TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 10, October 2017, Pages 7085–7105
http://dx.doi.org/10.1090/tran/6870
Article electronically published on March 30, 2017

MULTIPLE ERGODIC THEOREMS FOR ARITHMETIC SETS

NIKOS FRANTZIKINAKIS AND BERNARD HOST

Abstract. We establish results with an arithmetic flavor that generalize the polynomial multidimensional Szemerédi theorem and related multiple recurrence and convergence results in ergodic theory. For instance, we show that in all these statements we can restrict the implicit parameter n to those integers that have an even number of distinct prime factors or satisfy any other congruence condition. In order to obtain these refinements we study the limiting behavior of some closely related multiple ergodic averages with weights given by appropriately chosen multiplicative functions. These averages are then analyzed using a recent structural result for bounded multiplicative functions proved by the authors.

1. Introduction and main results

1.1. **Introduction.** The multidimensional Szemerédi theorem of H. Furstenberg and Y. Katznelson [14], stated in ergodic terms, asserts that if $T_1, \dots, T_\ell$ are commuting measure preserving transformations acting on the same probability space $(X, \mathcal{X}, \mu)$, then for every $A \in \mathcal{X}$ with $\mu(A) > 0$ there exists $n \in \mathbb{N}$ such that

$$\mu(T_1^{-n} A \cap \cdots \cap T_\ell^{-n} A) > 0.$$

More recently, T. Tao [23] established mean convergence for some closely related multiple ergodic averages by showing that for $F_1, \dots, F_\ell \in L^\infty(\mu)$ the averages

$$\frac{1}{N} \sum_{n=1}^{N} T_1^n F_1 \cdots T_\ell^n F_\ell$$

converge in the mean as $N \to \infty$. In this article we are interested in studying variants of such statements where the parameter n is restricted to certain subsets of the integers of arithmetic nature. For instance, we are interested in knowing whether the previous results remain true when we restrict the parameter n to those integers that have an even (or an odd) number of distinct prime factors. More generally, do they hold if we restrict n to those integers that have $a \bmod b$ distinct prime factors for some $a, b \in \mathbb{N}$?

We answer these questions affirmatively. In order to give some model results in this introductory section (extensions and related statements appear in Section 1.2) we introduce some notation. For $a, b \in \mathbb{N}$ we let $S_{a,b}$ consist of those $n \in \mathbb{N}$ whose number of distinct prime factors is congruent to $a \bmod b$. It can be shown that for

Received by the editors March 27, 2015 and, in revised form, March 30, 2015 and November 7, 2015.

2010 *Mathematics Subject Classification.* Primary 37A45; Secondary 05D10, 11B30, 11N37, 28D05.

Key words and phrases. Multiple ergodic averages, multiple recurrence, multiplicative functions, higher degree uniformity.

every $a \in \{0, \dots, b-1\}$ the set $S_{a,b}$ has density $1/b$ (see the second remark after Proposition 2.10).

Theorem A. *Let $T_1, \dots, T_\ell$ be commuting measure preserving transformations acting on the same probability space $(X, \mathcal{X}, \mu)$. Then for every $A \in \mathcal{X}$ with $\mu(A) > 0$ we have*

$$\mu(T_1^{-n} A \cap \dots \cap T_\ell^{-n} A) > 0$$

for a set of $n \in S_{a,b}$ with positive lower density.

We deduce from this ergodic statement, via the correspondence principle of H. Furstenberg (see Section 1.3), that every set of integers with positive upper density contains arbitrarily long arithmetic progressions with common difference taken from the set $S_{a,b}$; similar statements also hold for the multidimensional Szeméredi theorem, polynomial variants of it (see Theorem 1.5), and for any shift of the sets $S_{a,b}$.

Theorem B. *Let $T_1, \dots, T_\ell$ be commuting measure preserving transformations acting on the same probability space $(X, \mathcal{X}, \mu)$. Then for all $F_1, \dots, F_\ell \in L^\infty(\mu)$, the averages*

$$\frac{1}{N} \sum_{n \in S_{a,b} \cap [1,N]} T_1^n F_1 \cdots T_\ell^n F_\ell \tag{1}$$

converge in $L^2(\mu)$. In fact, the limit is equal to $\lim_{N\to\infty} \frac{1}{bN} \sum_{n=1}^N T_1^n F_1 \cdots T_\ell^n F_\ell$.

In order to analyze the averages (1) we do not use the theory of characteristic factors; even for averages of the form $\frac{1}{N} \sum_{n=1}^N T_1^n F_1 \cdots T_\ell^n F_\ell$ this theory is very intricate and not yet developed to an extent that facilitates our study. Instead, we proceed by comparing the averages (1) with the averages $\frac{1}{bN} \sum_{n=1}^N T_1^n F_1 \cdots T_\ell^n F_\ell$ and show that the difference converges to 0 in $L^2(\mu)$. To do this, we work with some weighted multiple ergodic averages with weights given by suitably chosen multiplicative functions. Then the asserted convergence to 0 is a consequence of the next statement.

Theorem C. *Let $f \in \mathcal{M}_{conv}$ be a multiplicative function (see definition in Section 1.2). If $T_1, \dots, T_\ell$ are commuting measure preserving transformations acting on the same probability space $(X, \mathcal{X}, \mu)$, then for all $F_1, \dots, F_\ell \in L^\infty(\mu)$, the averages*

$$\frac{1}{N} \sum_{n=1}^N f(n) \cdot T_1^n F_1 \cdots T_\ell^n F_\ell \tag{2}$$

converge in $L^2(\mu)$. Furthermore, the limit is zero if f is aperiodic (see definition in Section 1.2).

Let us briefly explain how we derive Theorem A and B from Theorem C. For $b \in \mathbb{N}$ we let ζ be a root of unity of order b and let f be the multiplicative function defined by $f(p^k) = \zeta$ for all primes p and all $k \in \mathbb{N}$. Note that

$$\mathbf{1}_{S_{a,b}}(n) = \frac{1}{b} \sum_{j=0}^{b-1} \zeta^{-aj} (f(n))^j. \tag{3}$$

It follows from Corollary 2.10 that for $j = 1, \dots, b-1$ the multiplicative function f^j is aperiodic. Combining this with (3) and Theorem C we get that the difference

$$\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{S_{a,b}}(n) \cdot T_1^n F_1 \cdots T_\ell^n F_\ell - \frac{1}{bN}\sum_{n=1}^{N} T_1^n F_1 \cdots T_\ell^n F_\ell \tag{4}$$

converges to 0 in $L^2(\mu)$. Using this and the aforementioned convergence result of T. Tao we deduce Theorem B. Furthermore, since the difference (4) converges to 0 in $L^2(\mu)$, letting $F_1 = \cdots = F_\ell = \mathbf{1}_A$ where $\mu(A) > 0$, integrating over X and using the multidimensional Szemerédi theorem of Furstenberg and Katznelson, we deduce Theorem A.

The proof of Theorem C depends upon a deep structural result for multiplicative functions proved by the authors in [11]. Roughly speaking, it asserts that the general bounded multiplicative function can be decomposed into two terms, one that is approximately periodic and another that contributes negligibly to the averages (2). The approximately periodic component vanishes if the multiplicative function is aperiodic. In the general case, a careful analysis of the contribution of the structured component allows us to conclude the proof of Theorem C.

We note that if one is only interested in the weak convergence of the averages (2), an alternate (and arguably simpler) approach is to use a decomposition result for multiple correlation sequences from [9]; we discuss this approach in more detail in Section 3.5.

1.2. **Recurrence and convergence results.** Our main results cover a vastly more general setting than the one described in the previous subsection. In order to facilitate exposition we introduce some definitions and notation. We start with some notions from number theory related to multiplicative functions.

Definition. A function $f\colon \mathbb{N} \to \mathbb{C}$ is called *multiplicative* if

$$f(mn) = f(m)f(n) \quad \text{whenever} \quad (m,n) = 1.$$

We let

$$\mathcal{M} := \{f\colon \mathbb{N} \to \mathbb{C} \text{ is multiplicative such that } |f(n)| \le 1 \text{ for every } n \in \mathbb{N}\}$$

and

$$\mathcal{M}_{conv} := \Big\{ f \in \mathcal{M}\colon \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(an+b) \text{ exists for every } a,b \in \mathbb{N} \Big\}.$$

We say that $f \in \mathcal{M}$ is *aperiodic* if $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(an+b) = 0$ for every $a, b \in \mathbb{N}$.

If a multiplicative function takes real values, then a well-known theorem of E. Wirsing [25] states that it has a mean value; furthermore it belongs to $\mathcal{M}_{conv}$. But there exist complex valued multiplicative functions that do not have a mean value, for example if $f(n) = n^{it}$ for some $t \neq 0$; then $\frac{1}{N}\sum_{n=1}^{N} f(n) = \frac{N^{it}}{1+it} + o_{N\to\infty}(1)$. Lending terminology from [18], it can be shown that $f \in \mathcal{M}_{conv}$ unless $f(n)$ "pretends" to be $n^{it}\chi(n)$ for some $t \in \mathbb{R}$ and Dirichlet character χ. Necessary and sufficient conditions for checking when a multiplicative function belongs to the set $\mathcal{M}_{conv}$ can be found in Theorem 2.9 below.

Next we introduce some notions from ergodic theory.

Definition. • A bounded sequence of complex numbers $(w(n))$ is a *good universal weight for polynomial multiple mean convergence* if for every $\ell, m \in \mathbb{N}$, probability space $(X, \mathcal{X}, \mu)$, invertible commuting measure preserving transformations $T_1, \dots, T_\ell \colon X \to X$, functions $F_1, \dots, F_m \in L^\infty(\mu)$, and polynomials $p_{i,j} \colon \mathbb{Z} \to \mathbb{Z}$, $i = 1, \dots \ell$, $j = 1, \dots, m$, the averages

$$\frac{1}{N} \sum_{n=1}^{N} w(n) \cdot (\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)}) F_1 \cdot \dots \cdot (\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)}) F_m$$

converge in $L^2(\mu)$, where $\prod_{i=1}^{\ell} S_i$ denotes the composition $S_1 \circ \dots \circ S_\ell$. A set of integers S is *a set of polynomial multiple mean convergence* if the sequence $(\mathbf{1}_S(n))$ is a good universal weight for polynomial multiple mean convergence.

• A set of integers S is a *set of polynomial multiple recurrence* if for every $\ell, m \in \mathbb{N}$, probability space $(X, \mathcal{X}, \mu)$, invertible commuting measure preserving transformations $T_1, \dots, T_\ell \colon X \to X$, set $A \in \mathcal{X}$ with $\mu(A) > 0$, and polynomials $p_{i,j} \colon \mathbb{Z} \to \mathbb{Z}$, $i = 1, \dots \ell$, $j = 1, \dots, m$, with $p_{i,j}(0) = 0$, we have

$$\mu\big((\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)}) A \cap \dots \cap (\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)}) A\big) > 0 \tag{5}$$

for a set of $n \in S$ with positive lower density.

Remark. All the statements in this article refer to sets of integers S with positive density; hence there is no need to normalize the relevant averages. Furthermore, although we always work under the assumption that the measure preserving transformations commute, with some additional work our arguments extend to the case where the transformations generate a nilpotent group; we discuss this in more detail in Section 3.4.

Our first result generalizes Theorem C from the introduction.

Theorem 1.1. *Let $f \in \mathcal{M}_{conv}$ be a multiplicative function. Then the sequence $(f(n))$ is a good universal weight for polynomial multiple mean convergence. Furthermore, if f is aperiodic, then the corresponding weighted ergodic averages converge to 0 in $L^2(\mu)$.*

Remark. Examples of periodic systems show that one does not have convergence if $f \notin \mathcal{M}_{conv}$. Nevertheless, following the method of [10] it is possible to show that for every $f \in \mathcal{M}$ there exist $t \in \mathbb{R}$ and a slowly varying sequence $\eta(n)$ (meaning $\max_{x \le n \le x^2} |\eta(n) - \eta(x)| \to 0$ as $x \to \infty$), where both t and η depend only on f, such that the corresponding weighted ergodic averages multiplied by $N^{-it} e(-\eta(n))$ converge in the mean.

If S is the set of square-free integers, applying Theorem 1.1 for the multiplicative function $f := \mathbf{1}_S$ we deduce that S is a set of polynomial multiple mean convergence.[1]

Next, we give generalizations of Theorems A and B from the introduction. They are consequences of a result that we state next. A subset of the unit interval or the unit circle is called *Riemann-measurable* if its indicator function is a Riemann integrable function. It is known that this condition is equivalent to having boundary of Lebesgue measure 0.

[1] This can also be deduced directly from Theorem 2.2 by approximating S in density by periodic sets.

Theorem 1.2. *Let $f \in \mathcal{M}$ be a multiplicative function taking values on the unit circle.*

(i) *If for some $k \in \mathbb{N}$ the function f takes values on k-th roots of unity and K is a non-empty subset of its range, then $f^{-1}(K)$ is a set of polynomial multiple mean convergence. If in addition f^j is aperiodic for $j = 1, \dots, k-1$, then any shift of the set $f^{-1}(K)$ is a set of polynomial multiple recurrence.*

(ii) *If f^j is aperiodic for all $j \in \mathbb{N}$ and K is a Riemann-measurable subset of the unit circle of positive measure, then any shift of the set $f^{-1}(K)$ is a set of polynomial multiple recurrence and polynomial multiple mean convergence.*

In fact, under the aperiodicity assumptions of part (i) *or* (ii) *we get that the set $f^{-1}(K)$ is Gowers uniform (see definition in Section* 2.2*).*

We denote by $\omega(n)$ the number of distinct prime factors of an integer n and by $\Omega(n)$ the number of prime factors of n counted with multiplicity. We let

$$S_{\omega,A,b} := \{n \in \mathbb{N} \colon \omega(n) \equiv a \mod b \text{ for some } a \in A\}$$

and similarly we define $S_{\Omega,A,b}$.

Corollary 1.3. *For every $b \in \mathbb{N}$ and $A \subset \{0, \dots, b-1\}$ non-empty, any shift of the sets $S_{\omega,A,b}$ and $S_{\Omega,A,b}$ is a set of polynomial multiple recurrence and polynomial multiple mean convergence. In fact, all these sets are Gowers uniform.*

Remark. As the polynomial $2^a n^b$ takes values in $S_{\Omega,a,b}$, the multiple recurrence property (5) for some $n \in S_{\Omega,a,b}$ can be inferred from the polynomial Szemerédi theorem. This argument does not apply for non-trivial shifts of the sets $S_{\Omega,a,b}$ and $S_{\omega,a,b}$. On the other hand, see Section 3.6 for an alternate argument that can be used to prove "linear" multiple recurrence statements by finding IP_k-patterns within any shift of $S_{\omega,a,b}$ and $S_{\Omega,a,b}$.

For $\alpha \in \mathbb{R}$ and $A \subset [0, 1/2]$ we let

$$S_{\omega,A,\alpha} := \{n \in \mathbb{N} \colon \|\omega(n)\alpha\| \in A\}, \quad S_{\Omega,A,\alpha} := \{n \in \mathbb{N} \colon \|\Omega(n)\alpha\| \in A\},$$

where $\|x\| := d(x, \mathbb{Z})$ for $x \in \mathbb{R}$.

Corollary 1.4. *For every irrational α and Riemann-measurable set $A \subset [0, 1/2]$ of positive measure, any shift of the sets $S_{\omega,A,\alpha}$ and $S_{\Omega,A,\alpha}$ is a set of polynomial multiple recurrence and mean convergence. In fact, all these sets are Gowers uniform.*

Remark. Similar results hold if $S_{\omega,A,\alpha}$ and $S_{\Omega,A,\alpha}$ are defined using fractional parts.

1.3. **Combinatorial implications.** We give some combinatorial implications of the previous multiple recurrence results. We define the *upper Banach density* $d^*(E)$ of a set $E \subset \mathbb{Z}^\ell$ as $d^*(E) := \limsup_{|I| \to \infty} \frac{|E \cap I|}{|I|}$, where the lim sup is taken over all parallelepipeds $I \subset \mathbb{Z}^\ell$ whose side lengths tend to infinity. We use the following modification of the correspondence principle of H. Furstenberg (the proof can be found in [4]).

Furstenberg correspondence principle ([13])**.** Let $\ell \in \mathbb{N}$ and $\mathbb{E} \subset \mathbb{Z}^\ell$. There exist a probability space $(X, \mathcal{X}, \mu)$, invertible commuting measure preserving transformations $T_1, \dots, T_\ell \colon X \to X$, and a set $A \in \mathcal{X}$ with $\mu(A) = d^*(E)$, such that

$$d^*((E-\vec{n}_1)\cap\cdots\cap(E-\vec{n}_m)) \geq \mu\big((\prod_{i=1}^{\ell} T_i^{n_{i,1}})A\cap\cdots\cap(\prod_{i=1}^{\ell} T_i^{n_{i,m}})A\big)$$

for all $m \in \mathbb{N}$ and $\vec{n}_j = (n_{1,j}, \dots, n_{\ell,j}) \in \mathbb{Z}^\ell$ for $j = 1, \dots, m$.

Using this result and Corollaries 1.3 and 1.4, we immediately deduce the following.

Theorem 1.5. *Let $\ell, m \in \mathbb{N}$, let $\vec{q}_1, \dots, \vec{q}_m \colon \mathbb{Z} \to \mathbb{Z}^\ell$ be polynomials with $\vec{q}_i(0) = \vec{0}$ for $i = 1, \dots, m$, and let $E \subset \mathbb{Z}^\ell$ with $d^*(E) > 0$. Then the set*

$$\big\{n \in \mathbb{N} \colon d^*\big((E-\vec{q}_1(n))\cap\cdots\cap(E-\vec{q}_m(n))\big) > 0\big\}$$

intersects any shift of the sets $S_{\omega,A,b}$, $S_{\Omega,A,b}$, $S_{\omega,A,\alpha}$, $S_{\Omega,A,\alpha}$ (we assume that the set A satisfies the assumptions of Corollaries 1.3 *and* 1.4*) on a set of positive lower density.*

1.4. **Pointwise convergence.** Variants of the previous mean convergence results that deal with pointwise convergence of multiple ergodic averages are, for the most part, completely open. The situation is only clear for the single term ergodic averages

$$\frac{1}{N}\sum_{n=1}^{N} f(n)\cdot F(T^n x) \tag{6}$$

where $F \in L^\infty(\mu)$. If f is the Möbius or the Liouville function, then it is shown in [1, Proposition 3.1] that these averages converge pointwise to 0. This is done by combining the spectral theorem with some classical quantitative bounds of H. Davenport [7] for averages of the form $\frac{1}{N}\sum_{n=1}^{N} f(n)\,\mathrm{e}(nt)$; note though that such bounds do not hold for general aperiodic multiplicative functions.

For more general $f \in \mathcal{M}_{conv}$ we can treat pointwise convergence of the averages (6) as follows: If F is orthogonal to the Kronecker factor of the system, then for every $f \in \mathcal{M}$ the averages (6) converge pointwise to 0. We can establish this by combining an orthogonality criterion of I. Kátai [21] with a result of J. Bourgain [3]. The former implies that the averages (6) converge to zero if $\frac{1}{N}\sum_{n=1}^{n} F(T^{an}x)\cdot \overline{F}(T^{bn}x) \to 0$ for every $a, b \in \mathbb{N}$ with $a \neq b$, and the latter confirms this property pointwise almost everywhere when F is orthogonal to the Kronecker factor of the system. On the other hand, suppose that F is an eigenfunction with eigenvalue $e(\alpha)$ for some $\alpha \in \mathbb{R}$. If α is irrational, then using a result of H. Daboussi [5,6] we deduce that for all $f \in \mathcal{M}$ the averages (6) converge to 0 pointwise. If α is rational, then they converge for all $f \in \mathcal{M}_{conv}$. Furthermore, in either case, the averages (6) converge to 0 if f is aperiodic. Combining the above and using an approximation argument, we get that if $f \in \mathcal{M}_{conv}$, then the averages (6) converge pointwise, and they converge to 0 if f is aperiodic. We deduce from this that all the sets $S_{\omega,A,b}$, $S_{\Omega,A,b}$, $S_{\omega,A,\alpha}$, $S_{\Omega,A,\alpha}$ defined in Section 1.2 are good for pointwise convergence of single ergodic averages, and under the obvious non-degeneracy assumptions for the set A we get that for ergodic systems the normalized averages converge to $\int F\,d\mu$ for all $F \in L^\infty(\mu)$. Furthermore, an approximation argument allows us to extend these results to all $F \in L^1(\mu)$.

We record here a related open problem regarding multiple ergodic averages with arithmetic weights (perhaps the simplest of this type).

Problem. Let $f \in \mathcal{M}_{conv}$ be a multiplicative function. Is it true that for every measure preserving system $(X, \mathcal{X}, \mu, T)$ and every $F, G \in L^\infty(\mu)$, the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(n) \cdot F(T^n x) \cdot G(T^{2n} x) \tag{7}$$

converge pointwise? Do they converge to 0 if f is aperiodic?

When $f = 1$ the averages (7) converge pointwise by a result of J. Bourgain [3]. In general, the problem is open even when f is the Liouville function; that is, it is not known whether the averages $\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_S(n) \cdot F(T^n x) \cdot G(T^{2n} x)$ converge pointwise when S is the set of integers that have an even number of prime factors counted with multiplicity.

1.5. **Notation and conventions.** For the reader's convenience, we gather here some notation that we use throughout the article. We denote by $\mathbb{N}$ the set of positive integers and by $\mathbb{P}$ the set of prime numbers. For $N \in \mathbb{N}$ we let $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ and $[N] := \{1, \dots, N\}$. We let $\mathrm{e}(t) := e^{2\pi i t}$. With $o_{N\to\infty}(1)$ we denote a quantity that converges to 0 when $N \to \infty$ and all other implicit parameters are fixed. Given transformations $T_i \colon X \to X$, $i = 1, \dots, \ell$, with $\prod_{i=1}^{\ell} T_i$ we denote the composition $T_1 \circ \cdots \circ T_\ell$. We use the letter f to denote a multiplicative function. A Dirichlet character, denoted by χ, is a completely multiplicative function that is periodic and satisfies $\chi(1) = 1$.

2. Main ingredients

2.1. **Multiple recurrence and convergence results.** In order to prove our main results we will use some well-known multiple recurrence and convergence results in ergodic theory. The first is the polynomial Szemerédi theorem stated in ergodic terms.

Theorem 2.1 (Bergelson, Leibman [4])**.** *The set of positive integers is a set of polynomial multiple recurrence.*

The second is a mean convergence result for multiple ergodic averages.

Theorem 2.2 (Walsh [24])**.** *The set of positive integers is a set of polynomial multiple mean convergence.*

2.2. **Gowers norms and estimates.** We recall the definition of the U^s-Gowers uniformity norms from [16].

Definition (Gowers norms on $\mathbb{Z}_N$ [16])**.** Let $N \in \mathbb{N}$ and $a \colon \mathbb{Z}_N \to \mathbb{C}$. For $s \in \mathbb{N}$ the *Gowers $U^s(\mathbb{Z}_N)$-norm* $\|a\|_{U^s(\mathbb{Z}_N)}$ of a is defined inductively as follows: For every $t \in \mathbb{Z}_N$ we write $a_t(n) := a(n+t)$. We let

$$\|a\|_{U^1(\mathbb{Z}_N)} := \Big|\frac{1}{N}\sum_{n\in\mathbb{Z}_N} a(n)\Big|$$

and for every $s \in \mathbb{N}$ we let

$$\|a\|_{U^{s+1}(\mathbb{Z}_N)} := \Big(\frac{1}{N}\sum_{t\in\mathbb{Z}_N} \|a \cdot \overline{a}_t\|^{2^s}_{U^s(\mathbb{Z}_N)}\Big)^{1/2^{s+1}}.$$

If $a\colon \mathbb{N}\to\mathbb{C}$ is an infinite sequence, then by $\|a\|_{U^s(\mathbb{Z}_N)}$ we denote the $U^s(\mathbb{Z}_N)$-norm of the restriction of a to the interval $[N]$, thought of as a function on $\mathbb{Z}_N$.

The following uniformity estimates will be used to analyze the limiting behavior of multiple ergodic averages.

Lemma 2.3 (Uniformity estimates [12, Lemma 3.5])**.** *Let $\ell,m\in\mathbb{N}$, let $(X,\mathcal{X},\mu)$ be a probability space, let $T_1,\dots,T_\ell\colon X\to X$ be invertible commuting measure preserving transformations, let $F_1,\dots,F_m\in L^\infty(\mu)$ be functions bounded by 1, and let $p_{i,j}\colon\mathbb{Z}\to\mathbb{Z}$, $i\in\{1,\dots,\ell\}$, $j\in\{1,\dots,m\}$, be polynomials. Let $w\colon\mathbb{N}\to\mathbb{C}$ be a sequence of complex numbers that is bounded by 1. Then there exists $s\in\mathbb{N}$, depending only on the maximum degree of the polynomials $p_{i,j}$ and the integers ℓ and m, such that*

$$\begin{aligned}\Big\|\frac{1}{N}\sum_{n=1}^{N}w(n)\cdot(\prod_{i=1}^{\ell}T_i^{p_{i,1}(n)})F_1\cdot\dots\cdot(\prod_{i=1}^{\ell}T_i^{p_{i,m}(n)})F_m\Big\|_{L^2(\mu)}\\ \ll\big\|\mathbf{1}_{[N]}\cdot w\big\|_{U^s(\mathbb{Z}_{sN})}+o_N(1).\end{aligned}$$

Furthermore, the implicit constant and the $o_N(1)$ term depend only on the integer s.

We also need the following result, which follows from Lemmas A.1 and A.2 in [11].

Lemma 2.4. *Let $s\geq 2$ be an integer and $\varepsilon,\kappa>0$. Then there exist $\delta>0$ and $N_0\in\mathbb{N}$, such that for all integers $N,\widetilde{N}$ with $N_0\leq N\leq\widetilde{N}\leq\kappa N$, every interval $J\subset[N]$, and $f\colon\mathbb{Z}_{\widetilde{N}}\to\mathbb{C}$ with $|f|\leq 1$, the following implication holds:*

$$\text{if}\quad \|f\|_{U^s(\mathbb{Z}_{\widetilde{N}})}\leq\delta,\quad\text{then}\quad \|\mathbf{1}_J\cdot f\|_{U^s(\mathbb{Z}_N)}\leq\varepsilon.$$

2.3. **Gowers uniform sets.** We introduce here the notion of a Gowers uniform subset of the integers that was used repeatedly in the statements of our main results.

Definition. We say that a set of positive integers S is *Gowers uniform* if there exists a positive constant c such that

$$\lim_{N\to\infty}\|\mathbf{1}_S-c\|_{U^s(\mathbb{Z}_N)}=0$$

for every $s\in\mathbb{N}$.

Remark. If such a constant exists, then applying the defining property for $s=1$ gives that c is the density of the set S.

If S is a Gowers uniform set, then applying Lemma 2.3 for the weight $w(n)=\mathbf{1}_S(n)-c$, $n\in\mathbb{N}$, and combining the definition of Gowers uniformity with Lemma 2.4, we deduce that for

$$V_n:=(\prod_{i=1}^{\ell}T_i^{p_{i,1}(n)})F_1\cdot\dots\cdot(\prod_{i=1}^{\ell}T_i^{p_{i,m}(n)})F_m,$$

we have

$$\frac{1}{|S\cap[N]|}\sum_{n\in S\cap[N]}V_n-\frac{1}{N}\sum_{n=1}^{N}V_n\to^{L^2(\mu)}0.$$

Using this, the recurrence result of Theorem 2.1, and the convergence result of Theorem 2.2, we deduce the following.

Proposition 2.5. *Suppose that the set $S \subset \mathbb{N}$ is Gowers uniform. Then any shift of S is a set of polynomial multiple recurrence and polynomial multiple mean convergence.*

2.4. **Structure theorem for multiplicative functions and aperiodicity.** Next we state a structural result from [11] that is crucial for our study. We first introduce some notation from [11, Section 3]. Given $f\colon \mathbb{N} \to \mathbb{C}$ and $N \in \mathbb{N}$ we let

$$f_N := f \cdot \mathbf{1}_{[N]}$$

and whenever appropriate we consider f_N as a function in $\mathbb{Z}_N$. The Fourier transform $\widehat{f_N}$ of f_N is defined by

$$\widehat{f_N}(\xi) := \frac{1}{N}\sum_{n=1}^{N} f(n)\,\mathrm{e}\big(-n\frac{\xi}{N}\big) \quad \text{for } \xi \in \mathbb{Z}_N.$$

By a *kernel* on $\mathbb{Z}_N$ we mean a non-negative function on $\mathbb{Z}_N$ with average 1. For every prime number N and $\theta > 0$, in [11] we defined two positive integers $Q = Q(\theta)$ and $V = V(\theta)$, and for $N > 2QV$, a function $\phi_{N,\theta}\colon \mathbb{Z}_N \to \mathbb{C}$ given by the formula

$$\phi_{N,\theta} := \sum_{\xi \in \Xi_{N,\theta}} \big(1 - \Big\|\frac{Q\xi}{N}\Big\| \frac{N}{QV}\big)\,\mathrm{e}\big(n\frac{\xi}{N}\big)$$

where

$$\Xi_{N,\theta} := \Big\{\xi \in \mathbb{Z}_N \colon \Big\|\frac{Q\xi}{N}\Big\| < \frac{QV}{N}\Big\}. \tag{8}$$

Then for every $\xi \in \mathbb{Z}_N$ we have

$$\widehat{\phi_{N,\theta}}(\xi) = \begin{cases} 1 - \Big\|\frac{Q\xi}{N}\Big\| \frac{N}{QV} & \text{if } \ \xi \in \Xi_{N,\theta}\ ; \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Theorem 2.6 (Structure theorem for multiplicative functions [11, Theorem 8.1])**.** *Let $s \in \mathbb{N}$ and $\varepsilon > 0$. Then there exist a real number $\theta > 0$ and $N_0 \in \mathbb{N}$, depending on s and ε only, such that for every prime $N \geq N_0$, every $f \in \mathcal{M}$ admits the decomposition*

$$f(n) = f_{N,\mathrm{st}}(n) + f_{N,\mathrm{un}}(n), \qquad \textit{for every } \ n \in [N],$$

where $f_{N,\mathrm{st}}, f_{N,\mathrm{un}}\colon [N] \to \mathbb{C}$ are bounded by 1 and 2 respectively and satisfy:

(i) *$f_{N,\mathrm{st}} = f_N * \phi_{N,\theta}$ where $\phi_{N,\theta}$ is the kernel on $\mathbb{Z}_N$ defined previously and the convolution product is defined in $\mathbb{Z}_N$;*

(ii) *$\|f_{N,\mathrm{un}}\|_{U^s(\mathbb{Z}_N)} \leq \varepsilon$.*

Remark. In [11] this result is stated with f multiplied by a certain cut-off. The cut-off is not needed for our purposes, and exactly the same argument proves the current version.

We think of $f_{N,\mathrm{st}}$ and $f_{N,\mathrm{un}}$ as the structured and uniform component of f respectively.

From this point on we assume that $N > 2QV$. When convenient we identify $\mathbb{Z}_N$ with the set $\{0, \dots, N-1\}$ and we denote by $(a, b) \bmod N$ the set that consists of

those $\xi \in \mathbb{Z}_N$ such that $\xi + kN \in (a,b)$ for some $k \in \mathbb{Z}$. Note that $\xi \in \Xi_{N,\theta}$ if and only if there exists $p \in \mathbb{Z}$ such that $\xi - \frac{p}{Q}N \in (-V,V) \bmod N$. Hence,

$$\Xi_{N,\theta} = \bigcup_{p=0}^{Q-1} \big(\frac{p}{Q}N - V, \frac{p}{Q}N + V\big) \bmod N.$$

We may choose to include or omit the endpoints of each interval (if they are integers), since for these values the Fourier transform of $\phi_{N,\theta}$ is 0. Hence, we can assume that

$$\Xi_{N,\theta} = \bigcup_{p=0}^{Q-1} \Xi_{N,\theta,p} \tag{10}$$

where for $p = 0, \dots, Q-1$ we have

$$\Xi_{N,\theta,p} := \big\{ \lfloor \frac{p}{Q}N \rfloor + j \bmod N\colon -V < j \le V \big\}.$$

Note that for fixed $N > 2QV$ and $\theta > 0$ the sets $\Xi_{N,\theta,p}$, $p = 0, \dots, Q-1$, are disjoint, each of cardinality $2V$, hence $|\Xi_{N,\theta}| = 2QV$. Furthermore, if $N \equiv 1 \bmod Q$, then

$$\Xi_{N,\theta,p} = \big\{ \frac{p}{Q}(N-1) + j \bmod N\colon -V < j \le V \big\}. \tag{11}$$

Restricting N to a specific congruence class mod$Q$ is needed in the proof of Lemma 3.3.

We will also use the following consequence of Theorem 2.6; it can be derived by combining Theorem 2.4 and Lemma A.1 in [11].

Theorem 2.7 (Aperiodic multiplicative functions [11])**.** *Let $f \in \mathcal{M}$ be an aperiodic multiplicative function, and for $N \in \mathbb{N}$ let I_N be a subinterval of $[N]$. Then*

$$\lim_{N\to\infty} \|\mathbf{1}_{I_N} \cdot f\|_{U^s(\mathbb{Z}_N)} = 0 \quad \textit{for every } s \in \mathbb{N}.$$

2.5. **Halász's theorem and consequences.** To facilitate exposition, we define the distance between two multiplicative functions as in [18]:

Definition. If $f, g \in \mathcal{M}$ we let $\mathbb{D}\colon \mathcal{M} \times \mathcal{M} \to [0,\infty]$ be given by

$$\mathbb{D}(f,g)^2 = \sum_{p\in\mathbb{P}} \frac{1}{p} \big(1 - \mathrm{Re}\big(f(p)\overline{g(p)}\big)\big).$$

Remark. Note that if $|f| = |g| = 1$, then $\mathbb{D}(f,g)^2 = \sum_{p\in\mathbb{P}} \frac{1}{2p} |f(p) - g(p)|^2$.

It can be shown (see [18]) that $\mathbb{D}$ satisfies the triangle inequality

$$\mathbb{D}(f,g) \le \mathbb{D}(f,h) + \mathbb{D}(h,g).$$

Also for all $f_1, f_2, g_1, g_2 \in \mathcal{M}$ we have (see [17, Lemma 3.1])

$$\mathbb{D}(f_1 f_2, g_1 g_2) \le \mathbb{D}(f_1, g_1) + \mathbb{D}(f_2, g_2). \tag{12}$$

We will also use that if $f \in \mathcal{M}$ is such that for some c in the unit circle we have $f(p) = c$ for all primes p, then $\mathbb{D}(f, n^{it}) = \infty$ for every $t \neq 0$. In particular we have $\mathbb{D}(1, n^{it}) = \infty$ for every $t \neq 0$. Using this and the triangle inequality, one deduces that for $f \in \mathcal{M}$ we have $\mathbb{D}(f, n^{it}) < \infty$ for at most one value of $t \in \mathbb{R}$. We will use the following celebrated result of G. Halász.

Theorem 2.8 (Halász [19]). *A multiplicative function $f \in \mathcal{M}$ has mean value zero if and only if for every $t \in \mathbb{R}$ we either have $\mathbb{D}(f, n^{it}) = \infty$ or $f(2^k) = -2^{ikt}$ for all $k \in \mathbb{N}$.*

Remark. Since f is aperiodic if and only if for every Dirichlet character χ the multiplicative function $f \cdot \chi$ has mean value zero, this result also gives necessary and sufficient conditions for aperiodicity.

Another consequence of the mean value theorem of Halász (see for example [8, Theorem 6.3]) is the following result that gives easy-to-check necessary and sufficient conditions for a multiplicative function to have a mean value (not necessarily zero).

Theorem 2.9. *Let $f \in \mathcal{M}$. Then f has a mean value if and only if we have either:*

(i) $\mathbb{D}(f, n^{it}) = \infty$ *for every* $t \in \mathbb{R}$, *or*
(ii) $\sum_{p \in \mathbb{P}} \frac{1}{p}(1 - f(p))$ *converges, or*
(iii) *for some* $t \in \mathbb{R}$ *we have* $\mathbb{D}(f, n^{it}) < \infty$ *and* $f(2^k) = -2^{ikt}$ *for all* $k \in \mathbb{N}$.

Remark. Since $f \in \mathcal{M}_{conv}$ if and only if for every Dirichlet character χ the multiplicative function $f \cdot \chi$ has a mean value, this result also gives necessary and sufficient conditions for a multiplicative function to be in $\mathcal{M}_{conv}$.

We deduce from the previous results the following criterion that will be used in the proof of Theorem 1.2 and the proof of Corollaries 1.3 and 1.4:

Proposition 2.10. *Let $f \in \mathcal{M}$.*

(i) *If for some $k \in \mathbb{N}$, f takes values on the k-th roots of unity, then $f \in \mathcal{M}_{conv}$.*
(ii) *If $\alpha \in \mathbb{R}$ is not an integer and $f(p) = e(\alpha)$ for all $p \in \mathbb{P}$, then f is aperiodic.*

Remarks. • Sharper results can be obtained using a theorem of R. Hall [20] and the argument in [17, Corollary 2]. For instance, it can be shown that if $f(p)$ takes values in a finite subset of the unit disc for all $p \in \mathbb{P}$, then $f \in \mathcal{M}_{conv}$, and if in addition $f(p) \neq 1$ for all $p \in \mathbb{P}$, then f is aperiodic.

• By taking averages in (3) and using that part (ii) of the previous result implies aperiodicity of f^j for $j = 1, \ldots, b-1$, we deduce that $d(S_{a,b}) = \frac{1}{b}$.

Proof. We prove (i). It suffices to show that for every Dirichlet character χ the multiplicative function $f \cdot \chi$ has a mean value. Note that χ takes values on roots of unity of fixed order for all but finitely many primes (on which it is 0). Hence, it suffices to show that if for some $m \in \mathbb{N}$ a multiplicative function g takes values on the m-th roots of unity for all but finitely many primes, then g has a mean value. So let g be such a multiplicative function. If $\mathbb{D}(g, n^{it}) = \infty$ for every $t \in \mathbb{R}$, then we are done by Theorem 2.9. Suppose that there exists $t \in \mathbb{R}$ such that $\mathbb{D}(g, n^{it}) < \infty$. Using that $\mathbb{D}(1, n^{it}) = \infty$ for every $t \neq 0$ we have by (12) that

$$m\mathbb{D}(g, n^{it}) \geq \mathbb{D}(g^m, n^{imt}) = \mathbb{D}(1, n^{imt}) + O(1) = \infty, \quad \text{for every } t \neq 0,$$

where the lower bound follows from (12). Hence, $t = 0$, which implies that

$$\sum_{p \in \mathbb{P}} \frac{1}{p} \left(1 - \mathrm{Re}(g(p))\right) < \infty.$$

Since g takes finitely many values on the unit disc, there exists $c > 0$ such that for all $p \in \mathbb{P}$ we either have $g(p) = 1$ or $1 - \mathrm{Re}(g(p)) \geq c$. Hence,

$$\sum_{p\in\mathbb{P}, g(p)\neq 1} \frac{1}{p} < \infty.$$

Since $|1 - g(p)| \leq 2$ for all $p \in \mathbb{P}$, we deduce that

$$\sum_{p\in\mathbb{P}} \frac{|1-g(p)|}{p} < \infty.$$

Theorem 2.9 again gives that g has a mean value, completing the proof of (i).

We prove (ii). Using the remark following Theorem 2.8, it suffices to show that for every $t \in \mathbb{R}$ and Dirichlet character χ we have $\mathbb{D}(f \cdot \chi, n^{it}) = \infty$. So let χ be a Dirichlet character. Then there exists $m \in \mathbb{N}$ such that $(\chi(p))^m = 1$ for all but a finite number of primes p. Then using (12) we get

$$m\mathbb{D}(f \cdot \chi, n^{it}) \geq \mathbb{D}(f^m \cdot \chi^m, n^{imt}) = \mathbb{D}(f^m, n^{imt}) + O(1) = \infty, \quad \text{for every } t \neq 0,$$

where the last distance is infinite since f^m is constant on primes and $mt \neq 0$. It remains to show that $\mathbb{D}(f \cdot \chi, 1) = \infty$. Suppose that χ has period d. Since $\chi(1) = 1$, we have $\chi(n) = 1$ whenever $n \equiv 1 \bmod d$, and since $f(p) = e(\alpha)$ for all $p \in \mathbb{P}$, we have

$$\mathbb{D}(f \cdot \chi, 1)^2 \geq (1 - \cos(2\pi\alpha)) \cdot \sum_{p\in\mathbb{P}\cap(d\mathbb{Z}+1)} \frac{1}{p} = \infty,$$

where we used that $(1 - \cos(2\pi\alpha)) \neq 0$ because $\alpha \notin \mathbb{Z}$ and the divergence of the last series follows from Dirichlet's theorem. This completes the proof of (ii). $\square$

3. Proof of main results

3.1. **Proof of Theorem 1.1.** We start with a few elementary lemmas.

Lemma 3.1. *Let (V_n) be a bounded sequence of elements of a normed space such that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} V_n = V.$$

For $N \in \mathbb{N}$ let $\widetilde{N} > N$ be integers such that the limit $\beta := \lim_{N\to\infty} \frac{N}{\widetilde{N}}$ exists. Then for every $\alpha \in \mathbb{R}$ we have

$$\frac{1}{N} \sum_{n=1}^{N} \mathrm{e}\big(n\frac{\alpha}{\widetilde{N}}\big)\, V_n = c \cdot V$$

where $c = \frac{1}{\beta}\int_0^\beta \mathrm{e}\big(\alpha y\big)\, dy$ if $\beta \neq 0$ and $c = 1$ if $\beta = 0$.

Proof. For $n \in \mathbb{N}$ let

$$S_n := \sum_{k=1}^{n} \big(V_k - V\big).$$

Our assumption gives that $S_n/n \to 0$ as $n \to \infty$. Using partial summation we get that the modulus of the average

$$\frac{1}{N} \sum_{n=1}^{N} \mathrm{e}\big(n\frac{\alpha}{\widetilde{N}}\big)\, \big(V_n - V\big)$$

is at most

$$\frac{1}{N}\Big(\sum_{n=2}^{N-1}\|S_n\|\,\big|\mathrm{e}\big((n+1)\frac{\alpha}{\widetilde{N}}\big)-\mathrm{e}\big(n\frac{\alpha}{\widetilde{N}}\big)\big|+\|S_N\|\Big)+o_{N\to\infty}(1).$$

Let $\varepsilon>0$. Since $S_n/n\to 0$ as $n\to\infty$, we have $|S_n|\le\varepsilon n$ for every sufficiently large n, and thus the last expression is bounded by

$$\frac{1}{N}\Big(\sum_{n=2}^{N-1}\varepsilon n\frac{|2\pi\alpha|}{\widetilde{N}}+\varepsilon N\Big)+o_{N\to\infty}(1)\le(\beta|\pi\alpha|+1)\varepsilon+o_{N\to\infty}(1).$$

Since ε is arbitrary, we get that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathrm{e}\big(n\frac{\alpha}{\widetilde{N}}\big)\,\big(V_n-V\big)=0.$$

Note also that if $\beta>0$, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathrm{e}\big(n\frac{\alpha}{\widetilde{N}}\big)=\frac{1}{\beta}\int_0^{\beta}\mathrm{e}\big(\alpha y\big)\,dy,$$

and the previous limit is 1 if $\beta=0$. Combining the above we get the asserted claim. □

Next we show that the discrete Fourier transform of elements of $\mathcal{M}_{conv}$ along certain "major arc" frequencies converges.

Lemma 3.2. *Let $f\in\mathcal{M}_{conv}$. Let $Q\in\mathbb{N}$, $p,\xi'\in\mathbb{Z}$, and*

$$\xi_N=\frac{p}{Q}N+\frac{\xi'}{Q},\quad N\in\mathbb{N}.$$

Then the averages

$$\frac{1}{N}\sum_{n=1}^{N}f(n)\,\mathrm{e}\big(-n\frac{\xi_N}{N}\big)\tag{13}$$

converge.

Proof. Notice first that the expression in (13) is equal to

$$\frac{1}{Q}\sum_{r=1}^{Q}\mathrm{e}\big(-r\frac{p}{Q}\big)\frac{1}{\lfloor N/Q\rfloor}\sum_{n=1}^{\lfloor N/Q\rfloor}f(Qn+r)\,\mathrm{e}\big(-(Qn+r)\frac{\xi'}{QN}\big)+o_{N\to\infty}(1).$$

Hence, it suffices to show that for every fixed Q,ξ', and $r\in[Q]$, the averages

$$\frac{1}{\lfloor N/Q\rfloor}\sum_{n=1}^{\lfloor N/Q\rfloor}f(Qn+r)\,\mathrm{e}\big(-(Qn+r)\frac{\xi'}{QN}\big)$$

converge. Since $\mathrm{e}(-r\xi'/(QN))\to 1$ as $N\to\infty$, it suffices to show that the averages

$$\frac{1}{\lfloor N/Q\rfloor}\sum_{n=1}^{\lfloor N/Q\rfloor}f(Qn+r)\,\mathrm{e}\big(-n\frac{\xi'}{N}\big)\tag{14}$$

converge. Since $f\in\mathcal{M}_{conv}$, we have that the averages

$$\frac{1}{N}\sum_{n=1}^{N}f(Qn+r)$$

converge, and using Lemma 3.1 for $a(n) := f(Qn+r)$, N replaced by $\lfloor N/Q \rfloor$, and $\widetilde{N}$ replaced by N, we deduce the needed convergence for the averages (14). This completes the proof. □

Next we analyze the asymptotic behavior of the relevant ergodic averages with weights given by the structured components (defined by Theorem 2.6) of an element of $\mathcal{M}_{conv}$.

Lemma 3.3. *Let $\theta > 0$, $Q \in \mathbb{N}$, and $f \in \mathcal{M}_{conv}$. For $N \in \mathbb{N}$ let $\widetilde{N} > N$ be a prime that satisfies $\widetilde{N} \equiv 1 \bmod Q$ and suppose that the limit $\beta := \lim_{N\to\infty} \frac{N}{\widetilde{N}}$ exists. Let $f_{\widetilde{N},\mathrm{st}} := f_{\widetilde{N}} * \phi_{\widetilde{N},\theta}$ where $\phi_{\widetilde{N},\theta}$ is defined by* (9) *and the convolution product is defined in $\mathbb{Z}_{\widetilde{N}}$. Then for every probability space $(X, \mathcal{X}, \mu)$, invertible commuting measure preserving transformations $T_1, \dots, T_\ell \colon X \to X$, functions $F_1, \dots, F_m \in L^\infty(\mu)$, and polynomials $p_{i,j} \colon \mathbb{Z} \to \mathbb{Z}$, $i = 1, \dots \ell$, $j = 1, \dots, m$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_{\widetilde{N},\mathrm{st}}(n) \cdot \Big(\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)}\Big)F_1 \cdot \dots \cdot \Big(\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)}\Big)F_m \tag{15}$$

converge in $L^2(\mu)$.

Proof. By the definition of $f_{\widetilde{N},\mathrm{st}}$ and $\phi_{\widetilde{N},\theta}$ we have that

$$f_{\widetilde{N},\mathrm{st}}(n) = \sum_{\xi \in \Xi_{\widetilde{N},\theta}} \widehat{f_{\widetilde{N}}}(\xi)\widehat{\phi_{\widetilde{N},\theta}}(\xi)\mathrm{e}\big(n\frac{\xi}{\widetilde{N}}\big), \quad n \in [\widetilde{N}],$$

where $\Xi_{\widetilde{N},\theta}$ is defined in (8). Since $\widetilde{N} \equiv 1 \bmod Q$, it follows from (10) and (11) that for $\widetilde{N} > 2QV$, if $\xi \in \Xi_{\widetilde{N},\theta}$, then ξ can be uniquely represented as

$$\xi = \frac{p}{Q}\widetilde{N} + \frac{\xi'}{Q}$$

for some $p \in \{0, \dots, Q-1\}$ and $\xi' \in \Xi'_{p,\theta}$ where for $p = 0, 1, \dots, Q-1$ we have

$$\Xi'_{p,\theta} := \big\{ -p + jQ \colon -V < j \le V \big\}.$$

Hence, it suffices to show that the averages (15) satisfy the asserted asymptotic when the (finite) sequence $(f_{\widetilde{N},\mathrm{st}}(n))_{n\in[N]}$ in (15) is replaced by the sequence

$$\widehat{f_{\widetilde{N}}}\big(\frac{p}{Q}\widetilde{N} + \frac{\xi'}{Q}\big) \cdot \widehat{\phi_{\widetilde{N},\theta}}\big(\frac{p}{Q}\widetilde{N} + \frac{\xi'}{Q}\big) \cdot \mathrm{e}\big(n\big(\frac{p}{Q} + \frac{\xi'}{Q\widetilde{N}}\big)\big), \quad n \in [N],$$

for all $p \in \{0, \dots, Q-1\}$ and $\xi' \in \Xi'_{p,\theta}$.

Recall that if $f \colon \mathbb{Z}_{\widetilde{N}} \to \mathbb{C}$ and $\xi \in \mathbb{Z}_{\widetilde{N}}$, then $\widehat{f}(\xi) = \frac{1}{\widetilde{N}}\sum_{n\in\mathbb{Z}_{\widetilde{N}}} f(n)\,\mathrm{e}\big(-n\frac{\xi}{\widetilde{N}}\big)$. By Lemma 3.2 the limit

$$\lim_{N\to\infty} \widehat{f_{\widetilde{N}}}\big(\frac{p}{Q}\widetilde{N} + \frac{\xi'}{Q}\big)$$

exists, and it follows from (9) that

$$\widehat{\phi_{\widetilde{N},\theta}}\big(\frac{p}{Q}\widetilde{N} + \frac{\xi'}{Q}\big) = 1 - \frac{\xi'}{QV}, \quad \text{whenever } \widetilde{N} \ge 2QV.$$

As a consequence, both terms can be factored out from the averaging operation.

It remains to deal with the term $\mathrm{e}\big(n(\frac{p}{Q}+\frac{\xi'}{Q\widetilde{N}})\big)$. Let

$$V_n := (\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)})F_1 \cdot \dots \cdot (\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)})F_m, \quad n \in \mathbb{N}.$$

By Theorem 2.2 we have that the averages

$$\frac{1}{N}\sum_{n=1}^{N} \mathrm{e}\big(n\,\frac{p}{Q}\big)\, V_n$$

converge in $L^2(\mu)$. Using Lemma 3.1 we deduce that for all $p, \xi' \in \mathbb{Z}$, $Q \in \mathbb{N}$, the averages

$$\frac{1}{N}\sum_{n=1}^{N} \mathrm{e}\big(n(\frac{p}{Q}+\frac{\xi'}{Q\widetilde{N}})\big)\, V_n$$

converge in $L^2(\mu)$. This completes the proof. □

We are ready now to prove Theorem 1.1.

Proof of Theorem 1.1. Let $\varepsilon > 0$. Without loss of generality we can assume that all functions are bounded by 1. We let $s \geq 2$ be the integer and C_s be the implicit constant defined in Lemma 2.3. We apply Lemma 2.4 for $\varepsilon/(2C_s)$ in place of ε and for $\kappa := 2$. We get that there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all integers $N, \widetilde{N}$ with $N_0 \leq sN \leq \widetilde{N} \leq 2sN$ and $f\colon \mathbb{Z}_{\widetilde{N}} \to \mathbb{C}$ with $|f| \leq 1$, the following implication holds:

$$\text{if} \quad \|f\|_{U^s(\mathbb{Z}_{\widetilde{N}})} \leq \delta, \quad \text{then} \quad \big\|\mathbf{1}_{[N]} \cdot f\big\|_{U^s(\mathbb{Z}_{sN})} \leq \frac{\varepsilon}{2C_s}. \tag{16}$$

We use the structural result of Theorem 2.6 for this δ in place of ε and for the previously defined s. We get that there exists $\theta = \theta(\delta, s) > 0$ such that for all large enough $N \in \mathbb{N}$, if $\widetilde{N}$ denotes the smallest prime such that $\widetilde{N} > sN$ and $\widetilde{N} \equiv 1 \mod Q$ (Q was introduced in Section 2.4 and depends only on θ), then we have the decomposition

$$f(n) = f_{\widetilde{N},\mathrm{st}}(n) + f_{\widetilde{N},\mathrm{un}}(n), \qquad n \in [\widetilde{N}], \tag{17}$$

where $f_{\widetilde{N},\mathrm{st}} = f_{\widetilde{N}} * \phi_{\widetilde{N},\theta}$ ($\phi_{\widetilde{N},\theta}$ is defined by (9)) and

$$\Big\|f_{\widetilde{N},\mathrm{un}}\Big\|_{U^s(\mathbb{Z}_{\widetilde{N}})} \leq \delta.$$

It follows from (16) that for all large enough N we have

$$\Big\|\mathbf{1}_{[N]} \cdot f_{\widetilde{N},\mathrm{un}}\Big\|_{U^s(\mathbb{Z}_{sN})} \leq \frac{\varepsilon}{2C_s}. \tag{18}$$

Note also that the prime number theorem on arithmetic progressions implies that

$$\lim_{N\to\infty} \frac{N}{\widetilde{N}} = \frac{1}{s}.$$

For $\ell, m \in \mathbb{N}$ let $(X, \mathcal{X}, \mu)$ be a probability space, let $T_1, \dots, T_\ell\colon X \to X$ be invertible commuting measure preserving transformations, let $F_1, \dots, F_m \in L^\infty(\mu)$, and let $p_{i,j}\colon \mathbb{Z} \to \mathbb{Z}$ be polynomials where $i = 1, \dots \ell$, $j = 1, \dots, m$. Let

$$V_n := (\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)})F_1 \cdot \dots \cdot (\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)})F_m, \quad n \in \mathbb{N}.$$

If $N \in \mathbb{N}$, for a given $a_N \colon [N] \to \mathbb{C}$ we define

$$A_N(a_N) := \frac{1}{N} \sum_{n=1}^{N} a_N(n) \cdot V_n.$$

Since $f_{\widetilde{N},\mathrm{un}}$ is bounded by 2 and the functions F_i are bounded by 1, it follows from Lemma 2.3 and (18) that

$$\limsup_{N\to\infty} \left\| A_N(f_{\widetilde{N},\mathrm{un}}) \right\|_{L^2(\mu)} \leq \varepsilon.$$

Hence,

$$\limsup_{N\to\infty} \left\| A_N(f) - A_N(f_{\widetilde{N},\mathrm{st}}) \right\|_{L^2(\mu)} \leq \varepsilon.$$

Furthermore, by Lemma 3.3 we have that the averages

$$A_N(f_{\widetilde{N},\mathrm{st}})$$

converge in $L^2(\mu)$ as $N \to \infty$. Combining the above we deduce that the sequence $(A_N(f))$ is Cauchy in $L^2(\mu)$ and hence it converges in $L^2(\mu)$. Therefore, the sequence $(f(n))$ is a good universal weight for polynomial multiple mean convergence.

Finally, we prove the last claim of Theorem 1.1. If the multiplicative function f is aperiodic, then by Theorem 2.7 we have that $\lim_{N\to\infty} \left\| \mathbf{1}_{[N]} \cdot f \right\|_{U^s(\mathbb{Z}_{sN})} = 0$ for every $s \geq 2$. Using Lemma 2.3 we deduce that the averages $A_N(f)$ converge to 0 in $L^2(\mu)$. This verifies the asserted convergence. □

3.2. **Proof of Theorem 1.2.**

Proof of part (i) *of Theorem* 1.2. Recall that the range of f is contained in a set of the form $R = \{1, \zeta, \dots, \zeta^{k-1}\}$ where ζ is a root of unity of order k. Then

$$\mathbf{1}_{f^{-1}(K)}(n) = \sum_{z\in K} \frac{1}{k} \sum_{j=0}^{k-1} z^{-j} (f(n))^j. \tag{19}$$

We establish the first claim of part (i). For $\ell, m \in \mathbb{N}$ let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \dots, T_\ell \colon X \to X$ be invertible commuting measure preserving transformations, $F_1, \dots, F_m \in L^\infty(\mu)$, $p_{i,j} \colon \mathbb{Z} \to \mathbb{Z}$ be polynomials where $i = 1, \dots, \ell$, $j = 1, \dots, m$. Using (19) we see that in order to verify the asserted convergence it suffices to show that for

$$V_n := (\prod_{i=1}^{\ell} T_i^{p_{i,1}(n)}) F_1 \cdot \dots \cdot (\prod_{i=1}^{\ell} T_i^{p_{i,m}(n)}) F_m, \quad n \in \mathbb{N},$$

the averages

$$\frac{1}{N} \sum_{n=1}^{N} (f(n))^j \, V_n$$

converge in $L^2(\mu)$ for $j = 0, \dots, k-1$. This follows from the first part of Theorem 1.1 and the fact that under the stated assumptions on the range of f we have $f^j \in \mathcal{M}_{conv}$ for $j = 0, \dots, k-1$ by part (i) of Proposition 2.10.

We establish now the second claim in part (i). Suppose that f^j is aperiodic for $j = 1, \dots, k-1$. By Proposition 2.5 it suffices to show that the set $f^{-1}(K)$ is Gowers uniform. Let $s \geq 2$ be an integer. We claim that

$$\lim_{N\to\infty} \left\| \mathbf{1}_{f^{-1}(K)} - \frac{|K|}{k} \right\|_{U^s(\mathbb{Z}_N)} = 0.$$

Using (19) and the triangle inequality for the U^s-norms, we see that in order to verify the claim it suffices to show that for $j = 1, \dots, k-1$ we have $\lim_{N\to\infty} \left\| f^j \right\|_{U^s(\mathbb{Z}_N)} = 0$ for every $s \in \mathbb{N}$. Since f^j is by assumption aperiodic for $j = 1, \dots, k-1$, this follows from Theorem 2.7. □

Proof of part (ii) *of Theorem* 1.2. We claim that if F is a Riemann integrable function on $\mathbb{T}$ with integral zero, then $\big((F \circ f)(n)\big)$ is a Gowers uniform sequence, meaning

$$\lim_{N\to\infty} \|F \circ f\|_{U^s(\mathbb{Z}_N)} = 0, \quad \text{for every } s \in \mathbb{N}. \tag{20}$$

Applying this for $F := \mathbf{1}_K - m_{\mathbb{T}}(K)$ and using that $m_{\mathbb{T}}(K) > 0$, we deduce that the set $f^{-1}(K)$ is Gowers uniform; hence by Proposition 2.5 it is a set of polynomial multiple recurrence and mean convergence.

We now verify (20). Without loss of generality we can assume that $\|F\|_\infty \leq 1/2$. Let $s \in \mathbb{N}$ and $\varepsilon > 0$.

We first claim that the sequence $(f(n))$ is equidistributed on the unit circle. Indeed, using Weyl's equidistribution criterion it suffices to show that for every non-zero $j \in \mathbb{Z}$ we have

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

This follows at once since by assumption f^j is aperiodic for $j \in \mathbb{N}$; hence it has average 0. Taking complex conjugates we get a similar property for all negative j as well.

Since F is Riemann integrable, bounded by $1/2$, and has zero mean, there exists a trigonometric polynomial P on $\mathbb{T}$, bounded by 1, with zero constant term, such that

$$\|F - P\|_{L^1(m_{\mathbb{T}})} \leq \big(\frac{\varepsilon}{2}\big)^{2^s}.$$

Since $(f(n))$ is equidistributed in $\mathbb{T}$ and the function $F - P$ is Riemann integrable, we deduce that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |F(f(n)) - P(f(n))| = \|F - P\|_{L^1(m_{\mathbb{T}})} \leq \big(\frac{\varepsilon}{2}\big)^{2^s}.$$

Using this, the fact that $|F - P|$ is bounded by 2, and the estimate

$$\|a\|_{U^s(\mathbb{Z}_N)}^{2^s} \leq \|a\|_{L^\infty(\mathbb{Z}_N)}^{2^s-1} \, \|a\|_{L^1(\mathbb{Z}_N)},$$

which can be easily proved using the inductive definition of the norms $U^s(\mathbb{Z}_N)$, we deduce that

$$\limsup_{N\to\infty} \|F \circ f - P \circ f\|_{U^s(\mathbb{Z}_N)} \leq \varepsilon. \tag{21}$$

We know by assumption that f^j is aperiodic for all $j \in \mathbb{N}$, and taking complex conjugates we get that f^j is aperiodic for all non-zero $j \in \mathbb{Z}$. Hence, Theorem 2.7

gives that $\lim_{N\to\infty} \|f^j\|_{U^s(\mathbb{Z}_N)} = 0$ for all non-zero $j \in \mathbb{Z}$. Since the trigonometric polynomial P has zero constant term, it follows by the triangle inequality that $\lim_{N\to\infty} \|P \circ f\|_{U^s(\mathbb{Z}_N)} = 0$. From this and (21) we deduce that

$$\limsup_{N\to\infty} \|F \circ f\|_{U^s(\mathbb{Z}_N)} \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we get $\lim_{N\to\infty} \|F \circ f\|_{U^s(\mathbb{Z}_N)} = 0$ and the proof is complete. □

3.3. **Proof of Corollaries 1.3 and 1.4.**

Proof of Corollary 1.3. Let $b \geq 2$ be an integer, ζ be a root of unity of order b, and $a \in \{0, \dots, b-1\}$.

In order to deal with the set $S_{\omega,A,b}$ we define the multiplicative function f_1 by $f_1(p^k) = \zeta$ for all $k \in \mathbb{N}$ and primes p. Using part (ii) of Proposition 2.10 we deduce that f_1^j is aperiodic for $j = 1, \dots, b-1$. Applying Theorem 1.2 for this multiplicative function and for $K := \{\zeta^a \colon a \in A\}$, we deduce the asserted claims for the set $S_{\omega,A,b}$.

In a similar fashion we prove the asserted claims for the set $S_{\Omega,A,b}$; the only difference is that we apply Theorem 1.2 for the multiplicative function f_2 defined by $f_2(p^k) = \zeta^k$, for all $k \in \mathbb{N}$ and primes p. □

Proof of Corollary 1.4. In order to deal with the set $S_{\Omega,A,\alpha}$ we define the multiplicative function f_1 by $f_1(p^k) = \mathrm{e}(\alpha)$ for all $k \in \mathbb{N}$ and primes p. Using part (ii) of Proposition 2.10 we deduce that f_1^j is aperiodic for all $j \in \mathbb{N}$. Applying Theorem 1.2 for this multiplicative function and for $K := \{\mathrm{e}(t) \colon t \in (-A) \cup A\}$, we deduce the asserted claims for the set $S_{\Omega,A,\alpha}$.

In a similar fashion, we prove the asserted claims for the set $S_{\Omega,A,\alpha}$; the only difference is that we apply Theorem 1.2 for the multiplicative function f_2 defined by $f_2(p^k) = \mathrm{e}(k\alpha)$ for all $k \in \mathbb{N}$ and primes p. □

3.4. **Extension to nilpotent groups.** Essentially the same arguments used in the previous subsections can be replicated in order to extend the main results of this article to the case where the transformations $T_1, \dots, T_\ell$ generate a nilpotent group. The only extra difficulty that we do not address here is to prove a variant of the uniformity estimates of Lemma 2.3 that deals with this more general setup. This requires a non-trivial modification of the PET induction argument used in [12, Lemma 3.5] along the lines of the argument used to prove [24, Theorem 4.2]. Assuming these estimates, substituting the convergence result of Theorem 2.2 with its nilpotent version (again due to M. Walsh) and the multiple recurrence result of Theorem 2.1 with a result of S. Leibman [22], the rest of the argument carries without any change.

3.5. **An alternate approach for weak convergence.** If one is satisfied with analyzing weak convergence of the multiple ergodic averages in our main results (which suffices for proving multiple recurrence), then an alternate way to proceed is as follows: Using the main result from [9] we get that sequences of the form $C(n) = \int F_0 \cdot T_1^n F_1 \cdots T_\ell^n F_\ell \, d\mu$, $n \in \mathbb{N}$, can be decomposed in two terms, one that is an ℓ-step nilsequence $(N(n))$ and another that contributes negligibly in evaluating weighted averages of the form $\frac{1}{N}\sum_{n=1}^N f(n)\, C(n)$. This reduces matters to analyzing the limiting behavior of averages of the form $\frac{1}{N}\sum_{n=1}^N f(n)\, N(n)$, a task

that has been carried out in [11]. This way one can prove a version of Theorem 1.1 and related corollaries that deal with weak convergence, avoiding the full strength of the main structural result in [11].

3.6. **An alternate approach for recurrence.** We mention here an alternate way to prove "linear" multiple recurrence results for shifts of the sets $S_{\omega,A,b}$, $S_{\Omega,A,b}$. After modifying the argument below along the lines of the proof of part (ii) of Theorem 1.2 we get similar results for the sets $S_{\omega,A,\alpha}$, $S_{\Omega,A,\alpha}$.

Definition. An IP_k*-set of integers* is a set of the form

$$\{a_{i_1} + \cdots + a_{i_\ell} \colon 1 \le \ell \le k,\ i_1 < i_2 < \cdots < i_\ell\}$$

where $a_1, \ldots, a_k$ are distinct positive integers.

For example, an IP_3-set has the form $\{m, n, r, m+n, m+r, n+r, m+n+r\}$ with $m, n, r \in \mathbb{N}$ distinct.

Proposition 3.4. *Let $d, \ell \in \mathbb{N}$ and let $L_1, \ldots, L_\ell \colon \mathbb{N}^d \to \mathbb{N}$ be pairwise independent linear forms. Let $b, c \in \mathbb{N}$ and $a \in \{0, \ldots, b-1\}$, and let $S_{a,b,c}$ be either $S_{\omega,a,b} + c$ or $S_{\Omega,a,b} + c$. Then the set*

$$\{\mathbf{m} \in \mathbb{N}^d \colon L_1(\mathbf{m}), \ldots, L_\ell(\mathbf{m}) \in S_{a,b,c}\} \tag{22}$$

has density $b^{-\ell}$. Therefore, the set $S_{a,b,c}$ contains IP_k-sets of integers for every $k \in \mathbb{N}$.

Remark. Similar results hold for affine linear forms and all sets defined in Theorem 1.2.

Proof. Note first that for $n > c$ we have

$$\mathbf{1}_{S_{a,b,c}}(n) = \frac{1}{b} \sum_{j=0}^{b-1} \zeta^{-aj} (f(n-c))^j \tag{23}$$

where f is the multiplicative function and ζ is the b-th root of unity defined in the proof of Corollary 1.3. Note also that the density of the set in (22) is equal to

$$\lim_{N\to\infty} \frac{1}{N^d} \sum_{\mathbf{m}\in[N]^d} \prod_{i=1}^{\ell} \mathbf{1}_{S_{a,b,c}}(L_i(\mathbf{m})); \tag{24}$$

the existence of the limit will be established momentarily. By [10, Theorem 1.1], if at least one of the functions $f_1, \ldots, f_\ell \in \mathcal{M}$ is aperiodic, then

$$\lim_{N\to\infty} \frac{1}{N^d} \sum_{\mathbf{m}\in[N]^d} \prod_{j=1}^{\ell} f_j(L_j(\mathbf{m})) = 0. \tag{25}$$

Using that f^j is aperiodic for $j = 1, \ldots, b-1$, in conjunction with (23) and (25), we deduce that the limit in (24) exists and is equal to $b^{-\ell}$. □

Theorem 3.5 (Furstenberg, Katznelson [15, Theorem 10.1])**.** *Let $T_1, \ldots, T_\ell$ be commuting measure preserving transformations acting on the same probability space $(X, \mathcal{X}, \mu)$. Then for every $A \in \mathcal{X}$ with $\mu(A) > 0$ there exists $k \in \mathbb{N}$ such that the set*

$$\{n \in \mathbb{N} \colon \mu(T_1^{-n} A \cap \cdots \cap T_\ell^{-n} A) > 0\}$$

intersects non-trivially every IP_k-set of integers.

Combining Proposition 3.4 with Theorem 3.5 we get that the set of return times in Theorem 3.5 intersects non-trivially each of the sets $S_{\omega,a,b,c}$ and $S_{\Omega,a,b,c}$. Unfortunately, a polynomial extension of Theorem 3.5 is not yet available, and we cannot get the full strength of the recurrence results of Corollaries 1.3 and 1.4 using such methods.

Acknowledgement

The authors would like to thank M. Lemanczyk for helpful remarks.

References

[1] E. el Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, *The Chowla and the Sarnak conjectures from ergodic theory point of view*, to appear in Discrete Contin. Dyn. Syst., arXiv:1410.1673.

[2] Antal Balog, Andrew Granville, and Kannan Soundararajan, *Multiplicative functions in arithmetic progressions* (English, with English and French summaries), Ann. Math. Qué. **37** (2013), no. 1, 3–30, DOI 10.1007/s40316-013-0001-z. MR3117735

[3] J. Bourgain, *Double recurrence and almost sure convergence*, J. Reine Angew. Math. **404** (1990), 140–161, DOI 10.1515/crll.1990.404.140. MR1037434

[4] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753, DOI 10.1090/S0894-0347-96-00194-4. MR1325795

[5] Hédi Daboussi, *Fonctions multiplicatives presque périodiques B* (French), Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974), Astérisque, No. 24–25, Soc. Math. France, Paris, 1975, pp. 321–324. D'après un travail commun avec Hubert Delange. MR0374074

[6] Hédi Daboussi and Hubert Delange, *Quelques propriétés des fonctions multiplicatives de module au plus égal à* 1 (French), C. R. Acad. Sci. Paris Sér. A **278** (1974), 657–660. MR0332702

[7] H. Davenport, *On some infinite series involving arithmetical functions II*, Quart. J. Math. Oxf. **8** (1937), 313–320.

[8] P. D. T. A. Elliott, *Probabilistic number theory. I: Mean-value theorems*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 239, Springer-Verlag, New York-Berlin, 1979. MR551361

[9] Nikos Frantzikinakis, *Multiple correlation sequences and nilsequences*, Invent. Math. **202** (2015), no. 2, 875–892, DOI 10.1007/s00222-015-0579-7. MR3418246

[10] Nikos Frantzikinakis and Bernard Host, *Asymptotics for multilinear averages of multiplicative functions*, Math. Proc. Cambridge Philos. Soc. **161** (2016), no. 1, 87–101, DOI 10.1017/S0305004116000116. MR3505671

[11] Nikos Frantzikinakis and Bernard Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. **30** (2017), no. 1, 67–157, DOI 10.1090/jams/857. MR3556289

[12] Nikos Frantzikinakis, Bernard Host, and Bryna Kra, *The polynomial multidimensional Szemerédi theorem along shifted primes*, Israel J. Math. **194** (2013), no. 1, 331–348, DOI 10.1007/s11856-012-0132-y. MR3047073

[13] Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR0498471

[14] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979), DOI 10.1007/BF02790016. MR531279

[15] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168, DOI 10.1007/BF02792547. MR833409

[16] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588, DOI 10.1007/s00039-001-0332-9. MR1844079

[17] Andrew Granville and K. Soundararajan, *Large character sums: pretentious characters and the Pólya-Vinogradov theorem*, J. Amer. Math. Soc. **20** (2007), no. 2, 357–384, DOI 10.1090/S0894-0347-06-00536-4. MR2276774

[18] A. Granville and K. Soundararajan, *Multiplicative number theory: The pretentious approach*, book manuscript in preparation.

[19] G. Halász, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen* (German), Acta Math. Acad. Sci. Hungar. **19** (1968), 365–403. MR0230694

[20] R. R. Hall, *A sharp inequality of Halász type for the mean value of a multiplicative arithmetic function*, Mathematika **42** (1995), no. 1, 144–157, DOI 10.1112/S0025579300011426. MR1346679

[21] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), no. 1-2, 223–225, DOI 10.1007/BF01949145. MR836415

[22] A. Leibman, *Multiple recurrence theorem for measure preserving actions of a nilpotent group*, Geom. Funct. Anal. **8** (1998), no. 5, 853–931, DOI 10.1007/s000390050077. MR1650102

[23] Terence Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 657–688, DOI 10.1017/S0143385708000011. MR2408398

[24] Miguel N. Walsh, *Norm convergence of nilpotent ergodic averages*, Ann. of Math. (2) **175** (2012), no. 3, 1667–1688, DOI 10.4007/annals.2012.175.3.15. MR2912715

[25] E. Wirsing, *Das asymptotische Verhalten von Summen über multiplikative Funktionen. II* (German), Acta Math. Acad. Sci. Hungar. **18** (1967), 411–467. MR0223318

Department of Mathematics, Voutes University Campus, University of Crete, Heraklion 71003, Greece

E-mail address: `frantzikinakis@gmail.com`

Université Paris-Est Marne-la-Vallée, Laboratoire d'analyse et de mathématiques appliquées, UMR CNRS 8050, 5 Bd Descartes, 77454 Marne la Vallée Cedex, France

E-mail address: `bernard.host@u-pem.fr`