g^*}(2\delta; X).$$

Taking limsups as $X \to \infty$ in these inequalities, it follows that g^* satisfies (22) whenever g does; that g^* also satisfies (23) is an immediate consequence of Lemma 3.6. Moreover, we see by the Cauchy–Schwarz inequality and Lemma 3.2 that

$$\frac{2}{X} \sum_{X/2 < n \le X} \left| \frac{1}{h} \sum_{n-h < m \le n} (G(m) - A_G(X)) \right|^2$$

$$\le \frac{1}{Xh} \sum_{X/2 < n \le X} \sum_{n-h < m \le n} |G(m) - A_G(X)|^2$$

$$\ll \frac{1}{X} \sum_{X/3 < m \le X} |G(m) - A_G(X)|^2 \ll B_G(X)^2 = o(B_g(X)^2).$$

Using the estimate $\frac{2}{X} \sum_{X/2 < n \le X} G(n) = A_G(X) + o(B_g(X))$ by Lemmas 3.1 and 3.4 (as in the proof of Lemma 4.1), we see that

$$\Delta_g(X, h) \ll \Delta_{g^*}(X, h) + \Delta_G(X, h) = \Delta_{g^*}(X, h) + o(B_g(X)),$$

so that if (24) holds for strongly additive $g^* \in \mathcal{A}_s$ then it also holds for all $g \in \mathcal{A}_s$. This completes the proof. $\qquad\square$

Until further notice we may thus assume that g is non-negative and strongly additive. For a fixed small parameter $\varepsilon > 0$, let $\delta \in (0, 1/100)$ be chosen such that $F_g(\delta) < \varepsilon$. Let X be a scale chosen sufficiently large so that

$$\sum_{\substack{p \leq X \\ g(p) > \delta^{-1} B_g(X)}} \frac{g(p)^2}{p} \leq 2 F_g(\delta) B_g(X)^2 < 2\varepsilon B_g(X)^2, \tag{25}$$

$$\sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{g(p^k)^2}{p^k} = \sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{g(p)^2}{p^k} < \varepsilon B_g(X)^2. \tag{26}$$

With these data, define

$$\mathcal{C} = \mathcal{C}(X, \delta) := \{p \leq X : g(p) \leq \delta^{-1} B_g(X)\}.$$

We decompose g as

$$g = g_{\mathcal{C}} + g_{\mathcal{P} \setminus \mathcal{C}},$$

where $g_{\mathcal{C}}$ and $g_{\mathcal{P} \setminus \mathcal{C}}$ are strongly additive functions defined at primes by

$$g_{\mathcal{C}}(p) := \begin{cases} g(p) & \text{if } p \in \mathcal{C} \\ 0 & \text{if } p \notin \mathcal{C}. \end{cases} \qquad g_{\mathcal{P} \setminus \mathcal{C}}(p) := \begin{cases} 0 & \text{if } p \in \mathcal{C} \\ g(p) & \text{if } p \notin \mathcal{C}. \end{cases}$$

We will consider the mean-squared errors

$$\frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h} \sum_{n-h < m \leq n} g_{\mathcal{A}}(m) - \frac{2}{X} \sum_{X/2 < m \leq X} g_{\mathcal{A}}(m) \right|^2$$

for $\mathcal{A} \in \{\mathcal{C}, \mathcal{P} \setminus \mathcal{C}\}$, separately. The fact that $g \in \mathcal{A}_s$ means, in particular, that $g_{\mathcal{P} \setminus \mathcal{C}}$ contributes little to $\Delta_g(X, h)$.

Lemma 5.2 *Let $g \in \mathcal{A}_s$ be a non-negative, strongly additive function. Assume that X and δ are chosen such that (25) and (26) both hold. Then we have*

$$\Delta_g(X, h) \ll \Delta_{g_{\mathcal{C}}}(X, h) + \varepsilon B_g(X)^2,$$

where $\mathcal{C} = \mathcal{C}(X, \delta)$.

Proof Arguing as in the proof of Lemma 5.1, we obtain from (25) and (26) that

$$\Delta_{g_{\mathcal{P} \setminus \mathcal{C}}}(X, h) \ll B_{g_{\mathcal{P} \setminus \mathcal{C}}}(X)^2 < 3\varepsilon B_g(X)^2.$$

Thus, by the Cauchy–Schwarz inequality, we obtain

$$\Delta_g(X, h) \ll \Delta_{g_C}(X, h) + \Delta_{g_{\mathcal{P} \setminus C}}(X, h) \ll \Delta_{g_C}(X, h) + \varepsilon B_g(X)^2,$$

as claimed. $\square$

In analogy to the work of the previous section, we will reduce the estimation of $\Delta_{g_C}(X)$ to that of the variance of short- and long-interval averages of certain multiplicative functions determined by g_C. These are defined as follows.
Fix $r \in (0, \delta^2]$. Given $z \in \mathbb{C}$ satisfying $|z - 1| = r$, define

$$F_z(n) := z^{g_C(n)/B_g(X)} \text{ for all } n \in \mathbb{N}.$$

Since g_C is strongly additive and satisfies $0 \le g_C(p) \le \delta^{-1} B_g(X)$ for all $p \le X$, we have

$$|F_z(p^k)| = |F_z(p)| \le (1 + \delta^2)^{\delta^{-1}} \le 2 \tag{27}$$

for all $p^k \le X$, and thus also

$$|F_z(n)| \le \left(\max_{p|n} |F_z(p)| \right)^{\omega(n)} \le 2^{\omega(n)} \le d(n), \tag{28}$$

where $d(n)$ is the divisor function. Furthermore, as $\delta \in (0, 1/100)$, for any $2 \le u \le v \le X$ we get

$$\sum_{u < p \le v} \frac{|F_z(p)|}{p} \ge (1 - \delta^2)^{\delta^{-1}} \sum_{u < p \le v} \frac{1}{p} \ge 0.98 \sum_{u < p \le v} \frac{1}{p}. \tag{29}$$

Our treatment of short sums of g_C will entail an analysis of corresponding short sums of F_z, for z lying in a small neighbourhood of 1. In preparation to apply a relevant result from the recent paper [14], we introduce some further notation. Given a multiplicative function $f : \mathbb{N} \to \mathbb{C}$ set

$$H(f; X) := \prod_{p \le X} \left(1 + \frac{(|f(p)| - 1)^2}{p} \right), \qquad \mathcal{P}_f(X) := \prod_{p \le X} \left(1 + \frac{|f(p)| - 1}{p} \right).$$

We also define the following variant of the pretentious distance:

$$\rho(f, n^{it}; X)^2 := \sum_{p \le X} \frac{|f(p)| - \mathrm{Re}(f(p)p^{-it})}{p}.$$

We let $t_0 = t_0(f, X)$ denote a real number $t \in [-X, X]$ that minimizes $t \mapsto \rho(f, n^{it}; X)^2$.

Theorem 5.3 [14, Thm. 2.1] *Let $0 < A \leq 2$, and let X be large. Let $f : \mathbb{N} \to \mathbb{C}$ be a multiplicative function that satisfies*

(i) *$|f(n)| \leq d(n)$ for all $n \leq X$, and in particular $|f(p)| \leq 2$ for all $p \leq X$,*
(ii) *for any $2 \leq u \leq v \leq X$,*

$$\sum_{u < p \leq v} \frac{|f(p)|}{p} \geq A \sum_{u < p \leq v} \frac{1}{p} - O\left(\frac{1}{\log u}\right).$$

Let $10 \leq h_0 \leq X/(10H(f; X))$, and put $h_1 := h_0 H(f; X)$ and $t_0 = t_0(f, X)$. Then there are constants $c_1, c_2 \in (0, 1/3)$, depending only on A, such that if $X/(\log X)^{c_1} < h_2 \leq X$,

$$\frac{2}{X} \int_{X/2}^{X} \left| \frac{1}{h_1} \sum_{x - h_1 < n \leq x} f(n) - \frac{1}{h_1} \int_{x-h_1}^{x} u^{it_0} du \cdot \frac{1}{h_2} \sum_{x - h_2 < n \leq x} f(n) n^{-it_0} \right|^2 dx$$

$$\ll_A \left(\left(\frac{\log \log h_0}{\log h_0}\right)^A + \left(\frac{\log \log X}{(\log X)^{c_2}}\right)^{\min\{1,A\}} \right) \mathcal{P}_f(X)^2.$$

Let $c_1 \in (0, 1/3)$ be the constant from Theorem 5.3, applied with $A = 0.98$. By Lemma 4.1, if $h_2 = \lceil X/(\log X)^{c_1} \rceil$ then for any $x \in (X/2, X]$

$$\frac{1}{h_2} \sum_{x - h_2 < n \leq x} g_{\mathcal{C}}(n) = \frac{2}{X} \sum_{X/2 < n \leq X} g_{\mathcal{C}}(n) + O\left(\frac{B_g(X)}{(\log X)^{1/6}}\right). \tag{30}$$

In view of Lemma 5.2, in order to prove Theorem 1.4 it suffices to show that as $X \to \infty$,

$$\Delta_{g_{\mathcal{C}}}(X; h_1, h_2) := \frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h_1} \sum_{n - h_1 < m \leq x} g_{\mathcal{C}}(m) - \frac{1}{h_2} \sum_{n - h_2 < m \leq n} g_{\mathcal{C}}(m) \right|^2$$

$$= o(B_g(X)^2),$$

where $h_1 = h$ and $h_2 = \lceil X/(\log X)^{c_1} \rceil$.
Using Theorem 5.3, we will prove the following.

Corollary 5.4 *Let $10 \leq h_1 \leq X/10$ be an integer and $h_2 := \lceil X/(\log X)^{c_1} \rceil$ as above. Then there is a constant $\gamma > 0$ such that*

$$\Delta_{g_{\mathcal{C}}}(X; h_1, h_2) \ll \delta^{-4} \left(\left(\frac{\log \log h}{\log h}\right)^{0.98} + (\log X)^{-\gamma} \right) B_g(X)^2.$$

The conditions (i) and (ii) of Theorem 5.3 were verified for $f = F_z$ in (27), (28) and (29) (with $A = 0.98$), and it remains to elucidate information about $t_0(F_z, X)$,

$H(F_z; X)$ and the size of the Euler product $\mathcal{P}_{F_z}(X)$. This is provided by the following lemma.

Lemma 5.5 *Fix* $r \in (0, \delta^2]$ *and let* $z \in \mathbb{C}$ *satisfy* $|z - 1| = r$. *Then*

(a) $t_0(F_z, X) \ll 1/\log X$,
(b) $H(F_z; X) \asymp 1$ *and*
(c) $\mathcal{P}_{F_z}(X)^2 \ll \prod_{p \leq X} \left(1 + \frac{|F_z(p)|^2 - 1}{p} \right)$.

Proof (a) Applying [14, (7)] with $A = 0.98$, $B = 2$ and $C = 1$ (which is a straight-forward consequence of [10, Lem. 5.1(i)]), we see that if $(\log X)|t_0(F_z; X)| \geq D$ for a suitably large constant $D > 0$ then by minimality of t_0,

$$\rho(F_z, 1; X)^2 \geq \sigma \min\{\log \log X, 3 \log(|t_0(F_z; X)| \log X + 1)\} + O(1) \geq 100,$$

say, where $\sigma > 0$ is an absolute constant.

To obtain a contradiction, observe next that for any $z = re(\theta)$ with $\theta \in [0, 1]$, we have already shown that $|F_z(p)| \leq 2$ for $\delta \in (0, 1/100)$. Thus, writing $F_z(p) = |F_z(p)|e(\theta g_{\mathcal{C}}(p)/B_g(X))$ and applying the inequality $0 \leq 1 - \cos x \leq x^2/2$ for all $x \geq 0$, we find

$$\rho(F_z, 1; X)^2 = \sum_{p \leq X} |F_z(p)| \frac{1 - \cos(2\pi \theta g_{\mathcal{C}}(p)/B_g(X))}{p}$$
$$\leq \frac{4\pi^2}{B_g(X)^2} \sum_{p \leq X} \frac{g(p)^2}{p} \leq 4\pi^2.$$

This contradiction implies that $|t_0(F_z; X)| \log X \leq D$ for some constant D, and the claim follows.

(b) By Taylor expansion, $|z|^{g_{\mathcal{C}}(p)/B_g(X)} = 1 + O(\delta g(p)/B_g(X))$, and thus

$$H(F_z; X) \ll \exp \left(\sum_{p \leq X} \frac{(|z|^{g_{\mathcal{C}}(p)/B_g(X)} - 1)^2}{p} \right)$$
$$= \exp \left(O \left(\frac{\delta^2}{B_g(X)^2} \sum_{p \leq X} \frac{g(p)^2}{p} \right) \right) \ll 1.$$

The corresponding lower bound is trivial from the definition of $H(F_z; X)$.

(c) Since $|F_z(p)| \leq 2$ for all $p \leq X$, we have the upper bounds

$$\mathcal{P}_{F_z}(X)^2 \ll \prod_{p \leq X} \left(1 + \frac{2(|F_z(p)| - 1)}{p} \right) \leq \prod_{p \leq X} \left(1 + \frac{|F_z(p)|^2 - 1}{p} \right),$$

the latter of which arises from $(|F_z(p)| - 1)^2 \geq 0$ for all p. $\qquad\square$

Lemma 5.6 *Let g be non-negative and strongly additive. Let $r \in (0, \delta^2]$, and set $h_1 = h$ and $h_2 = \lceil X/(\log X)^{c_1} \rceil$ as above. Then there is $z_0 \in \mathbb{C}$ with $|z_0 - 1| = r$ such that as $X \to \infty$,*

$$
\Delta_{g_{\mathcal{C}}}(X; h_1, h_2)
$$
$$
\ll \frac{B_g(X)^2}{r^2} |z_0|^{-2\frac{A_g(X)}{B_g(X)}}
$$
$$
\cdot \frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h_1} \sum_{n-h_1 < m \leq n} F_{z_0}(m) - I(n; t_0, h_1) \cdot \frac{1}{h_2} \sum_{n-h_2 < m \leq n} F_{z_0}(m) m^{-it_0} \right|^2
$$
$$
+ \frac{B_g(X)^2}{r^2} (\log X)^{-c_1 + o(1)},
$$

where $t_0 = t_0(F_{z_0}, X)$ and $I(x; t, h) := \frac{1}{h} \int_{x-h}^{x} u^{it} \, du$ as in the previous section.

Proof For each $n \in (X/2, X]$, $z \in \mathbb{C}$ and $j = 1, 2$, define the maps

$$
\phi_n(z; h_j) := \frac{1}{h_j} \sum_{n-h_j < m \leq n} z^{(g_{\mathcal{C}}(m) - A_{g_{\mathcal{C}}}(X))/B_g(X)},
$$

which are analytic in z. Note that

$$
\frac{1}{h_j} \sum_{n-h_j < m \leq n} \left(\frac{g_{\mathcal{C}}(m) - A_{g_{\mathcal{C}}}(X)}{B_g(X)} \right) = \frac{\mathrm{d}}{\mathrm{d}z} \phi_n(z; h_j) \Big|_{z=1}.
$$

Recall that $h_1, h_2 \in \mathbb{Z}$. Thus, by Cauchy's integral formula we have

$$
\frac{1}{h_1} \sum_{n-h_1 < m \leq n} g_{\mathcal{C}}(m) - \frac{1}{h_2} \sum_{n-h_2 < m \leq n} g_{\mathcal{C}}(m)
$$
$$
= \frac{1}{h_1} \sum_{n-h_1 < m \leq n} (g_{\mathcal{C}}(m) - A_{g_{\mathcal{C}}}(X)) - \frac{1}{h_2} \sum_{n-h_2 < m \leq n} (g_{\mathcal{C}}(m) - A_{g_{\mathcal{C}}}(X))
$$
$$
= \frac{B_g(X)}{2\pi i} \int_{|z-1|=r} (\phi_n(z; h_1) - \phi_n(z; h_2)) \frac{\mathrm{d}z}{(z-1)^2}.
$$

By Cauchy–Schwarz and the definition of F_z, we obtain

$$
\Delta_{g_{\mathcal{C}}}(X; h_1, h_2) \ll \frac{B_g(X)^2}{r^2} \max_{|z-1|=r} \frac{2}{X} \sum_{X/2 < n \leq X} |\phi_n(z; h_1) - \phi_n(z; h_2)|^2
$$
$$
= \frac{B_g(X)^2}{r^2} |z_0|^{-2\frac{A_{g_{\mathcal{C}}}(X)}{B_g(X)}} \frac{2}{X} \sum_{X/2 < n \leq X} \left| \frac{1}{h_1} \sum_{n-h_1 < m \leq n} F_{z_0}(m) - \frac{1}{h_2} \sum_{n-h_2 < m \leq n} F_{z_0}(m) \right|^2,
$$

for some $z_0 \in \mathbb{C}$ with $|z_0 - 1| = r$. To complete the proof, note that by Taylor expansion and Lemma 5.5(a),

$$\frac{1}{h_1} \int_{n-h_1}^{n} u^{it_0} \mathrm{d}u = n^{it_0} \frac{1 - (1 - h_1/n)^{1+it_0}}{(1 + it_0)h_1/n} = n^{it_0} + O(h_1/X),$$

and also

$$m^{-it_0} = n^{-it_0} + O(|t_0| \log(n/m)) = n^{-it_0} + O(h_2/X)$$

uniformly in $n - h_2 < m \leq n$. It follows that

$$\frac{1}{h_2} \sum_{n-h_2 < m \leq n} F_{z_0}(m) = I(n; t_0, h_1) \cdot \frac{1}{h_2} \sum_{n-h_2 < m \leq n} F_{z_0}(m) m^{-it_0}$$

$$+ O\left(\frac{1}{X} \sum_{n-h_2 < m \leq n} |F_{z_0}(m)|\right).$$

The error term is, by Shiu's theorem [35, Thm. 1] and the Cauchy–Schwarz inequality,

$$\ll \frac{h_2}{X} \exp\left(\sum_{p \leq X} \frac{|F_{z_0}(p)| - 1}{p}\right) \ll \frac{h_2}{X} \exp\left(\frac{2r}{B_g(X)} \sum_{p \leq X} \frac{g(p)}{p}\right)$$

$$\ll \frac{h_2}{X} \exp\left(2\sqrt{\log \log X}\right),$$

which suffices to prove the claim.   □

We are now in a position to apply Theorem 5.3 in order to prove Corollary 5.4.

Proof of Corollary 5.4 Let z_0 be chosen as in Lemma 5.6. Since $h_1, h_2 \in \mathbb{Z}$ we may replace the discrete average in Lemma 5.6 by an integral average at the cost of an error term of size

$$\ll \frac{B_g(X)^2}{r^2} \max_{x \in [X/2, X]} \left(|I(x; t_0, h_1) - I(\lfloor x \rfloor; t_0, h_1)| \cdot \frac{1}{h_2} \sum_{\lfloor x \rfloor - h_2 < n \leq \lfloor x \rfloor} |F_{z_0}(n)|\right)^2$$

$$\ll \frac{B_g(X)^2}{h_1^2 r^2} |z_0|^{-2\frac{A g_{\mathcal{C}}(X)}{B_g(X)}} \mathcal{P}_{F_{z_0}}(X)^2,$$

again by Shiu's bound [35, Thm. 1]. Using the data from Lemma 5.5, Theorem 5.3 therefore yields

$$\Delta_{g_{\mathcal{C}}}(X; h_1, h_2) \ll \frac{B_g(X)^2}{\delta^4} \left(\left(\frac{\log\log h_1}{\log h_1} \right)^{0.98} + \left(\frac{\log\log X}{(\log X)^{c_2}} \right)^{0.98} \right) \tag{31}$$

$$\cdot |z_0|^{-2\frac{A_{g_{\mathcal{C}}}(X)}{B_g(X)}} \prod_{p \leq X} \left(1 + \frac{|F_{z_0}(p)|^2 - 1}{p} \right). \tag{32}$$

As g is strongly additive,

$$A_{g_{\mathcal{C}}}(X) = \sum_{p \leq X} \frac{g_{\mathcal{C}}(p)}{p} + O\left(\sum_{p \leq X} \frac{g(p)}{p^2} \right) = \sum_{p \leq X} \frac{g_{\mathcal{C}}(p)}{p} + O(B_g(X)),$$

the bound in the error term arising from the Cauchy–Schwarz inequality. Now put $\rho := \log|z_0| \in (-10\delta^2, 10\delta^2)$, say. Using the estimates $\log(1 + x) = x + O(x^2)$ and $|e^x - 1 - x| \leq \frac{1}{2}|x|^2$ for $|x| \leq 1/2$, the factors in (32) can be estimated as

$$\ll \exp\left(\sum_{p \leq X} \left(-\frac{2\rho}{p} \frac{g_{\mathcal{C}}(p)}{B_g(X)} + \log\left(1 + \frac{e^{2\rho g_{\mathcal{C}}(p)/B_g(X)} - 1}{p} \right) \right) \right)$$

$$\ll \exp\left(\sum_{p \leq X} \frac{1}{p} \left(e^{2\rho g_{\mathcal{C}}(p)/B_g(X)} - 1 - 2\rho \frac{g_{\mathcal{C}}(p)}{B_g(X)} \right) \right)$$

$$\leq \exp\left(\frac{2\rho^2}{B_g(X)^2} \sum_{p \leq X} \frac{g_{\mathcal{C}}(p)^2}{p} \right) \ll 1.$$

The claimed bound now follows with any $0 < \gamma < 0.98c_2$ (changing the implicit constant as needed). □

Proof of Theorem 1.4 Let $g \in \mathcal{A}_s$. By Lemma 5.1 we may assume that g is non-negative and strongly additive. Let $\varepsilon > 0$ and pick $\delta > 0$ and $X_0 = X_0(\delta)$ such that if $X \geq X_0$ then (25) and (26) both hold, and also define $\mathcal{C}$ as above. Set also $h_1 = h$ and $h_2 := \left\lceil \frac{X}{(\log X)^{c_1}} \right\rceil$ as above. Combining Lemma 5.2 and (30), we have

$$\Delta_g(X, h) \ll \Delta_{g_{\mathcal{C}}}(X, h) + \varepsilon B_g(X)^2 \ll \Delta_{g_{\mathcal{C}}}(X; h_1, h_2) + \left(\varepsilon + (\log X)^{-1/6} \right) B_g(X)^2.$$

Applying Corollary 5.4 in this estimate, we find that there is a $\gamma \in (0, 1/6)$ for which

$$\Delta_g(X, h) \ll \left(\varepsilon + \delta^{-4} \left(\left(\frac{\log\log h}{\log h} \right)^{0.98} + (\log X)^{-\gamma} \right) \right) B_g(X)^2.$$

Selecting $h \geq \exp\left(\delta^{-5}\varepsilon^{-2} \log(1/(\delta\varepsilon)) \right)$, picking X_0 larger if necessary, we deduce that $\Delta_g(X, h) \ll \varepsilon B_g(X)$. Since ε was arbitrary, we deduce that $\Delta_g(X, h) = o(B_g(X))$, as claimed. □

6 Gaps and moments

In this section, we will prove Theorem 1.11.

6.1 Small gaps and small first moments are equivalent: proof of Theorem 1.11(a)

We start by proving the following quantitative ℓ^1 gap result.

Proposition 6.1 *Let $0 < \varepsilon < 1/3$ and let X be large. Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function. Assume that*

$$\frac{1}{Y} \sum_{n \leq Y} |g(n) - g(n-1)| \ll \varepsilon B_g(X) \tag{33}$$

for all $X/\log X < Y \leq X$. Then we have

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)| \ll \left(\sqrt{\frac{\log \log(1/\varepsilon)}{\log(1/\varepsilon)}} + (\log X)^{-1/800} \right) B_g(X).$$

Proof Let $h = \lfloor \min\{X/(2 \log X), \varepsilon^{-1/2}\} \rfloor$, and let $X/\log X < Y \leq X$. By the triangle inequality and (33), for any $1 \leq m \leq h$ we obtain

$$\frac{1}{Y} \sum_{h < n \leq Y} |g(n) - g(n-m)| \leq \frac{1}{Y} \sum_{0 \leq j \leq m-1} \sum_{j < n \leq Y} |g(n-j) - g(n-j-1)|$$

$$\ll \frac{h}{Y} \sum_{1 \leq n \leq Y} |g(n) - g(n-1)| \ll \varepsilon^{1/2} B_g(X).$$

Averaging over $1 \leq m \leq h$ and then applying the triangle inequality once again, we obtain

$$\frac{1}{Y} \sum_{Y/2 < n \leq Y} \left| g(n) - \frac{1}{h} \sum_{1 \leq m \leq h} g(n-m) \right|$$

$$\leq \frac{1}{Y} \sum_{h < n \leq Y} \left| g(n) - \frac{1}{h} \sum_{1 \leq m \leq h} g(n-m) \right| \ll h^{-1} B_g(X).$$

Applying Theorem 1.1, we deduce from this that

$$\frac{1}{Y} \sum_{Y/2 < n \leq Y} \left| g(n) - \frac{2}{Y} \sum_{Y/2 < m \leq Y} g(m) \right| \ll B_g(X) \left(\sqrt{\frac{\log \log h}{\log h}} + (\log X)^{-1/800} \right).$$

Now, for each such Y, Lemmas 3.1 and 3.4 combine to yield

$$\frac{2}{Y} \sum_{Y/2 < m \le Y} g(m) = 2A_g(Y) - A_g(Y/2) + O\left(\frac{B_g(Y)}{\sqrt{\log Y}}\right)$$

$$= A_g(X) + O\left(B_g(X)\sqrt{\frac{\log \log X}{\log X}}\right),$$

and thus combining this estimate into the previous expression and summing over all dyadic subintervals of $[X/\log X, X]$, we obtain

$$\frac{1}{X} \sum_{\frac{X}{\log X} < n \le X} |g(n) - A_g(X)| \le \sum_{\frac{X}{\log X} < Y = X/2^j \le X} \frac{Y}{2X} \cdot \frac{2}{Y} \sum_{Y/2 < n \le Y} |g(n) - A_g(X)|$$

$$\ll B_g(X)\left(\sqrt{\frac{\log \log h}{\log h}} + (\log X)^{-1/800}\right).$$

Applying Lemma 3.2 and the Cauchy–Schwarz inequality on $[1, X/\log X]$, we obtain

$$\frac{1}{X} \sum_{1 \le n \le \frac{X}{\log X}} |g(n) - A_g(X)| \le \frac{1}{\sqrt{\log X}} \left(\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)|^2\right)^{\frac{1}{2}} \ll \frac{B_g(X)}{\sqrt{\log X}}.$$

The latter two estimates together imply the claim. $\qquad\square$

Proof of Theorem 1.11(a) By the triangle inequality, we see that if

$$\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)| = o(B_g(X))$$

then, consequently,

$$\frac{1}{X} \sum_{n \le X} |g(n) - g(n-1)| \le \frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)| + \frac{1}{X} \sum_{m \le X-1} |g(m) - A_g(X)|$$

$$= o(B_g(X)).$$

The converse implication follows immediately from Proposition 6.1. $\qquad\square$

6.2 A gap theorem for the second moment

In parallel to the results of the previous subsection, we will apply Theorem 1.4 to prove the following result.

Proposition 6.2 *Let $g \in \mathcal{A}_s$ Then for any integer $10 \leq h \leq \frac{X}{10 \log X}$ we have*

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 \ll \frac{h^2}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 + o_{h \to \infty}(B_g(X)^2).$$

Proof Given our assumptions about g, we may apply Theorem 1.4 to obtain

$$\sum_{Y/2 < n \leq Y} \left| \frac{1}{h} \sum_{n-h < m \leq n} g(m) - \frac{2}{Y} \sum_{Y/2 < m \leq Y} g(m) \right|^2 = o_{h \to \infty}(Y B_g(X)^2),$$

for any $X/\log X < Y \leq X$. Applying Lemmas 3.1 and 3.4, we deduce that

$$\frac{1}{X} \sum_{Y/2 < n \leq Y} |g(n) - A_g(Y)|^2$$

$$\ll \frac{1}{X} \sum_{Y/2 < n \leq Y} \left| g(n) - \frac{1}{h} \sum_{n-h < m \leq n} g(m) \right|^2 + o_{h \to \infty} \left(\frac{Y}{X} B_g(X)^2 \right).$$

By telescoping,

$$g(n) - \frac{1}{h} \sum_{n-h < m \leq n} g(m) = \frac{1}{h} \sum_{0 \leq j \leq h-1} (g(n) - g(n-j))$$

$$= \sum_{0 \leq j \leq h-2} \left(1 - \frac{j+1}{h} \right) (g(n-j) - g(n-j-1)).$$

Squaring both sides and applying Cauchy–Schwarz, we obtain

$$\frac{1}{X} \sum_{Y/2 < n \leq Y} \left| g(n) - \frac{1}{h} \sum_{n-h < m \leq n} g(m) \right|^2$$

$$\ll \frac{h}{X} \sum_{0 \leq j \leq h-2} \sum_{Y/2 < n \leq Y} (g(n-j) - g(n-j-1))^2$$

$$\leq \frac{h^2}{X} \sum_{\frac{Y}{2} - h < n \leq Y < n \leq Y} (g(n) - g(n-1))^2.$$

Combined with the previous estimates, we obtain

$$\frac{1}{X} \sum_{Y/2 < n \leq Y} |g(n) - A_g(Y)|^2$$

$$\ll \frac{h^2}{X} \sum_{\frac{Y}{2} - h < n \leq Y} |g(n) - g(n-1)|^2 + o_{h \to \infty} \left(\frac{Y}{X} B_g(X)^2 \right).$$

By Lemma 3.4, for each $X/\log X < Y \leq X$ we have

$$|A_g(X) - A_g(Y)| \ll B_g(X) \sqrt{\frac{\log \log X}{\log X}},$$

so that, summing over dyadic subintervals of $[X/\log X, X]$ and noting that by our assumption $h \leq X/(10 \log X)$ at most two dyadic intervals contain any point of

$$\bigcup_{\frac{X}{\log X} < Y = X/2^j \leq X} [Y/2 - h, Y],$$

we find

$$\frac{1}{X} \sum_{\frac{X}{\log X} < n \leq X} |g(n) - A_g(X)|^2$$

$$\ll \sum_{\frac{X}{\log X} < Y = X/2^j \leq X} \frac{1}{X} \sum_{Y/2 < n \leq Y} \left(|g(n) - A_g(Y)|^2 + B_g(X)^2 \cdot \frac{\log \log X}{\log X} \right)$$

$$\ll \frac{h^2}{X} \sum_{\frac{X}{4 \log X} < n \leq X} |g(n) - g(n-1)|^2 + o_{h \to \infty}(B_g(X)^2).$$

Applying Lemma 3.2 trivially to the segment $[1, X/\log X]$ gives

$$\frac{1}{X} \sum_{n \leq \frac{X}{\log X}} |g(n) - A_g(X)|^2 \ll \frac{B_g(X)^2}{\log X},$$

so combining these two estimates implies the claim. □

Proof of Theorem 1.11(b) To obtain the theorem, we note first the trivial estimate

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 \ll \frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 + \frac{1}{X} \sum_{m \leq X-1} |g(m) - A_g(X)|^2$$

$$\ll \frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2,$$

so that if the RHS is $o(B_g(X)^2)$ then so is the LHS.

Conversely, suppose that

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 \leq \xi(X) B_g(X)^2,$$

for some function $\xi(Y) \to 0$ as $Y \to \infty$. Set $h := \lfloor \xi(X)^{-1/3} \rfloor$. By Proposition 6.2,

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 \ll \xi(X)^{-2/3} \cdot \xi(X) B_g(X)^2 + o(B_g(X)^2) = o(B_g(X)^2)$$

as $X \to \infty$, as required. $\square$

Proof of Corollary 1.12 Let $g \in \mathcal{A}_s$ and suppose that

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)|^2 = o(B_g(X)^2).$$

By Theorem 1.11(b), we obtain that

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 = o(B_g(X)^2).$$

Now, by Lemma 3.3, this implies that there is $\lambda = \lambda(X) \in \mathbb{R}$ such that

$$B_{g_\lambda}(X)^2 + |\lambda|^2 = o(B_g(X)^2),$$

where $g_\lambda = g - \lambda \log$. Since the left-hand side is $\geq B_{g_\lambda}(X)^2$, the claim follows immediately. $\square$

7 Erdős' almost everywhere monotonicity problem

Let $g : \mathbb{N} \to \mathbb{R}$ be additive. For convenience, set $g(0) := 0$, and recall the definitions

$$\mathcal{B} := \{n \in \mathbb{N} : g(n) < g(n-1)\}, \qquad \mathcal{B}(X) := \mathcal{B} \cap [1, X].$$

In this and the following section, we will study functions g such that $|\mathcal{B}(X)| = o(X)$.

7.1 The second moment along sparse subsets

To prove Theorem 1.8 we will eventually need control over a sparsely-supported sum such as

$$\frac{1}{X} \sum_{n \in \mathcal{B}(X)} |g(n) - A_g(X)|^2,$$

with the objective of obtaining savings over the trivial bound $O(B_g(X)^2)$ from Lemma 3.2. The purpose of this subsection is to determine sufficient conditions in order to achieve a non-trivial estimate of this kind.

Given a set of positive integers $\mathcal{S}$, a positive real number $X \geq 1$ and a prime power $p^k \leq X$, write $\mathcal{S}(X) := \mathcal{S} \cap [1, X]$ and $\mathcal{S}_{p^k}(X) := \{n \in \mathcal{S}(X) : p^k \mid n\}$.

Proposition 7.1 *Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function belonging to $\mathcal{A}$. Let $\mathcal{S}$ be a set of integers with $|\mathcal{S}(X)| = o(X)$, and let $\varepsilon \in (0, 1)$ satisfy the conditions*

$$|\mathcal{S}(X)| < \tfrac{\varepsilon}{2} X, \qquad \sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{|g(p)|^2 + |g(p^k)|^2}{p^k} < \varepsilon B_g(X)^2.$$

Then the following bound holds:

$$\frac{1}{X} \sum_{n \in \mathcal{S}(X)} |g(n) - A_g(X)|^2 \ll B_g(X)^2 \left(\varepsilon + \varepsilon^{-1} \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2} \right)$$

$$+ \sum_{\substack{p \leq X \\ |\mathcal{S}_p(X)| > \varepsilon X / p}} \frac{|g(p)|^2}{p}.$$

Moreover, we have

$$\sum_{\substack{p \leq X \\ |\mathcal{S}_p(X)| > \varepsilon X / p}} \frac{1}{p} \ll \varepsilon^{-2} \frac{|\mathcal{S}(X)|}{X}. \tag{34}$$

Remark 7.2 Proposition 7.1 states that if the bulk of the contribution to the variance of $g(n)$ occurs along a sparse subset $\mathcal{S}(X) \subseteq [1, X]$ then $B_g(X)$ is dominated by primes p of which $\mathcal{S}(X)$ has many multiples $\leq X$. For sufficiently small primes p this is ruled out by the sparseness of $\mathcal{S}(X)$, but it may occur for large enough primes.

Our proof will proceed by applying variants of the large sieve and Turán–Kubilius inequalities. The first of these is due to Elliott.

Lemma 7.3 (Elliott's Dual Turán–Kubilius Inequality) *Let $\{a(n)\}_n \subset \mathbb{C}$ be a sequence and let $X \geq 2$. Then*

$$\sum_{p \leq X} p \left| \sum_{\substack{n \leq X \\ p \mid n}} a(n) - \frac{1}{p} \sum_{n \leq X} a(n) \right|^2 \ll X \sum_{n \leq X} |a(n)|^2.$$

Proof This is [22, Lemma 5.2] (taking $\sigma = 0$ there). □

A variant of the latter result, for divisibility by products of two large primes, is as follows.

Lemma 7.4 (Variant of Dual Turán–Kubilius) *Let* $\{a(n)\}_n \subset \mathbb{C}$. *Then*

$$\sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} pq \left| \sum_{\substack{n \leq X \\ pq|n}} a(n) - \frac{1}{pq} \sum_{n \leq X} a(n) \right|^2 \ll X \sum_{n \leq X} |a(n)|^2.$$

Proof By including the factor pq into the square, we observe that this establishes an $\ell^2 \to \ell^2$ operator norm for the matrix with entries

$$M = \left((pq)^{1/2} 1_{pq|n} - (pq)^{-1/2} \right)_{\substack{X^{1/4} < p,q \leq X, p \neq q \\ n \leq X}}.$$

Thus, by the duality principle [31, Sec. 7.1] it suffices to show that for any sequence $\{b(p,q)\}_{p,q \text{ prime}} \subset \mathbb{C}$ we have

$$\sum_{n \leq X} \left| \sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} \frac{b(p,q)}{pq} (pq 1_{pq|n} - 1) \right|^2 \ll X \sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} \frac{|b(p,q)|^2}{pq}. \quad (35)$$

Expanding the square on the LHS and swapping orders of summation, we obtain

$$\sum_{\substack{X^{1/4} < p_1,p_2,q_1,q_2 \leq X \\ p_j \neq q_j \\ j=1,2}} \frac{b(p_1,q_1)\overline{b}(p_2,q_2)}{p_1 q_1 p_2 q_2}$$

$$\cdot \sum_{n \leq X} \left(p_1 q_1 p_2 q_2 1_{p_1 q_1|n} 1_{p_2 q_2|n} - p_1 q_1 1_{p_1 q_1|n} - p_2 q_2 1_{p_2 q_2|n} + 1 \right).$$

Fix the quadruple (p_1, q_1, p_2, q_2) for the moment, and consider the inner sum over $n \leq X$. If $(p_1 q_1, p_2 q_2) = 1$ then as $p_1 q_1 p_2 q_2 > X$ the sum is

$$\ll p_1 q_1 \left\lfloor \frac{X}{p_1 q_1} \right\rfloor + p_2 q_2 \left\lfloor \frac{X}{p_2 q_2} \right\rfloor + X \ll X.$$

If $(p_1 q_1, p_2 q_2) = p_1$ (so that $q_1 \neq q_2$), say, then the sum is

$$\ll p_1^2 q_1 q_2 \left\lfloor \frac{X}{p_1 q_1 q_2} \right\rfloor + p_1 q_1 \left\lfloor \frac{X}{p_1 q_1} \right\rfloor + p_2 q_2 \left\lfloor \frac{X}{p_2 q_2} \right\rfloor + X \ll p_1 X.$$

By symmetry, the analogous result holds if $(p_1q_1, p_2q_2) = q_1$. Finally, if $p_1q_1 = p_2q_2$ then similarly the bound is $\ll p_1q_1 X$. We thus obtain from these cases that the LHS in (35) is bounded above by

$$
\begin{aligned}
&\ll X \sum_{\substack{X^{1/4}<p_1,q_1,p_2,q_2\leq X \\ p_1\neq q_1, p_2\neq q_2 \\ (p_1q_1,p_2q_2)=1}} \frac{|b(p_1,q_1)||b(p_2,q_2)|}{p_1p_2q_1q_2} \\
&+ \sum_{\substack{X^{1/4}<p,q,r\leq X \\ p\neq q, p\neq r, q\neq r}} \frac{|b(p,q)|(|b(p,r)|+|b(r,q)|)}{pqr} \\
&+ X \sum_{\substack{X^{1/4}<p,q\leq X \\ p\neq q}} \frac{|b(p,q)|(|b(p,q)|+|b(q,p)|)}{pq}.
\end{aligned}
$$

By the AM–GM inequality we simply have $2|b(p,q)||b(p',q')| \leq |b(p,q)|^2 + |b(p',q')|^2$ for any pairs of primes p, q and p', q', so invoking Mertens' theorem and symmetry the above expressions are

$$
\ll X \sum_{\substack{X^{1/4}<p,q\leq X \\ p\neq q}} \frac{|b(p,q)|^2}{pq},
$$

and the claim follows. $\qquad\square$

Proof of Proposition 7.1 Let $g \in \mathcal{A}$, and let g^* be the strongly additive function equal to g at primes, provided by Lemma 3.6. Applying Lemma 3.2, then following the proof of Lemma 3.6, we find

$$
\frac{1}{X} \sum_{n\in\mathcal{S}(X)} |(g-g^*) - A_{g-g^*}(X)|^2 \ll B_{g-g^*}(X)^2
$$

$$
\ll \sum_{\substack{p^k\leq X \\ k\geq 2}} \frac{|g(p)|^2 + |g(p^k)|^2}{p^k} \ll \varepsilon B_g(X)^2,
$$

by assumption. It follows that

$$
\frac{1}{X} \sum_{n\in\mathcal{S}(X)} |g(n) - A_g(X)|^2 \ll \varepsilon B_g(X)^2 + \frac{1}{X} \sum_{n\in\mathcal{S}(X)} |g^*(n) - A_{g^*}(X)|^2,
$$

so replacing g by g^*, we may assume in what follows that g is strongly additive. Fix $z = X^{1/4}$ and split $g = g_{\leq z} + g_{>z}$, where $g_{\leq z}$ is the strongly additive function defined by $g_{\leq z}(p^k) := g(p)1_{p\leq z}$ at primes powers p^k. By Cauchy–Schwarz, we seek to estimate

$$\frac{1}{X} \sum_{n \in \mathcal{S}(X)} |g_{\leq z}(n) - A_{g_{\leq z}}(X)|^2 + \frac{1}{X} \sum_{n \in \mathcal{S}(X)} |g_{>z}(n) - A_{g_{>z}}(X)|^2. \qquad (36)$$

We begin with the first expression. Writing

$$g_{\leq z}(n) - A_{g_{\leq z}}(X) = \sum_{p \leq z} g(p) \sum_{1 \leq k \leq \log X / \log p} \left(1_{p^k \| n} - p^{-k}(1 - 1/p)\right)$$

$$= \sum_{p \leq z} g(p)(1_{p|n} - 1/p) + O(X^{-1} \sum_{p \leq z} |g(p)|) \qquad (37)$$

for each $n \leq X$ and expanding the square, the first sum in (36) is

$$= \frac{1}{X} \sum_{n \in \mathcal{S}(X)} \sum_{\substack{p,q \leq z \\ p \neq q}} \frac{g(p)\overline{g(q)}}{pq} (p1_{p|n} - 1)(q1_{q|n} - 1)$$

$$+ O\left(\sum_{p \leq z} \frac{|g(p)|^2}{p^2} \frac{1}{X} \sum_{n \in \mathcal{S}(X)} (p1_{p|n} - 1)^2\right)$$

$$+ O(B_g(z)^2 X^{-1}(z\pi(z))^{1/2})$$

$$=: O + D + O(B_g(X)^2 X^{-3/4}).$$

Consider the off-diagonal term O. Observe that for any two distinct primes p and q, the Chinese remainder theorem implies that

$$(p1_{p|n} - 1)(q1_{q|n} - 1) = \left(\sum_{\substack{a \pmod{p} \\ a \neq 0}} e(an/p)\right)\left(\sum_{\substack{b \pmod{q} \\ b \neq 0}} e(bn/q)\right)$$

$$= \sum_{c \pmod{pq}}^{*} e(cn/pq),$$

where the asterisked sum is over reduced residues modulo pq. Note that for any two distinct products p_1q_1 and p_2q_2 the gap between fractions with these denominators satisfies

$$\left|\frac{c_1}{p_1q_1} - \frac{c_2}{p_2q_2}\right| \geq \frac{1}{p_1q_1p_2q_2} \geq \frac{1}{z^4} = \frac{1}{X},$$

and the number of pairs yielding the same product pq is ≤ 2. Using this expression in O, applying the Cauchy–Schwarz inequality twice followed by the large sieve inequality [31, Lem. 7.11], we obtain

$$O \leq \frac{1}{X} \sum_{\substack{p,q \leq z \\ p \neq q}} \frac{|g(p)g(q)|}{pq} \left| \sum_{c \pmod{pq}}^{*} \sum_{n \leq X} 1_{\mathcal{S}}(n) e(cn/pq) \right|$$

$$\leq \frac{1}{X} \left(\sum_{p \leq z} \frac{|g(p)|^2}{p} \right) \left(\sum_{\substack{pq \leq \sqrt{X} \\ p \neq q}} \frac{1}{pq} \left| \sum_{c \pmod{pq}}^{*} \sum_{n \leq X} 1_{\mathcal{S}}(n) e(cn/pq) \right|^2 \right)^{1/2}$$

$$\ll \frac{1}{X} B_g(X)^2 \left(\sum_{\substack{pq \leq \sqrt{X} \\ p \neq q}} \sum_{c \pmod{pq}}^{*} \left| \sum_{n \leq X} 1_{\mathcal{S}}(n) e(cn/pq) \right|^2 \right)^{1/2}$$

$$\ll \frac{1}{X} B_g(X)^2 \left(X|\mathcal{S}(X)| \right)^{1/2} = \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2} B_g(X)^2.$$

Expanding $g_{>z}(n) - A_{g>z}(X)$ in a similar way as in (37), then inserting this and the previous estimate into (36) we obtain the upper bound,

$$\ll \frac{1}{X} \sum_{p \leq z} \frac{|g(p)|^2}{p^2} \sum_{n \in \mathcal{S}(X)} |p1_{p|n} - 1|^2 + \frac{1}{X} \sum_{n \in \mathcal{S}(X)} \left| \sum_{X^{1/4} < p \leq X} \frac{g(p)}{p} (p1_{p|n} - 1) \right|^2$$

$$+ \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2} B_g(X)^2$$

$$\ll \left| \sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} \frac{g(p)\overline{g(q)}}{pq} \frac{1}{X} \sum_{n \in \mathcal{S}(X)} (p1_{p|n} - 1)(q1_{q|n} - 1) \right|$$

$$+ \sum_{p \leq X} \frac{|g(p)|^2}{p^2} \frac{1}{X} \sum_{n \in \mathcal{S}(X)} |p1_{p|n} - 1|^2 + \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2} B_g(X)^2. \tag{38}$$

Denote by T the expression in (38), so that

$$T = \sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} \frac{g(p)\overline{g(q)}}{pq} \left(\frac{pq}{X} |\mathcal{S}_{pq}(X)| - \frac{p}{X} |\mathcal{S}_p(X)| - \frac{q}{X} |\mathcal{S}_q(X)| + \frac{|\mathcal{S}(X)|}{X} \right)$$

$$=: \sum_{\substack{X^{1/4} < p,q \leq X \\ p \neq q}} \frac{g(p)\overline{g(q)}}{pq} T_{p,q}(X),$$

say. We split the pairs of primes $X^{1/4} < p, q \leq X$, $p \neq q$ in the support of T as follows. Given a squarefree integer d, call $E_d(\varepsilon)$ the condition $\frac{d}{X}|\mathcal{S}_d(X)| \leq \varepsilon$, and let $L_d(\varepsilon)$ be the converse condition $\frac{d}{X}|\mathcal{S}_d(X)| > \varepsilon$. If, simultaneously, the three conditions $E_{pq}(\varepsilon)$, $E_p(\varepsilon)$ and $E_q(\varepsilon)$ all hold, then as $|\mathcal{S}(X)|/X < \varepsilon$ we have $T_{p,q}(X) \ll \varepsilon$; otherwise, we trivially have $T_{p,q}(X) \ll 1$. We thus find by the Cauchy–Schwarz inequality that

$$
T \ll \varepsilon \sum_{\substack{X^{1/4} < p, q \leq X \\ p \neq q}} \frac{|g(p)g(q)|}{pq} + \sum_{\substack{X^{1/4} < p, q \leq X \\ p \neq q \\ L_p(\varepsilon), L_q(\varepsilon) \text{ or } L_{pq}(\varepsilon) \\ \text{holds}}} \frac{|g(p)g(q)|}{pq}
$$

$$
\ll B_g(X)^2 \left(\varepsilon \sum_{X^{1/4} < p \leq X} \frac{1}{p} + \left(\sum_{\substack{X^{1/4} < p, q \leq X \\ p \neq q \\ L_p(\varepsilon), L_q(\varepsilon) \text{ or } L_{pq}(\varepsilon) \\ \text{holds}}} \frac{1}{pq} \right)^{1/2} \right). \tag{39}
$$

Now suppose $L_d(\varepsilon)$ holds for some $d \geq 2$. As $\varepsilon > 2\frac{|\mathcal{S}(X)|}{X}$ we have

$$
\frac{1}{d} = \frac{4d}{(\varepsilon X)^2} \left(\frac{\varepsilon X}{2d} \right)^2 < \frac{4d}{(\varepsilon X)^2} \left(|\mathcal{S}_d(X)| - \frac{|\mathcal{S}(X)|}{d} \right)^2.
$$

Using this with $d = p$ for each $X^{1/4} < p \leq X$, and applying Lemma 7.3, we get

$$
\sum_{\substack{X^{1/4} < p \leq X \\ L_p(\varepsilon) \text{ holds}}} \frac{1}{p} \ll \frac{1}{(\varepsilon X)^2} \sum_{p \leq X} p \left| \sum_{\substack{n \leq X \\ p|n}} 1_{\mathcal{S}}(n) - \frac{1}{p} \sum_{n \leq X} 1_{\mathcal{S}}(n) \right|^2 \ll \varepsilon^{-2} \frac{|\mathcal{S}(X)|}{X};
$$

this, by the way, establishes (34). Similarly, by Lemma 7.4 we get

$$
\sum_{\substack{X^{1/4} < p, q \leq X \\ p \neq q \\ L_{pq}(\varepsilon) \text{ holds}}} \frac{1}{pq} \ll \frac{1}{(\varepsilon X)^2} \sum_{\substack{X^{1/4} < p, q \leq X \\ p \neq q}} pq \left| \sum_{\substack{n \leq X \\ pq|n}} 1_{\mathcal{S}}(n) - \frac{1}{pq} \sum_{n \leq X} 1_{\mathcal{S}}(n) \right|^2
$$

$$
\ll \varepsilon^{-2} \frac{|\mathcal{S}(X)|}{X}.
$$

Combining these estimates in (39) shows that

$$T \ll B_g(X)^2 \left(\varepsilon + \varepsilon^{-1} \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2} \right).$$

Thus, putting $\delta(X) := \varepsilon + \varepsilon^{-1} \left(\frac{|\mathcal{S}(X)|}{X} \right)^{1/2}$, we finally conclude that

$$\frac{1}{X} \sum_{n \in \mathcal{S}(X)} |g(n) - A_g(X)|^2$$

$$\ll \sum_{p \leq X} \frac{|g(p)|^2}{p^2} \frac{1}{X} \sum_{n \in \mathcal{S}(X)} |p 1_{p|n} - 1|^2 + \delta(X) B_g(X)^2$$

$$= \sum_{p \leq X} \frac{|g(p)|^2}{p} \left(\frac{(p-2)|\mathcal{S}_p(X)|}{X} + \frac{|\mathcal{S}(X)|}{pX} \right) + \delta(X) B_g(X)^2$$

$$\ll \sum_{\substack{p \leq X \\ |\mathcal{S}_p(X)| > \varepsilon X/p}} \frac{|g(p)|^2}{p} + B_g(X)^2 \left(\varepsilon + \frac{|\mathcal{S}(X)|}{X} \right) + \delta(X) B_g(X)^2$$

$$\ll \sum_{\substack{p \leq X \\ |\mathcal{S}_p(X)| > \varepsilon X/p}} \frac{|g(p)|^2}{p} + \delta(X) B_g(X)^2,$$

and the claim follows. $\qquad \square$

Corollary 7.5 *Let $g : \mathbb{N} \to \mathbb{R}$ be an additive function in $\mathcal{A}_s$. Let $\mathcal{S} \subset \mathbb{N}$ be a set with $|\mathcal{S}(X)| = o(X)$. Then for any fixed $j \in \mathbb{Z}$,*

$$\frac{1}{X} \sum_{n+j \in \mathcal{S}(X)} |g(n) - A_g(X)|^2 = o(B_g(X)^2).$$

Proof By appealing to Proposition 7.1, we will show that for any $\varepsilon > 0$ there is $X_0(\varepsilon)$ such that if $X \geq X_0(\varepsilon)$ then

$$\frac{1}{X} \sum_{n+j \in \mathcal{S}(X)} |g(n) - A_g(X)|^2 \ll \varepsilon B_g(X)^2,$$

for any fixed $j \in \mathbb{Z}$.
First, note that as $B_g(X) \to \infty$, if $X \geq X_0(\varepsilon)$ then

$$\sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{g(p)^2}{p^k} \ll \sum_{p \leq X} \frac{g(p)^2}{p^2} \leq B_g(\log X)^2 + \frac{1}{\log X} \sum_{\log X < p \leq X} \frac{g(p)^2}{p} \leq \frac{1}{2} \varepsilon B_g(X)^2.$$

Taking X larger if necessary, we may combine this with (23) to deduce that

$$\sum_{\substack{p^k \leq X \\ k \geq 2}} \frac{g(p)^2 + g(p^k)^2}{p^k} \leq \varepsilon B_g(x)^2.$$

Next, for any fixed $j \in \mathbb{Z}$ we have $|(\mathcal{S} - j)(X)| \leq |\mathcal{S}(X)| + j = o(X)$ as $X \to \infty$, and so for X taken even larger if needed we have $|(\mathcal{S} - j)(X)|/X < \varepsilon^4 < \varepsilon/2$, for ε sufficiently small. We claim that

$$\sum_{\substack{p \leq X \\ |(\mathcal{S}-j)_p(X)| > \varepsilon X/p}} \frac{g(p)^2}{p} \ll \varepsilon B_g(X)^2 \tag{40}$$

for sufficiently large X. Assuming this, Proposition 7.1 will imply that

$$\frac{1}{X} \sum_{n+j \in \mathcal{S}(X)} |g(n) - A_g(X)|^2 \ll B_g(X)^2 \left(\varepsilon + \varepsilon^{-1} (\varepsilon^4)^{1/2} \right) + \varepsilon B_g(X)^2 \ll \varepsilon B_g(X)^2,$$

as required.

We may split the sum in (40) according to whether or not $|g(p)| > \delta^{-1} B_g(X)$, where $\delta = \delta(\varepsilon) > 0$ is to be chosen. In light of (34), we obtain

$$\sum_{\substack{p \leq X \\ |(\mathcal{S}-j)_p(X)| > \varepsilon X/p \\ |g(p)| \leq \delta^{-1} B_g(X)}} \frac{g(p)^2}{p} \leq \delta^{-2} B_g(X)^2 \sum_{\substack{p \leq X \\ |(\mathcal{S}-j)_p(X)| > \varepsilon X/p}} \frac{1}{p} \ll (\varepsilon\delta)^{-2} \frac{|\mathcal{S}(X)|}{X} B_g(X)^2,$$

so that this is $\ll \varepsilon B_g(X)^2$ if $X \geq X_0(\varepsilon)$.

On the other hand, by our assumption (22),

$$\sum_{\substack{p \leq X \\ |(\mathcal{S}-j)_p(X)| > \varepsilon X/p \\ |g(p)| > \delta^{-1} B_g(X)}} \frac{g(p)^2}{p} \leq \sum_{\substack{p \leq X \\ |g(p)| > \delta^{-1} B_g(X)}} \frac{g(p)^2}{p} \leq 2 F_g(\delta) B_g(X)^2,$$

provided $X \geq X_0(\delta)$. For $\delta = \delta(\varepsilon)$ sufficiently small we can make this $\ll \varepsilon B_g(X)^2$ whenever $X \geq X_0(\varepsilon)$ (with $X_0(\varepsilon)$ taken larger if necessary). The claim now follows. $\qquad\square$

We are now able to prove the first part of Theorem 1.8, namely that there is a parameter $\lambda = \lambda(X) \ll \frac{B_g(X)}{\log X}$ such that

$$\sum_{p^k \leq X} \frac{|g(p^k) - \lambda(X) \log p^k|^2}{p^k} = o(B_g(X)^2).$$

The proof of the slow variation condition $\lambda(X^u) = \lambda(X) + o(\frac{B_g(X)}{\log X})$, for $0 < u \le 1$ fixed, is postponed to the next section.

Proof of Theorem 1.8: Part I In light of Lemma 3.3, we begin by showing that

$$\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)|^2 = o(B_g(X)^2). \tag{41}$$

By Lemma 3.4, when $X/\log X < Y \le X$ we have

$$|A_g(X) - A_g(Y)| \ll B_g(X)\sqrt{\frac{\log \log X}{\log X}},$$

so that upon dyadically decomposing the range $[X/\log X, X]$ and applying Lemma 3.2 to the range $[1, X/\log X]$ in the sum on the LHS of (41), we get

$$\le \frac{1}{X} \sum_{n \le \frac{X}{\log X}} |g(n) - A_g(X)|^2 + \sum_{1 \le 2^j \le \log X} 2^{-j} \cdot \frac{2^j}{X} \sum_{X/2^{j+1} < n \le X/2^j} |g(n) - A_g(X)|^2$$

$$= \sum_{1 \le 2^j \le \log X} 2^{-j} \cdot \frac{2^j}{X} \sum_{X/2^{j+1} < n \le X/2^j} |g(n) - A_g(X/2^j)|^2 + O\left(\frac{B_g(X)^2 \log \log X}{\log X}\right).$$

It thus suffices to show that, uniformly over $1 \le 2^j \le 2 \log X$,

$$\frac{2^j}{X} \sum_{X/2^{j+1} < n \le X/2^j} |g(n) - A_g(X/2^j)|^2 = o(B_g(X)^2).$$

Fix $1 \le 2^k \le 2 \log X$, set $Y_k := X/2^k$ and introduce a parameter $1 \le R \le (\log X)^{1/2}$, which will eventually be chosen as slowly growing as a function of X. Let

$$\mathcal{B}_R := \bigcup_{|i| \le R} (\mathcal{B} + i), \qquad \mathcal{G}_R(Y_k) := [Y_k/2, Y_k] \setminus \mathcal{B}_R, \qquad \mathcal{B}_R(Y_k) := \mathcal{B}_R \cap [Y_k/2, Y_k].$$

We observe that if $n \in \mathcal{G}_R(Y_k)$ then we have

$$g(n - R) \le g(n - R + 1) \le \cdots \le g(n) \le g(n + 1) \le \cdots \le g(n + R).$$

We divide $\mathcal{G}_R(Y_k)$ further into the sets

$$\mathcal{G}_R^+(Y_k) := \{n \in \mathcal{G}_R(Y_k) : g(n) \ge A_g(Y_k)\}, \quad \mathcal{G}_R^-(Y_k) := \{n \in \mathcal{G}_R(Y_k) : g(n) < A_g(Y_k)\}.$$

Suppose $n \in \mathcal{G}_R^+(Y_k)$. Since

$$0 \le g(n) - A_g(Y_k) \le \frac{1}{R} \sum_{1 \le j \le R} g(n + j) - A_g(Y_k),$$

we deduce from the monotonicity of the map $y \mapsto y^2$ for $y \geq 0$ that (shifting $n \mapsto n + R =: n'$)

$$\frac{2}{Y_k} \sum_{\substack{n \in \mathcal{G}_R^+(Y_k) \\ n+R \leq Y_k}} |g(n) - A_g(Y_k)|^2$$

$$\leq \frac{2}{Y_k} \sum_{\substack{n'-R \in \mathcal{G}_R^+(Y_k) \\ n' \leq Y_k}} \left| \frac{1}{R} \sum_{n'-R < m \leq n'} g(m) - \frac{2}{Y_k} \sum_{Y_k/2 < m \leq Y_k} g(m) \right|^2 + O\left(\frac{B_g(X)^2}{\log X}\right),$$

where the error term comes from replacing $A_g(Y_k)$ by the sum over $[Y_k/2, Y_k]$ by applying Lemma 3.4. Similarly, if $n \in \mathcal{G}_R^-(Y_k)$ then

$$0 \leq A_g(Y_k) - g(n) \leq A_g(Y_k) - \frac{1}{R} \sum_{0 \leq j \leq R-1} g(n-j),$$

and so by the same argument we obtain

$$\frac{2}{Y_k} \sum_{\substack{n \in \mathcal{G}_R^-(Y_k) \\ n-R \geq Y_k/2}} |g(n) - A_g(Y_k)|^2$$

$$\leq \frac{2}{Y_k} \sum_{\substack{n \in \mathcal{G}_R^-(Y_k) \\ n-R \geq Y_k/2}} \left| \frac{1}{R} \sum_{n-R < m \leq n} g(m) - \frac{2}{Y_k} \sum_{Y_k/2 < m \leq Y_k} g(m) \right|^2$$

$$+ O\left(\frac{B_g(X)^2}{\log X}\right).$$

The above sums cover all elements of $\mathcal{G}_R(Y_k)$ besides those in $[Y_k/2, Y_k/2 + R) \cup (Y_k - R, Y_k]$. To deal with these, we define

$$\mathcal{S} := \bigcup_{\substack{j \in \mathbb{Z} \\ 0 < 2^j \leq X}} [X/2^j - R, X/2^j + R] \cap \mathbb{N}, \qquad \mathcal{S}(Z) := \mathcal{S} \cap [1, Z]$$

for $Z \geq 1$. We see that $|\mathcal{S}(Z)| \ll R \log Z = o(Z)$, and $\mathcal{S}$ contains $[Y_k/2, Y_k/2 + R] \cup [Y_k - R, Y_k]$ for each k. By Corollary 7.5 (taking $j = 0$ there), we thus obtain

$$\frac{1}{Y_k} \sum_{n \in \mathcal{G}_R(Y_k) \cap \mathcal{S}(Y_k)} |g(n) - A_g(Y_k)|^2 = o(B_g(X)^2)$$

uniformly over all $X/\log X < Y_k \leq X$, provided X is large enough in terms of R. Combining the foregoing estimates and using positivity, we find that

$$\frac{2}{Y_k} \sum_{n \in \mathcal{G}_R(Y_k)} |g(n) - A_g(Y_k)|^2$$

$$\ll \frac{1}{Y_k} \sum_{Y_k/2 < n \leq Y_k} \left| \frac{1}{R} \sum_{n-R < m \leq n} g(m) - \frac{2}{Y} \sum_{Y_k/2 < m \leq Y_k} g(m) \right|^2 + o(B_g(X)^2).$$

By Theorem 1.4, this gives $o_{R \to \infty}(B_g(X)^2)$, uniformly over $X/\log X < Y_k \leq X$. It remains to estimate the contribution from $n \in \mathcal{B}_R(Y_k)$. By the union bound, we have

$$\frac{2}{Y_k} \sum_{n \in \mathcal{B}_R(Y_k)} |g(n) - A_g(Y_k)|^2 \leq R \max_{|i| \leq R} \frac{2}{Y_k} \sum_{\substack{Y_k/2 < n \leq Y_k \\ n+i \in \mathcal{B}}} |g(n) - A_g(Y_k)|^2.$$

By Corollary 7.5, the above expression is $o(B_g(X))$, again provided X is sufficiently large in terms of R.

To conclude, for any $\varepsilon > 0$ we can find R large enough in terms of ε and X_0 sufficiently large in terms of ε and R such that if $X \geq X_0$ then

$$\frac{2}{Y_k} \sum_{Y_k/2 < n \leq Y_k} |g(n) - A_g(Y_k)|^2 \ll \varepsilon B_g(X)^2$$

uniformly in $X/\log X < Y_k = X/2^k \leq X$, and (41) follows.

Now, applying Lemma 3.3, we deduce that

$$B_{g_{\lambda_0}}(X)^2 + \lambda_0(X)^2 = \sum_{p^k \leq X} \frac{|g(p^k) - \lambda_0(X) \log p^k|^2}{p^k} + \lambda_0(X)^2$$

$$\ll \frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^2 = o(B_g(X)^2),$$

where we recall that $g_{\lambda_0}(n) = g(n) - \lambda_0 \log n$ for all $n \geq 1$, and

$$\lambda_0(X) := \frac{2}{(\log X)^2} \sum_{p \leq X} \frac{g(p) \log p}{p}. \tag{42}$$

Note that by Cauchy–Schwarz and the prime number theorem,

$$\lambda_0(X)^2 \ll \frac{1}{(\log X)^2} \left(\sum_{p \leq X} \frac{g(p)^2}{p} \right) \left(\frac{1}{(\log X)^2} \sum_{p \leq X} \frac{(\log p)^2}{p} \right) \ll \frac{B_g(X)^2}{(\log X)^2}.$$

Thus, $|\lambda_0(X)| \ll B_g(X)/\log X$, and $B_{g_{\lambda_0}}(X)^2 = o(B_g(X)^2)$, as wanted. We will verify that $\lambda_0(X)$ is slowly varying in the next section (immediately following the proof of Proposition 8.1). □

8 Rigidity properties for almost everywhere monotone functions

We continue to assume that g is almost everywhere monotone in the sense of the previous section. Theorem 1.8 claims that an additive function $g \in \mathcal{A}$ is well approximated by a constant times a logarithm, assuming $g(p)$ is not frequently much larger than $B_g(X)$ for $p \leq X$. In this section, we will complete the proof of this theorem, along with those of Corollary 1.7 and Theorem 1.7, all of which are consequences of the almost everywhere monotonicity property. A key input in this direction is Proposition 8.1, which is a structure theorem for the asymptotic mean value $A_g(X)$.

8.1 The structure of $A_g(X)$

The first main result of this section is the following.

Proposition 8.1 *Let $g : \mathbb{N} \to \mathbb{R}$ be an additive function satisfying $B_g(X) \to \infty$ as $X \to \infty$. Assume that $\mathcal{B} := \{n \in \mathbb{N} : g(n) < g(n-1)\}$ satisfies $|\mathcal{B}(X)| := |\mathcal{B} \cap [1, X]| = o(X)$ as $X \to \infty$. Then for each X sufficiently large there is $\lambda = \lambda(X) \in \mathbb{R}$ such that for any $\frac{\log \log X}{\sqrt{\log X}} < \delta \leq 1/4$,*

$$\sum_{X^\delta < p^k \leq X} \frac{1}{p^k} |A_g(X) - A_g(X/p^k) - \lambda \log p^k| = o(B_g(X)(\log(1/\delta))^{1/2}), \quad (43)$$

and also

$$\sum_{X^\delta < p^k \leq X} \frac{|g(p^k) - \lambda \log p^k|}{p^k} = o(B_g(X)(\log(1/\delta))^{1/2}). \quad (44)$$

Furthermore, $A_g(X)$ and $\lambda(X)$ satisfy the following properties:

(i) $\lambda(X) \ll B_g(X)/\log X$,
(ii) for X sufficiently large and any $X^\delta < t_1 \leq t_2 \leq X$,

$$A_g(t_2) = A_g(t_1) + \lambda(X) \log(t_1/t_2) + o((\log(1/\delta)^{1/2} B_g(X)),$$

(iii) for every $u \in (\delta, 1]$ we have

$$\lambda(X) = \lambda(X^u) + o\left((\log(1/\delta))^{1/2} \delta^{-1} \frac{B_g(X)}{\log X}\right).$$

Remark 8.2 It would be desirable to determine $A_g(t)$ directly as a function of t in some range, say $X^\delta < t \leq X$. Proposition 8.1 provides the approximation $A_g(t) =$

$A_g(X^\delta) + (1 - \delta u)\lambda(X) \log t + o(B_g(X))$, where $u := \log X / \log t$, but this still contains a reference to a second value $A_g(X^\delta)$. We might iterate this argument to obtain (using the slow variation of λ) a further approximation in terms of $A_g(X^{\delta^2})$, $A_g(X^{\delta^3})$, and so forth, but without further data about g (say, $A_g(X^{1/1000}) = o(B_g(X))$) it is not obvious that this argument yields an asymptotic formula for $A_g(t)$ alone.

To prove this proposition we will require a few lemmas.

Lemma 8.3 *Let* $g : \mathbb{N} \to \mathbb{C}$ *be an additive function satisfying* $B_g(X) \to \infty$ *as* $X \to \infty$. *Let* $\alpha \in (1, 2)$ *and* $\frac{\log \log X}{\sqrt{\log X}} < \delta \leq 1/4$. *Then*

$$\sum_{X^\delta < p^k \leq X} \frac{1}{p^k} |g(p^k) - A_g(X) + A_g(X/p^k)|$$

$$\ll_\alpha (\log(1/\delta))^{1/2} \left(\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^\alpha \right)^{1/\alpha} + \frac{B_g(X)}{(\log X)^{1/4}}.$$

Proof We will estimate the quantity

$$\mathcal{M} := \frac{1}{X} \sum_{X^\delta < p^k \leq X} \left| \sum_{\substack{n \leq X \\ p^k \| n}} g(n) - \frac{1}{p^k} \left(1 - \frac{1}{p} \right) \sum_{n \leq X} g(n) \right|$$

in two different ways.

First, we will obtain a lower bound for $\mathcal{M}$ as follows. Given $2 \leq Y \leq X$ observe that for any fixed prime p

$$S_p(Y) := \sum_{\substack{m \leq Y \\ p \nmid m}} g(m) = \sum_{m \leq Y} g(m) - \sum_{j \geq 1} \sum_{\substack{r \leq Y/p^j \\ p \nmid r}} (g(r) + g(p^j))$$

$$= \sum_{m \leq Y} g(m) - \sum_{m \leq Y/p} g(m) + \sum_{j \geq 1} \sum_{\substack{r \leq Y/p^j \\ p \nmid r}} (g(p^{j-1}) - g(p^j))$$

$$= Y \left(1 - \frac{1}{p} \right) A_g(Y) + \frac{Y}{p} (A_g(Y) - A_g(Y/p))$$

$$+ \sum_{j \geq 1} (g(p^{j-1}) - g(p^j)) \left(\left\lfloor \frac{Y}{p^j} \right\rfloor - \left\lfloor \frac{Y}{p^{j+1}} \right\rfloor \right) + O \left(\frac{Y B_g(Y)}{\sqrt{\log Y}} \right),$$

where the last equality arises from Lemma 3.1.

Using this estimate with $Y = X/p^k$ for $p^k \le X$, we obtain

$$\sum_{\substack{n \le X \\ p^k \| n}} g(n) = \sum_{\substack{mp^k \le X \\ p \nmid m}} g(mp^k) = g(p^k) \left(\left\lfloor \frac{X}{p^k} \right\rfloor - \left\lfloor \frac{X}{p^{k+1}} \right\rfloor \right) + S_p(X/p^k)$$

$$= \frac{X}{p^k} \left(1 - \frac{1}{p} \right) \left(g(p^k) + A_g(X/p^k) \right)$$

$$+ O \left(\frac{X}{p^{k+1}} |A_g(X/p^k) - A_g(X/p^{k+1})| \right)$$

$$+ O \left(|g(p^k)| + X \sum_{\substack{j \ge 1 \\ p^j \le X/p^k}} \frac{|g(p^{j-1})| + |g(p^j)|}{p^{k+j}} + \frac{X B_g(X/p^k)}{p^k \sqrt{\log(2X/p^k)}} \right),$$

noting that the third error term is 0 unless $p^k \le X/2$. Similarly, again by Lemma 3.1 we have

$$\frac{1}{p^k} \left(1 - \frac{1}{p} \right) \sum_{n \le X} g(n) = \frac{X}{p^k} \left(1 - \frac{1}{p} \right) \left(A_g(X) + O \left(\frac{B_g(X)}{\sqrt{\log X}} \right) \right).$$

We thus deduce that

$$\mathscr{M} = \sum_{X^\delta < p^k \le X} \frac{1}{p^k} \left(1 - \frac{1}{p} \right) \left| g(p^k) + A_g(X/p^k) - A_g(X) \right| + O\left(\mathcal{R}(X) \right) \quad (45)$$

$$\ge \frac{1}{2} \sum_{X^\delta < p^k \le X} \frac{1}{p^k} \left| g(p^k) + A_g(X/p^k) - A_g(X) \right| + O\left(\mathcal{R}(X) \right), \quad (46)$$

where we have set

$$\mathcal{R}(X) := \sum_{X^\delta < p^k \le X} \frac{1}{p^k} \left(1 + \sum_{\substack{j \ge 1 \\ p^j \le X/p^k}} \frac{|g(p^j)|}{p^j} \right) + \sum_{X^\delta < p^k \le X/2} \frac{B_g(X)}{p^k \sqrt{\log(X/p^k)}}$$

$$+ \frac{1}{X} \sum_{p^k \le X} |g(p^k)| + \sum_{X^\delta < p^k \le X} \frac{1}{p^{k+1}} |A_g(X/p^k) - A_g(X/p^{k+1})|$$

$$=: \mathcal{R}_1(X) + \mathcal{R}_2(X) + \mathcal{R}_3(X) + \mathcal{R}_4(X).$$

As in the proof of Lemma 3.1,

$$\mathcal{R}_3(X) \le \frac{1}{\sqrt{X}} \sum_{p^k \le X} \frac{|g(p^k)|}{p^{k/2}} \le \left(\frac{\pi(X)}{X} \right)^{1/2} B_g(X) \ll \frac{B_g(X)}{\sqrt{\log X}}.$$

Next, we may upper bound $\mathcal{R}_2(X)$ as

$$\mathcal{R}_2(X) \leq B_g(X) \left(\frac{1}{(\log X)^{1/4}} \sum_{X^\delta < p^k \leq Xe^{-\sqrt{\log X}}} \frac{1}{p^k} + \sum_{Xe^{-\sqrt{\log X}} < p^k \leq \frac{X}{2}} \frac{1}{p^k} \right)$$

$$\ll \frac{B_g(X)}{(\log X)^{1/4}}.$$

To treat $\mathcal{R}_1(X)$, we use $|g(p^j)|/p^j \leq B_g(p^j) p^{-j/2}$ to get

$$\sum_{X^\delta < p^k \leq X} \sum_{\substack{j \geq 1 \\ p^j \leq X/p^k}} \frac{|g(p^j)|}{p^{j+k}} \leq B_g(X) \sum_{X^\delta < p^k \leq X} \sum_{\substack{j \geq 1 \\ p^j \leq X/p^k}} p^{-(k+j/2)}$$

$$\ll B_g(X) \left(X^{-\delta} \sum_{\substack{X^\delta < p^k \leq X \\ k \geq 2/\delta}} 1 + \sum_{\substack{X^\delta < p^k \leq X \\ p > X^{\delta^2/2}}} p^{-k} \sum_{\substack{j \geq 1 \\ p^j \leq X/p^k}} \frac{1}{p^{j/2}} \right)$$

$$\ll B_g(X) \left(X^{-\delta/2} + \sum_{p > X^{\delta^2/2}} p^{-3/2} \right) \ll B_g(X) X^{-\delta^2/4}.$$

Furthermore, we have

$$\sum_{X^\delta < p^k \leq X} \frac{1}{p^{k+1}} \leq \sum_{1 \leq k \leq \log X} \sum_{\max\{2, X^{\delta/k}\} \leq p \leq X^{1/k}} \frac{1}{p^{k+1}} \ll \sum_{1 \leq k \leq \log X} k X^{-\delta} \ll \frac{(\log X)^2}{X^\delta}.$$
$$\tag{47}$$

Since $B_g(X) \gg 1$ these last two bounds combine to give

$$\mathcal{R}_1(X) \ll B_g(X) X^{-\delta^2/4}.$$

Finally, to bound $\mathcal{R}_4(X)$ we use Lemma 3.4 to obtain

$$|A_g(X/p^k) - A_g(X/p^{k+1})| \ll B_g(X/p^k)(\log\log X)^{1/2} \leq B_g(X)(\log\log X)^{1/2}$$

uniformly over $p^k \leq X$, and thus using (47) we find

$$\mathcal{R}_4(X) \ll B_g(X)(\log\log X)^{1/2} \sum_{X^\delta < p^k \leq X} \frac{1}{p^{k+1}} \ll B_g(X) \frac{(\log X)^3}{X^\delta}.$$

Combining the estimates for $\mathcal{R}_j(X)$, $1 \leq j \leq 4$, in view of the range of δ we finally obtain

$$\mathcal{R}(X) \ll B_g(X)/(\log X)^{1/4}$$

in (45). Thus,

$$\sum_{X^\delta < p^k \leq X} \frac{1}{p^k} |g(p^k) + A_g(X/p^k) - A_g(X)| \leq 2\mathcal{M} + O(B_g(X)/(\log X)^{1/4}).$$

Next, we execute the second estimation of $\mathcal{M}$. If $p^k \in (X^\delta, X]$ define

$$\Delta_g(X; p^k) := \frac{p^k}{X} \left| \sum_{\substack{n \leq X \\ p^k || n}} g(n) - \frac{1}{p^k} \left(1 - \frac{1}{p}\right) \sum_{n \leq X} g(n) \right|.$$

Set $g'(n) := g(n) - A_g(X)$ for $n \leq X$, and note that

$$\Delta_g(X; p^k) = \Delta_{g'}(X; p^k) + O\left(\frac{p^k}{X} |A_g(X)|\right).$$

Thus, as $|A_g(X)| \leq B_g(X)\sqrt{\log \log X}$ by Lemma 3.4, we find

$$\mathcal{M} = \sum_{X^\delta < p^k \leq X} \frac{1}{p^k} \Delta_{g'}(X; p^k) + O\left(\frac{|A_g(X)|\pi(X)}{X}\right)$$

$$= \sum_{X^\delta < p^k \leq X} \frac{1}{p^k} \Delta_{g'}(X; p^k) + O\left(\frac{B_g(X)\sqrt{\log \log X}}{\log X}\right). \tag{48}$$

Recall that $\alpha \in (1, 2)$. Let us now partition the set of prime powers $X^\delta < p^k \leq X$ into the sets

$$\mathcal{P}_1 := \left\{ X^\delta < p^k \leq X : \Delta_{g'}(X; p^k) > \left(\frac{1}{X} \sum_{n \leq X} |g'(n)|^\alpha\right)^{1/\alpha} \right\}$$

$$\mathcal{P}_2 := \left\{ X^\delta < p^k \leq X : \Delta_{g'}(X; p^k) \leq \left(\frac{1}{X} \sum_{n \leq X} |g'(n)|^\alpha\right)^{1/\alpha} \right\}.$$

Note that by Mertens' theorem,

$$\sum_{X^\delta < p^k \le X} \frac{1}{p^k} \le \sum_{X^\delta < p \le X} \frac{1}{p} + X^{-\delta} \sum_{\substack{X^\delta < p^k \le X \\ k \ge 2/\delta}} 1 + \sum_{\substack{X^\delta < p^k \le X \\ 2 \le k \le 2/\delta \\ p > X^{\delta^2/2}}} \frac{1}{p^k}$$

$$\ll \log(1/\delta)\left(1 + \delta^{-1} X^{-\delta^2/2}\right) + X^{-\delta/2} \ll \log(1/\delta). \qquad (49)$$

Using this and Hölder's inequality, we obtain

$$\sum_{X^\delta < p^k \le X} p^{-k} \Delta_{g'}(X; p^k)$$

$$\ll_\alpha (\log(1/\delta))^{1-\frac{1}{\alpha}} \left(\sum_{\substack{X^\delta < p^k \le X \\ p^k \in \mathcal{P}_1}} p^{-k} \Delta_{g'}(X; p^k)^\alpha\right)^{\frac{1}{\alpha}}$$

$$+ (\log(1/\delta))^{\frac{1}{2}} \left(\sum_{\substack{X^\delta < p^k \le X \\ p^k \in \mathcal{P}_2}} p^{-k} \Delta_{g'}(X; p^k)^2\right)^{\frac{1}{2}}.$$

By Theorem 3.1 of [36] this is bounded by

$$\ll (\log(1/\delta))^{\frac{1}{2}} \left(\frac{1}{X} \sum_{n \le X} |g'(n)|^\alpha\right)^{\frac{1}{\alpha}}.$$

Combining this with (45) and (48) completes the proof of the lemma. $\qquad \square$

Next, we show that, in an ℓ^1 sense, $g(p^k)$ is well approximated by $\lambda \log p^k$ on average over the prime powers $X^\delta < p^k \le X$, for some function $\lambda = \lambda(X)$. This will be the λ that appears in the statement of Proposition 8.1.

Lemma 8.4 *There is a parameter $\lambda = \lambda(X) \in \mathbb{R}$ such that the following holds. For any $\alpha \in (1, 2)$,*

$$\sum_{X^\delta < p^k \le X} \frac{1}{p^k} |g(p^k) - \lambda \log p^k| \ll_\alpha (\log(1/\delta))^{1/2} \left(\frac{1}{X} \sum_{n \le X} |g(n) - A_g(X)|^\alpha\right)^{1/\alpha}.$$

Proof By [1, Théorème 1], there are $\lambda = \lambda(X)$ and $c = c(X)$ (both depending on g but independent of α) such that

$$\left(\frac{1}{X}\sum_{n\leq X}|g(n) - A_g(X)|^\alpha\right)^\alpha \gg_\alpha \left(\sum_{p^k\leq X}\frac{|g''_{\lambda,c}(p^k)|^\alpha}{p^k}\right)^{1/\alpha} + \left(\sum_{p^k\leq X}\frac{|g'_{\lambda,c}(p^k)|^2}{p^k}\right)^{1/2}.$$
(50)

Here, writing $g_\lambda(n) = g(n) - \lambda\log(n)$, we have set

$$g'_{\lambda,c}(p^k) := \begin{cases} g_\lambda(p^k) & \text{if } |g_\lambda(p^k)| \leq c, \\ 0 & \text{otherwise;} \end{cases} \qquad g''_{\lambda,c}(p^k) := \begin{cases} 0 & \text{if } |g_\lambda(p^k)| \leq c, \\ g_\lambda(p^k) & \text{otherwise.} \end{cases}$$

Since $g_\lambda(p^k) = g'_{\lambda,c}(p^k) + g''_{\lambda,c}(p^k)$ for each p^k, by Hölder's inequality and (49) once again,

$$\sum_{X^\delta < p^k \leq X}\frac{|g_\lambda(p^k)|}{p^k} \ll_\alpha (\log(1/\delta))^{1-1/\alpha}\left(\sum_{X^\delta < p^k \leq X}\frac{|g''_{\lambda,c}(p^k)|^\alpha}{p^k}\right)^{1/\alpha}$$
$$+ (\log(1/\delta))^{1/2}\left(\sum_{X^\delta < p^k \leq X}\frac{|g'_{\lambda,c}(p^k)|^2}{p^k}\right)^{1/2}.$$

The result now follows upon combining this last estimate with (50) and using positivity. □

To make use of the previous two lemmas we establish the following upper bound for moments of order $\alpha \in [1, 2)$ that crucially uses the almost-everywhere monotonicity property of g.

Lemma 8.5 *Assume that $\mathcal{B} := \{n \in \mathbb{N} : g(n) < g(n-1)\}$ satisfies $|\mathcal{B}(X)| = o(X)$, where $\mathcal{B}(X) = \mathcal{B} \cap [1, X]$. Then for any $\alpha \in [1, 2)$,*

$$\frac{1}{X}\sum_{n\leq X}|g(n) - A_g(X)|^\alpha \ll \left(\frac{\log\log(1/r(X))}{\log(1/r(X))} + (\log X)^{-\frac{1}{800}}\right)^{2-\alpha} B_g(X)^\alpha,$$

where $r(X) := \max_{\frac{X}{\log X} < Y \leq X}\left(\left(\frac{|\mathcal{B}(Y)|}{Y}\right)^{1/2} + \frac{\log Y}{\sqrt{Y}}\right)$.

Proof By Hölder's inequality, for any $\alpha \in (1, 2)$ we have

$$\frac{1}{X}\sum_{n\leq X}|g(n) - A_g(X)|^\alpha \leq \left(\frac{1}{X}\sum_{n\leq X}|g(n) - A_g(X)|\right)^{2-\alpha}\cdot\left(\frac{1}{X}\sum_{n\leq X}|g(n) - A_g(X)|^2\right)^{\alpha-1},$$

an inequality that is vacuously also true when $\alpha = 1$. Applying Lemma 3.2 to the second bracketed expression, we obtain the upper bound

$$\ll B_g(X)^{2\alpha-2} \left(\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)| \right)^{2-\alpha}.$$

Next, we show that for all $X/\log X < Y \leq X$ we get

$$\frac{1}{Y} \sum_{n \leq Y} |g(n) - g(n-1)| \ll \left(\left(\frac{|\mathcal{B}(Y)|}{Y} \right)^{1/2} + \frac{\log Y}{\sqrt{Y}} \right) B_g(Y) \leq r(X) B_g(X). \quad (51)$$

This will imply the claim of the lemma, since by Proposition 6.1 the latter bound gives

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)| \ll \left(\frac{\log \log(1/r(X))}{\log(1/r(X))} + (\log X)^{-\frac{1}{800}} \right) B_g(X).$$

To prove (51), we note that for all $1 \leq n \leq Y$,

$$|g(n) - g(n-1)| = g(n) - g(n-1) + 2|g(n) - g(n-1)| 1_{n \in \mathcal{B}(Y)}.$$

It follows from this and telescoping that

$$\frac{1}{Y} \sum_{n \leq Y} |g(n) - g(n-1)| = \frac{g(\lfloor Y \rfloor)}{Y} + \frac{2}{Y} \sum_{n \in \mathcal{B}(Y)} |g(n) - g(n-1)|.$$

By Lemma 3.5, $g(\lfloor Y \rfloor)/Y \ll B_g(Y)(\log Y)Y^{-1/2}$. Owing to Lemma 3.2 and the triangle and Cauchy–Schwarz inequalities, we also obtain

$$\frac{1}{Y} \sum_{n \in \mathcal{B}(Y)} |g(n) - g(n-1)| \leq 2 \left(\frac{|\mathcal{B}(Y)|}{Y} \right)^{1/2} \left(\frac{1}{Y} \sum_{n \leq Y} |g(n) - A_g(Y)|^2 \right)^{1/2}$$

$$\ll \left(\frac{|\mathcal{B}(Y)|}{Y} \right)^{1/2} B_g(Y).$$

This implies (51), and completes the proof of the lemma. $\qquad \square$

Proof of Proposition 8.1 We begin with the proof of (44). Set $\alpha = 3/2$, say. Combining Lemmas 8.4 and 8.5, there is a $\lambda = \lambda(X) \in \mathbb{R}$ such that

$$\sum_{X^\delta < p^k \leq X} \frac{|g(p^k) - \lambda \log p^k|}{p^k} \ll (\log(1/\delta))^{1/2} \left(\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^{3/2} \right)^{2/3}$$

$$= o((\log(1/\delta))^{1/2} B_g(X)). \quad (52)$$

We use this to obtain (43). Indeed, this time combining Lemmas 8.3 and 8.5 , we get

$$\sum_{X^\delta < p^k \leq X} \frac{|g(p^k) - A_g(X) + A_g(X/p^k)|}{p^k}$$

$$\ll (\log(1/\delta))^{1/2} \left(\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)|^{3/2} \right)^{2/3} + \frac{B_g(X)}{(\log X)^{1/4}}$$

$$= o((\log(1/\delta))^{1/2} B_g(X)).$$

Combining this with (52) and applying the triangle inequality in the form

$$|A_g(X) - A_g(X/p^k) - \lambda \log p^k| \leq |g(p^k) - A_g(X) + A_g(X/p^k)| + |g(p^k) - \lambda \log p^k|$$

for each $X^\delta < p^k \leq X$, we quickly deduce (43).
Next, we proceed to the proofs of properties (i)–(iii).
(i) By the triangle inequality and positivity, we obtain

$$|\lambda(X)| \sum_{X^{1/4} < p \leq X^{1/2}} \frac{\log p}{p} \leq \sum_{X^{1/4} < p \leq X} \frac{|A_g(X) - A_g(X/p) - \lambda(X) \log p|}{p}$$

$$+ \sum_{X^{1/4} < p \leq X^{1/2}} \frac{|A_g(X) - A_g(X/p)|}{p}. \qquad (53)$$

By Mertens' theorem,

$$\sum_{X^{1/4} < p \leq X^{1/2}} \frac{\log p}{p} = \frac{1}{4} \log X + O\left(\frac{1}{\log X} \right) \gg \log X,$$

and by the Cauchy–Schwarz inequality we have, for $X^{1/4} \leq p \leq X^{1/2}$,

$$|A_g(X) - A_g(X/p)| \ll B_g(X) \left(\sum_{X/p \leq q^k \leq X} \frac{1}{q^k} \right)^{1/2} \ll B_g(X) \left(\frac{\log p}{\log X} \right)^{1/2}.$$

Using the above, (43), the prime number theorem and partial summation in (53), we find that

$$|\lambda(X)| \log X \ll o(B_g(X)) + \frac{B_g(X)}{\sqrt{\log X}} \sum_{p \leq X^{1/2}} \frac{(\log p)^{1/2}}{p} \ll B_g(X),$$

and (i) follows immediately.

(ii) We observe, using (i) and (44) that if $X^\delta < t_1 \le t_2 \le X$,

$$
\begin{aligned}
&\left| A_g(t_2) - A_g(t_1) - \lambda(X) \log(t_2/t_1) \right| \\
&= \left| \sum_{t_1 < p^k \le t_2} \left(1 - \frac{1}{p}\right) \frac{g(p^k) - \lambda(X) \log p^k}{p^k} \right| \\
&\quad + O\left(\lambda(X) \left(\frac{1}{\delta \log X} + \sum_{t_1 < p^k \le t_2} \frac{\log p^k}{p^{k+1}} \right) \right) \\
&\le \sum_{X^\delta < p^k \le X} \frac{|g(p^k) - \lambda(X) \log p^k|}{p^k} + O\left(B_g(X) \left(\frac{1}{\delta (\log X)^2} + \frac{(\log X)^3}{x^\delta} \right) \right) \\
&= o((\log(1/\delta))^{1/2} B_g(X)),
\end{aligned}
$$

where in the penultimate line the second error term is estimated similarly to (47). This proves the required estimate.

(iii) Applying (ii) with $(t_1, t_2) = (X^y, X^z)$, where, in sequence, $(y, z) = (u, 1)$, $(y, z) = (uv, 1)$ and $(y, z) = (uv, u)$ for any $v \in (\delta/u, 1/2]$ and $u \in (\delta, 1]$, we get

$$
\begin{aligned}
A_g(X) - A_g(X^u) &= (1 - u)\lambda(X) \log X + o((\log(1/\delta))^{1/2} B_g(X)) \\
A_g(X) - A_g(X^{uv}) &= (1 - uv)\lambda(X) \log X + o((\log(1/\delta))^{1/2} B_g(X)) \\
A_g(X^u) - A_g(X^{uv}) &= (u - uv)\lambda(X^u) \log X + o((\log(1/\delta))^{1/2} B_g(X^u)).
\end{aligned}
$$

We subtract the second equation from the first and combine the result with the third equation. Using $B_g(X^u) \le B_g(X)$ we conclude that

$$
u(1 - v)\lambda(X) \log X = u(1 - v)\lambda(X^u) \log X + o((\log(1/\delta))^{1/2} B_g(X)).
$$

Since $1 - v \ge 1/2$ and $u > \delta$, the claim follows immediately upon rearranging (with a potentially larger implicit constant in the error term). □

Proof of Theorem 1.8: Part II The work at the end of Sect. 7 implies that

$$
\sum_{p^k \le X} \frac{|g(p^k) - \lambda_0(X) \log p^k|^2}{p^k} = B_{g_{\lambda_0}}(X)^2 = o(B_g(X)^2), \tag{54}
$$

where λ_0 is as in (42). Now, by Proposition 8.1 we have

$$
\sum_{X^{1/2} < p^k \le X} \frac{|g(p^k) - \lambda(X) \log p^k|}{p^k} = o(B_g(X)), \tag{55}
$$

where $\lambda(X)$ satisfies

$$
\lambda(X) = \lambda(X^u) + o\left(\tfrac{B_g(X)}{\log X}\right), \quad 0 < u \le 1 \text{ fixed.}
$$

Thus, by (54) and (55), Cauchy–Schwarz and Mertens' theorem, whenever $Y = X^u$ with $0 < u \leq 1$ fixed we have

$$|\lambda(Y) - \lambda_0(Y)| \log Y \ll |\lambda(Y) - \lambda_0(Y)| \sum_{Y^{1/2} < p \leq Y} \frac{\log p}{p} \tag{56}$$

$$\leq \sum_{Y^{1/2} < p \leq Y} \frac{|g(p) - \lambda_0(Y) \log p|}{p}$$

$$+ \sum_{Y^{1/2} < p \leq Y} \frac{|g(p) - \lambda(Y) \log p|}{p}$$

$$\leq B_{g_{\lambda_0}}(Y) \left(\sum_{Y^{1/2} < p \leq Y} \frac{1}{p} \right)^{1/2} + o(B_g(Y)) = o(B_g(X)). \tag{57}$$

We thus deduce that $\lambda(X^u) = \lambda_0(X^u) + o(B_g(X)/\log X)$ for all $0 < u \leq 1$ fixed, and therefore also that

$$\lambda_0(X^u) = \lambda(X^u) + o\left(\frac{B_g(X)}{\log X} \right) = \lambda(X) + o\left(\frac{B_g(X)}{\log X} \right) = \lambda_0(X) + o\left(\frac{B_g(X)}{\log X} \right),$$

and the second claim of Theorem 1.8 is proved. □

8.2 Growth of $B_g(X)$ and the proof of Corollary 1.7

In this subsection we prove Corollary 1.7. The key step will be to show that if there is a $\lambda(X)$ such that $B_{g_\lambda}(X) = o(B_g(X))$ (which follows from Theorem 1.8) then $B_g(X)$ grows roughly like $\log X$.

We begin by showing that this is the case assuming in addition that $\lambda(X)$ is fairly large (this assumption is subsequently removed in Lemma 8.7).

Until further notice, assume that $B_g(X) \gg 1$ as X gets large.

Lemma 8.6 *Assume that there is a $C > 0$ such that $\lambda(X) \geq C B_g(X)/\log X$ for all X sufficiently large in the conclusion of Proposition 8.1. Then for any $\varepsilon > 0$, $(\log X)^{1-\varepsilon} \ll_\varepsilon B_g(X) \ll_\varepsilon (\log X)^{1+\varepsilon}$.*

Proof By Proposition 8.1 and our assumption $\lambda(X) \gg B_g(X)/\log X$,

$$\lambda(X) = \lambda(X^u) + o(B_g(X)/\log X) = \lambda(X^u) + o(\lambda(X)) \tag{58}$$

whenever $0 < u \leq 1$ is fixed. This implies in particular that $\lambda(X) \ll \lambda(X^u)$. Setting $Y := X^u$ and $v := 1/u \geq 1$, we see also that

$$\lambda(Y^v) = \lambda(Y) + o(\lambda(Y^v)) = \lambda(Y) + o(\lambda(Y^{uv})) = \lambda(Y) + o(\lambda(Y)).$$

Thus, (58) holds for all fixed $u \geq 1$ as well, and thus for all fixed $u > 0$. We thus deduce that for each $u > 0$ fixed and $\varepsilon > 0$ there is $X_0(\varepsilon, u)$ such that if $X \geq X_0(\varepsilon, u)$,

$$\left| \frac{\lambda(X^u)}{\lambda(X)} - 1 \right| < \varepsilon.$$

Set $u = 1/2$, put $X_0 = X_0(\varepsilon, 1/2)$ and for each $k \geq 1$ define $X_k := X_0^{2^k}$. Let K be large. Then

$$\frac{\lambda(X_0)}{\lambda(X_K)} = \prod_{1 \leq k \leq K} \frac{\lambda(X_{k-1})}{\lambda(X_k)} \in [(1-\varepsilon)^K, (1+\varepsilon)^K].$$

As $K \leq 2 \log\log X_K$ for large enough K and X_0, we find

$$|\lambda(X_K)| \leq |\lambda(X_0)| \exp\left(-K \log(1-\varepsilon)\right) \ll_\varepsilon \exp\left(4\varepsilon \log\log X_K\right) = (\log X_K)^{4\varepsilon},$$

$$|\lambda(X_K)| \geq |\lambda(X_0)| \exp\left(-K \log(1+\varepsilon)\right) \gg_\varepsilon \exp\left(-4\varepsilon \log\log X_K\right) = (\log X_K)^{-4\varepsilon}.$$

Thus, we have $B_g(X_K) \ll |\lambda(X_K)|(\log X_K) \ll_\varepsilon (\log X_K)^{1+4\varepsilon}$ by assumption, and by Proposition 8.1(i) we have $B_g(X_K) \gg |\lambda(X_K)| \log X_K \gg_\varepsilon (\log X_k)^{1-4\varepsilon}$. Since $\log X_K \asymp \log X_{K+1}$, by monotonicity of B_g we also have

$$B_g(X) \leq B_g(X_{K+1}) \ll_\varepsilon (\log X_{K+1})^{1+4\varepsilon} \ll_\varepsilon (\log X_K)^{1+4\varepsilon} \leq (\log X)^{1+4\varepsilon}$$

for any $X_K < X < X_{K+1}$. Similarly, we obtain $B_g(X) \gg_\varepsilon (\log X)^{1-4\varepsilon}$ on the same interval. Since $\varepsilon > 0$ was arbitrary, the claim now follows. ∎

Lemma 8.7 *Assume* $B_{g_v}(X) = o(B_g(X))$ *for some* $v = v(X)$ *that satisfies* $|v| \ll B_g(X)/\log X$. *Then for any* $\varepsilon > 0$, $(\log X)^{1-\varepsilon} \ll_\varepsilon B_g(X) \ll_\varepsilon (\log X)^{1+\varepsilon}$.

Proof By Cauchy–Schwarz, we have

$$B_g(X)^2 = \sum_{p^k \leq X} \frac{|g(p^k)|^2}{p^k} \leq 2 \left(v(X)^2 \sum_{p^k \leq X} \frac{(\log p^k)^2}{p^k} + \sum_{p^k \leq X} \frac{|g_v(p^k)|^2}{p^k} \right)$$

$$= 2v(X)^2 (\log X)^2 + o(B_g(X)^2).$$

It follows that $|v(X)| \geq \frac{1}{4} B_g(X)/\log X$ when X is sufficiently large. The conclusion follows from Lemma 8.6, provided we can show that $\lambda(X) = v(X) + o(B_g(X)/\log X)$ for all large X, where $\lambda(X)$ is the function from the conclusion of Proposition 8.1. But this can be verified by the same argument as that which leads to (57), so the claim follows. ∎

Proof of Corollary 1.7 Suppose $g : \mathbb{N} \to \mathbb{R}$ is a completely additive function that satisfies

$$F_g(\varepsilon) \to 0 \text{ as } \varepsilon \to 0^+, \text{ and } |\mathcal{B}(X)| \leq \frac{X}{(\log X)^{2+\eta}}, \text{ for some } \eta > 0.$$

Suppose first that $B_g(X) \to \infty$, so that $g \in \mathcal{A}_s$. By Theorem 1.8 there is a parameter $\lambda_0(X)$ with $|\lambda_0(X)| \ll B_g(X)/\log X$, such that $B_{g_{\lambda_0}}(X) = o(B_g(X))$, as $X \to \infty$. By Lemma 8.7, we deduce that $B_g(X) \ll_\varepsilon (\log X)^{1+\varepsilon}$. Now, applying (51), we obtain

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| \ll B_g(X) \left(\left(\frac{|\mathcal{B}(X)|}{X} \right)^{\frac{1}{2}} + \frac{\log X}{\sqrt{X}} \right)$$

$$\ll_\eta (\log X)^{1+\frac{\eta}{3}} \cdot (\log X)^{-\frac{1}{2}(2+\eta)} \ll (\log X)^{-\frac{\eta}{6}}.$$

By Theorem 3.7, we deduce that there is a constant $c \in \mathbb{R}$ such that $g(n) = c \log n$ for all n, as required.

If, instead, $B_g(X) \ll 1$ then we again deduce (even if $g \notin \mathcal{A}_s$) from (51) that

$$\frac{1}{X} \sum_{n \leq X} |g(n) - g(n-1)| = o(B_g(X)) = o(1),$$

and so the claim follows (necessarily with $c = 0$) by Theorem 3.7. $\qquad\square$

8.3 Proof of Theorem 1.9

To prove Theorem 1.9 we will appeal to the following result due to Elliott, which will be useful for us in light of our Proposition 8.1.

Theorem [26, Thm. 6] *Let $0 < a < b \leq 1$. Let $g : \mathbb{N} \to \mathbb{C}$ be an additive function, and for $y \geq 10$ define*

$$\theta(y) := \sum_{y^a < p^k \leq y^b} \frac{1}{p^k} |g(p^k) - A_g(y) + A_g(y/p^k)|.$$

Then for all $\varepsilon, B > 0$ there exist $X_0 = X_0(a, b, \varepsilon, B)$ and $c > 0$ such that if $X \geq X_0$ then, uniformly over $X^\varepsilon < t \leq X$,

$$A_g(t) = G(X) \log t - \eta(X) + O(Y(X)),$$

where G, η are measurable functions and

$$Y(X) := \sup_{X^c < w \leq X} \theta(w) + (\log X)^{-B} \sum_{p^k \leq X} \frac{|g(p^k)|}{p^k} + \max_{X^c \leq p^k \leq X} |g(p^k)| p^{-k}.$$

Corollary 8.8 *Let $\delta \in (0, 1/2)$. Suppose $g : \mathbb{N} \to \mathbb{R}$ is an additive function such that $|\mathcal{B}(X)| = o(X)$. Then, uniformly over all $X^\delta \leq t \leq X$ we have*

$$A_g(t) = \lambda(X) \log t - \eta(X) + o(B_g(X)),$$

where $\lambda(X)$ and $\eta(X)$ are measurable functions such that for each fixed $0 < u \le 1$,

$$\lambda(X^u) = \lambda(X) + o(\tfrac{B_g(X)}{\log X}), \qquad \eta(X^u) = \eta(X) + o(B_g(X)).$$

Proof By combining Lemmas 8.3 and 8.5 , we have

$$\sum_{X^\delta \le p^k \le X} \frac{|g(p^k) - A_g(X) + A_g(X/p^k)|}{p^k} = o(B_g(X)),$$

for any fixed $\delta > 0$. Applying Elliott's theorem with $a = \varepsilon = \delta$, $b = 1$, $B = 1$, we have

$$Y(X) = o(B_g(X)) + O\left(\frac{\sqrt{\log \log X}}{\log X} B_g(X) + B_g(X) X^{-\delta/2}\right) = o(B_g(X)),$$

using Lemma 3.4 to treat the second error term, and the bound $|g(p^k)| p^{-k/2} \le B_g(X)$ for all $p^k \le X$ in the third. We thus deduce the existence of $G(X)$ such that

$$A_g(t) = G(X) \log t - \eta(X) + o(B_g(X)). \tag{59}$$

For $X^\delta < t_1 \le t_2 \le X$,

$$A_g(t_2) - A_g(t_1) = G(X) \log(t_2/t_1) + o(B_g(X)),$$

so we have removed the term $\eta(X)$. Now, by Proposition 8.1 we also know that in the same range,

$$A_g(t_2) - A_g(t_1) = \lambda(X) \log(t_2/t_1) + o(B_g(X)).$$

Applying this with $t_1 = X^{1/2}$, $t_2 = X$, we deduce readily that

$$G(X) = \lambda(X) + o\left(\frac{B_g(X)}{\log X}\right),$$

and hence, from (59), that for all $X^\delta < t \le X$,

$$A_g(t) = \lambda(X) \log t - \eta(X) + o(B_g(X)).$$

The slow variation of $\lambda(X)$ is a consequence of Proposition 8.1(iii). To obtain the corresponding property for η we evaluate $A_g(X^u)$ in (59), once as written and once with X replaced by X^u, obtaining

$$A_g(X^u) = u\lambda(X^u) \log X - \eta(X^u) + o(B_g(X^u)) = u\lambda(X) \log X - \eta(X) + o(B_g(X)),$$

from which it also follows, using the slow variation of λ, that

$$\eta(X^u) = \eta(X) + u(\lambda(X^u) - \lambda(X)) \log X + o(B_g(X)) = \eta(X) + o(B_g(X))$$

for each fixed $\delta \leq u \leq 1$, as required. $\square$

Proof of Theorem 1.9 By Lemma 8.5 (with $\alpha = 1$),

$$\frac{1}{X} \sum_{n \leq X} |g(n) - A_g(X)| = o(B_g(X))$$

so that for all but $o(X)$ integers $n \leq X$ we have

$$g(n) = A_g(X) + o(B_g(X)). \tag{60}$$

By Corollary 8.8 and Proposition 8.1(i), we deduce that

$$g(n) = \lambda(X) \log X - \eta(X) + o(B_g(X)) = \lambda(X) \log n - \eta(X) + o(B_g(X))$$

for all but $o(X)$ integers $X/\log X < n \leq X$, and thus for all but $o(X)$ integers $n \leq X$, proving the claim of Theorem 1.9. $\square$

References

1. Hildebrand, A.: Sur les moments d'une fonction additive. Ann. de l'Institut Fourier **33**(3), 1–22 (1983)
2. Erdös, P., Wintner, A.: Additive arithmetical functions and statistical independence. Am. J. Math. **61**, 713–721 (1939)
3. Erdos, P., Kac, M.: The Gaussian law of errors in the theory of additive number theoretic functions. Am. J. Math. **62**, 738–742 (1940)
4. Wirsing, E.: Das asymptotische Verhalten von Summen über multiplikative Funktionen II. Acta Math. Acad. Sci. Hung. **18**, 411–467 (1967)
5. Halász, G.: Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen. Acta Math. Acad. Sci. Hung. **19**, 365–403 (1968)

6. Chowla, S.: The Riemann Hypothesis and Hilbert's Tenth Problem. Gordon and Breach Science Publishers, New York (1965)
7. Landau, E.: Über den Zusammenhang einiger neuer Sätze der analytischen Zahlentheorie. Math. Klasse **115**, 589–632 (1906)
8. Matomäki, K., Radziwiłł, M.: Multiplicative functions in short intervals. Ann. Math. (2) **183**(3), 1015–1056 (2016)
9. Matomäki, K., Radziwiłł, M., Tao, T.: An averaged form of Chowla's conjecture. Algebra Number Theory **9**(9), 2167–2196 (2015)
10. Matomäki, K., Radziwiłł, M.: Multiplicative functions in short intervals II. arXiv: 2007.04290 [math.NT]
11. Tao, T.: The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. Forum Math. Pi **4**, 8–36 (2016)
12. Tao, T.: The Erdős discrepancy problem. Discrete Anal. **1**, 29 (2016)
13. Kułaga-Przymus, J., Lemańczyk, M.: Sarnak's conjecture from the ergodic theory point of view. arXiv: 2009.04757 [math.DS]
14. Mangerel, A.P.: Divisor-bounded multiplicative functions in short intervals. arXiv: 2108.11401 [math.NT]
15. Goudout, E.: Théorème d'Erdös-Kac dans presque tous les petits intervalles. Acta Arith. **182**(2), 101–116 (2018)
16. Goudout, E.: Lois locales de la fonction ω dans presque tous les petits intervalles. Proc. Lond. Math. Soc. **115**(3), 599–637 (2017)
17. Erdős, P.: On the distribution function of additive functions. Ann. Math. **47**(2), 1–20 (1946)
18. Kátai, I.: On a problem of P. Erdős. J. Number Theory **2**(1), 1–6 (1970)
19. Wirsing, E.: Characterization of the logarithm as an additive function. Am. Math. Soc. Proc. Symp. Pure Math. **20**, 375–381 (1971)
20. Hildebrand, A.: An Erdös-Wintner theorem for differences of additive functions. Trans. Am. Math. Soc. **310**(1), 257–276 (1988)
21. Wirsing, E.: A characterization of $\log n$ as an additive arithmetic function. Symp. Math. dell' Istituto Nazionale di Aha Mathematica Roma **IV**, 45–57 (1970)
22. Elliott, P.D.T.A.: Arithmetic Functions and Integer Products, vol. 272. Springer, Berlin (1985)
23. Klurman, O.: Correlations of multiplicative functions and applications. Compos. Math. **153**(8), 1622–1657 (2017). https://doi.org/10.1112/S0010437X17007163
24. Klurman, O., Mangerel, A.P.: Rigidity theorems for multiplicative functions. Math. Ann. **372**(1–2), 651–697 (2018). https://doi.org/10.1007/s00208-018-1724-6
25. Kátai, I.: Continuous homomorphisms as arithmetical functions, and sets of uniqueness. In: Trends in Mathematics, pp. 183–200. Birkhäuse, Basel (2000)
26. Elliott, P.D.T.A.: Functional analysis of additive arithmetic functions. Bull. Am. Math. Soc. **16**(2), 179–223 (1987)
27. Hildebrand, A.: Additive functions at consecutive integers. J. Lond. Math. Soc. **35**(2), 217–232 (1987)
28. Mangerel, A.P.: On the bivariate Erdös-Kac theorem and correlations of the Möbius function. Math. Proc. Camb. Phil. Soc. **169**(3), 547–605 (2020)
29. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Cambridge University Press, Cambridge (1987)
30. Ruzsa, I.Z.: On the variance of additive functions. In: Studies in Pure Mathematics: To the Memory of Paul Turán, 576–586. Editors: Paul Erdös, László Alpár, Gábor Halász and András Sarközy, Springer Basel AG
31. Iwaniec, H., Kowalski, E.: Analytic Number Theory, vol. 53, p. 615. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence (2004)
32. Granville, A., Soundararajan, K.: Large character sums: pretentious characters and the Pólya-Vinogradov theorem. J. Am. Math. Soc **20**(2), 357–384 (2007)
33. Granville, A., Soundararajan, K.: Decay of mean values of multiplicative functions. Can. J. Math. **55**(6), 1191–1230 (2003)
34. Klurman, O., Mangerel, A.P., Teräväinen, J.: Multiplicative functions in short arithmetic progressions. arXiv:1909.12280 [math.NT]
35. Shiu, P.: A Brun-Titchmarsh theorem for multiplicative functions. J. Reine Angew. Math. **313**, 161–170 (1980)

36. Elliott, P.D.T.A.: Duality in Analytic Number Theory. Cambridge Tracts in Mathematics, vol. 122. Cambridge University Press, Cambridge (1997)