MÖBIUS ORTHOGONALITY OF SEQUENCES WITH MAXIMAL ENTROPY

MICHAEL DRMOTA, CHRISTIAN MAUDUIT, JOËL RIVAT, AND LUKAS SPIEGELHOFER

ABSTRACT. We prove that strongly b-multiplicative functions of modulus 1 along squares are asymptotically orthogonal to the Möbius function. This provides examples of sequences having maximal entropy and satisfying this property.

1. INTRODUCTION

Sarnak's conjecture [33, 34] is concerned with the Möbius μ-function, defined by $\mu(n) = (-1)^{\omega(n)}$ is n is squarefree, and $\mu(n) = 0$ otherwise, where $\omega(n)$ is the number of different prime factors of n. It can also be defined as the Dirichlet inverse of the constant function 1. Sarnak's conjecture states that every bounded *deterministic* sequence $f : \mathbb{N} \to \mathbb{C}$ is orthogonal to the Möbius function,

$$\sum_{n<N} \mu(n) f(n) = o(N).$$

Deterministic sequences f can be defined by the property that for all $\varepsilon > 0$, the set of k-tuples

$$\left\{ \big(f(n+0), \ldots, f(n+k-1) \big) : n \geq 0 \right\} \subseteq \mathbb{C}^k$$

can be covered by $\exp(o(k))$ many balls of radius ε, as k goes to infinity. For functions f having values in a finite set, it is equivalent to demand that f has subexponential *factor complexity* p_k: the number of contiguous finite subsequences of f of length k should be bounded by $\exp(o(k))$. For example, this is the case for all *automatic sequences* [1], which have a factor complexity bounded by Ck (where $C > 0$ is a constant depending on the sequence), and Sarnak's conjecture has been verified for this class of sequences by Müllner [29]. Sarnak's conjecture has been verified for other classes of sequences, see for example [2, 3, 4, 5, 6, 16, 17, 18, 19, 12, 20, 14, 15, 21, 23, 24, 32, 35, 36].

In this work, we are concerned with Möbius orthogonality for *non-deterministic* sequences — in particular, we are concerned with the *normal sequence* $\mathbf{t}(n^2)$, where $\mathbf{t}$ is the *Thue–Morse sequence*.

It is known that there exist (many) normal sequences that are Möbius disjoint; in particular, *each* measure-theoretic dynamical system $(X, \mathcal{B}, \lambda, T)$ is almost everywhere Möbius orthogonal: for each $f \in L^1(X)$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \leq N} f(T^n x) \mu(n) = 0 \tag{1}$$

for almost all $x \in X$ [33]; see [12] for a proof and [11] for a polynomial extension. Considering the Bernoulli shift on $\{0,1\}^{\mathbb{Z}}$ we obtain many normal sequences with the desired property. Moreover, Möbius orthogonality for dynamical systems having large positive entropy (close to the maximal value) was considered recently by Downarowicz and Serafin [7, 8]. The added value of our paper lies in an explicit, simple construction of a normal number that is Möbius orthogonal. We thank Mariusz Lemańczyk for pointing out this remark to us.

2000 *Mathematics Subject Classification.* Primary: 11A63, 11L03, 11N05, Secondary: 11N60, 11L20, 60F05.

Key words and phrases. sum-of-digits function, primes, exponential sums.

The first and fourth author is supported by the Austrian Science Foundation FWF, SFB F55-02 "Subsequences of Automatic Sequences and Uniform Distribution". This work was supported by the joint ANR-FWF-project ANR-14-CE34-0009, I-1751 MuDeRa.

The Thue–Morse sequence can be defined via the binary sum-of-digits function s_2, which counts the number of powers of two needed to represent a natural number as their sum. We define $\mathbf{t}(n) = (-1)^{s_2(n)}$, which is the Thue–Morse sequence on the two symbols $1, -1$. This sequence is automatic and as such has factor complexity $p_k \le Ck$; however, when we extract the subsequence along the squares, the resulting sequence is normal. That is, each finite word of length k on $\{1, -1\}$ occurs with asymptotic frequency 2^{-k} along this subsequence. This has been proved by the first three authors [9], strengthening a result of Moshe [28], who showed that each block $b \in \{1, -1\}^k$ occurs at least once in $\mathbf{t}(n^2)$.

Besides providing an example of a sequence having maximal topological entropy and being orthogonal to μ, our interest in the sum $\sum_{n<N} \mu(n)\mathbf{t}(n^2)$ has its origin in the study of the digits of prime numbers.

The second and third authors [26] proved in particular that the base-b sum-of-digits of prime numbers is uniformly distributed in residue classes; this was accomplished by studying the sum $\sum_{n<N} \Lambda(n) \exp(2\pi i\vartheta s_b(n))$, where Λ is the von Mangoldt function (defined by $\Lambda(n) = \log p$ if $n = p^k$ with $k \in \mathbb{N}$, $k \ge 1$ and $\Lambda(n) = 0$ otherwise). Moreover, the same authors [25] studied the sum of digits of the sequence of squares. It is therefore a natural problem to attack the sum of digits of *squares of primes*; for the Thue–Morse sequence, this can be accomplished by studying the sum $\sum_{n<N} \Lambda(n)\mathbf{t}(n^2)$. For the time being, we do not have a solution for this problem; a replacement is the (easier) sum $\sum_{n<N} \mu(n)\mathbf{t}(n^2)$.

1.1. **Notation.** We denote by $\mathbb{N}$ the set of non-negative integers, and by $\mathbb{U}$ the set of complex numbers of modulus 1. For $n \in \mathbb{N}$, $n \ge 1$, we denote by $\tau(n)$ the number of divisors of n, by $\omega(n)$ the number of distinct prime factors of n, by and by $\mu(n)$ the Möbius function (defined by $\mu(n) = (-1)^{\omega(n)}$ if n is squarefree and $\mu(n) = 0$ otherwise).

For $x \in \mathbb{R}$ we denote by $\pi(x)$ the number of prime numbers less or equal to x, by $\|x\|$ the distance of x to the nearest integer, and we set $\mathrm{e}(x) = \exp(2i\pi x)$. If f and g are two functions taking strictly positive values such that f/g is bounded, we write $f = O(g)$ or $f \ll g$.

Furthermore let $s_2(n)$ denote the binary sum-of-digits function.

1.2. **Main Result.** Let $t(n) = s_2(n) \bmod 2$ denote the Thue–Morse sequence on the alphabet $\{0, 1\}$. It has been shown by the three first authors that subsequence $t(n^2)$ is a normal sequence, that is, each binary block $B \in \{0, 1\}^L$, $L \ge 1$, appears with asymptotic frequency 2^{-L} as a factor in $t(n^2)$. In particular this shows that $t(n^2)$ has maximal (positive) entropy $\log 2$.

The main purpose of this paper is to show that $t(n^2)$ is orthogonal to the Möbius function.

Theorem 1. *Let $t(n)$ denote the Thue–Morse sequence. Then we have, as $N \to \infty$,*

$$\sum_{n<N} \mu(n)t(n^2) = o(N). \tag{2}$$

This is actually not the first explicit example of a positive entropy sequence that is orthogonal to the Möbius function. A previously considered example is given by the sequence $\mu(n)^2$ (detecting the square-free integers), which has topological entropy $\frac{6}{\pi^2} \log 2$([31, 33], see also [10, 13]) and obviously $\mu(n)^2$ is orthogonal to $\mu(n)$.

Nevertheless, our result is one of the first explicit examples of a (binary) sequence with maximal entropy $\log 2$ that has this orthogonality property.

This kind of examples is in particular interesting in view of the Sarnak conjecture [33, 34] which says that every bounded zero entropy sequence is orthogonal to the Möbius function.

In this context, we note that due to normality, the *symbolic dynamical system* $(X, \mathcal{B}, \lambda, T)$ defined by $t(n^2)$ is the full shift: clearly, there exist sequences $x \in X$ that are not orthogonal to the Möbius function. On the other hand, note that Downarowicz and Serafin [7, 8] study dynamical systems with entropy close to the maximum and still obtain Möbius orthogonality.

The Thue–Morse sequence is sometimes defined by $g(n) = (-1)^{s_2(n)}$. That is, the values $0, 1$ are replaced by 1 and -1. Since $g(n) = 1-2t(n)$ and $\sum_{n<N} \mu(n) = o(N)$ the relation (2) is equivalent to

$$\sum_{n<N} \mu(n)(-1)^{s_2(n)} = o(N). \tag{3}$$

The function $g(n) = (-1)^{s_2(n)}$ is a so-called *strongly 2-multiplicative function.* More generally, a strongly b-multiplicative functions (where $b \ge 2$ is a fixed integer) is defined by the relation

$$g(kb+a) = g(k)g(a) \qquad (a, k \in \mathbb{N},\, 0 \le a < b).$$

Actually, Theorem 1 can be generalized to all complex valued strongly b-multiplicative functions of modulus 1.

Theorem 2. *Let $b \ge 2$ be a given integer. Then for all complex valued strongly b-multiplicative functions $g(n)$ of modulus 1 we have*

$$\sum_{n<N} \mu(n) g(n^2) = o(N). \tag{4}$$

Note that this theorem gives many more examples of sequences with maximal entropy that are orthogonal to Möbius: Müllner [30] proved in particular that q-multiplicative functions with values in $\{\exp(2\pi i j/m) : 0 \le j < m\}$ are normal along the squares under certain weak conditions.

1.3. **Strongly b-Multiplicative Functions.** It is clear that strongly b-multiplicative functions $g(n)$ of modulus 1 satisfy $g(0) = 1$ and that $g(1), \ldots, g(b-1)$ determine all other values of $g(n)$:

$$g(n) = \prod_{j \ge 0} g(\varepsilon_j) \quad \text{with} \quad n = \sum_{j \ge 0} \varepsilon_j b^j.$$

We will distinguish between two different classes of b-multiplicative functions, namely periodic ones and non-periodic ones.

Proposition 1. *A b-multiplicative function g having values in $\{z \in \mathbb{C} : |z| = 1\}$ is periodic if and only if*

$$g(\ell) = g(1)^\ell \quad (0 \le \ell \le b-1) \quad \textit{and} \quad g(b-1) = 1. \tag{5}$$

While the difficult part of the proof (the "only if"-part) of this statement rests on Proposition 2 proved later, we will not use this direction in the sequel and thus there is no circular argument involved.

Proof. Suppose first that (5) holds, that is, $g(\ell) = \mathrm{e}(\ell j_0/(b-1))$ for some integer j_0. Then

$$g(n) = \mathrm{e}(n j_0/(b-1))$$

for all $n \ge 0$. This follows from the fact that $\mathrm{e}(b^j/(b-1)) = \mathrm{e}(1/(b-1))$. This means that in this case $g(n)$ is periodic with a period dividing $b-1$. Conversely, suppose, in order to obtain a contradiction, that (5) is not satisfied and that g is periodic with period L. By Proposition 2 below we have

$$F_\lambda(h) = o(1) \tag{6}$$

as $\lambda \to \infty$, for all $h \in \mathbb{Z}$. By periodicity,

$$F_\lambda(h) = \frac{1}{b^\lambda} \sum_{0 \le u < b^\lambda} g(u)\, \mathrm{e}(-hu/L) = \mathcal{O}(L/b^\lambda) + \frac{1}{L} \sum_{0 \le u < L} g(u)\, \mathrm{e}(-hu/L)$$

and (6) implies that $\sum_{0\le u<L} g(u)\,\mathrm{e}(-hu/L)=0$ for $0\le h<L$. By inversion, we obtain $g(u)=0$ for all u, which contradicts $|g(u)|=1$. This completes the proof. □

1.4. **Plan of the Proofs.** If $g(n)$ is periodic then Dirichlet's prime number theorem implies Theorem 2. Hence, it is sufficient to suppose that $g(n)$ is not periodic.

In order to prove Theorem 2 we apply the Daboussi–Kátai criterion (Lemma 5 below). This criterion says that

$$\sum_{n<N} g(p^2n^2)\overline{g(q^2n^2)} = o(N), \tag{7}$$

where p,q are different (and sufficiently large) prime numbers, implies Theorem 2.

At this stage we will apply a general theorem by the second and third authors [27] that gives sufficient conditions for functions $f(n)$ (with $|f(n)|\le 1$) such that

$$\sum_{n<N} f(n^2)\,\mathrm{e}(\theta n) = o(N).$$

In our case we want to apply this theorem for

$$f(n) = g(p^2n)\overline{g(q^2n)}$$

and $\theta=0$. In particular one has to check a *carry property* and a *Fourier property*. In our case the carry property is easy to check (see Section 3), whereas the Fourier property needs non-trivial bounds for the Fourier-terms

$$F_\lambda(t) = \frac{1}{b^\lambda}\sum_{0\le u<b^\lambda} f(u)\,\mathrm{e}(-ut) = \frac{1}{b^\lambda}\sum_{0\le u<b^\lambda} g(p^2u)\overline{g(q^2u)}\,\mathrm{e}(-ut),$$

We will derive the necessary bounds in Section 2. This will be then the main ingredient for the proof of Theorem 2 which will be summarized in Section 3.

2. Fourier bounds

In this section, we are concerned with strongly b-multiplicative functions $g:\mathbb{N}\to\mathbb{U}$ that do not satisfy (5). In the main result of this section, Proposition 2 below, we will prove that they possess Fourier coefficients $F_\lambda(t)$ that converge to zero uniformly in t.

We suppose that P,Q are positive and coprime integers that are also coprime to b — later we will apply our results for $P=p^2$ and $Q=q^2$, where p,q are different primes. In order to obtain upper bounds for $F_\lambda(t)$ we define more generally

$$F_\lambda^{i,j}(t) = \frac{1}{b^\lambda}\sum_{0\le u<b^\lambda} g(Pu+i)\overline{g(Qu+j)}\,\mathrm{e}(-ut),$$

where $0\le i\le P-1$ and $0\le j\le Q-1$.

The following recurrence follows directly from the definition.

Lemma 1. *Suppose that P,Q are positive and coprime integers that are also coprime to b and that $0\le i\le P-1$, $0\le j\le Q-1$, and $\lambda\ge 1$. Then we have for all $t\in\mathbb{R}$*

$$F_\lambda^{i,j}(t) = \frac{1}{b}\sum_{r=0}^{b-1} g(Pr+i \bmod b)\overline{g(Qr+j \bmod b)}\,\mathrm{e}(-rt)\,F_{\lambda-1}^{\lfloor\frac{i+rP}{b}\rfloor,\lfloor\frac{j+Qr}{b}\rfloor}(bt). \tag{8}$$

Proof. By distinguishing between residue classes modulo b we obtain

$$F_\lambda^{i,j}(t) = \frac{1}{b^\lambda}\sum_{r=0}^{b-1}\sum_{0\le u<b^{\lambda-1}} g(P(bu+r)+i)\overline{g(Q(bu+r)+j)}\,\mathrm{e}(-(bu+r)t)$$
$$= \frac{1}{b^\lambda}\sum_{r=0}^{b-1} g(Pr+i \bmod b)\overline{g(Qr+j \bmod b)}\,\mathrm{e}(-rt)$$
$$\times \sum_{0\le u<b^{\lambda-1}} g(bPu+b\lfloor (Pr+i)/b\rfloor)\overline{g(bQu+b\lfloor (Qr+j)/b\rfloor)}\,\mathrm{e}(-but)$$
$$= \frac{1}{b}\sum_{r=0}^{b-1} g(Pr+i \bmod b)\overline{g(Qr+j \bmod b)}\,\mathrm{e}(-rt)F_{\lambda-1}^{\lfloor\frac{i+rP}{b}\rfloor,\lfloor\frac{j+Qr}{b}\rfloor}(bt).$$

□

Actually we are interested in the behaviour of $F_\lambda^{0,0}(t) = F_\lambda(t)$. Thus, we have to study the action

$$T:(i,j)\to\left\{\left(\left\lfloor\frac{i}{b}\right\rfloor,\left\lfloor\frac{j}{b}\right\rfloor\right),\left(\left\lfloor\frac{i+P}{b}\right\rfloor,\left\lfloor\frac{j+Q}{b}\right\rfloor\right),\ldots,\left(\left\lfloor\frac{i+(b-1)P}{b}\right\rfloor,\left\lfloor\frac{j+(b-1)Q}{b}\right\rfloor\right)\right\},$$

where we start with $(0,0)$. In this context it is convenient to consider the di-graph D with vertices (i,j) $(0\le i\le P-1,\ 0\le j\le Q-1)$ and edges

$$(i,j)\to\left(\left\lfloor\frac{i}{b}\right\rfloor,\left\lfloor\frac{j}{b}\right\rfloor\right),\ (i,j)\to\left(\left\lfloor\frac{i+P}{b}\right\rfloor,\left\lfloor\frac{j+Q}{b}\right\rfloor\right),\ldots,(i,j)\to\left(\left\lfloor\frac{i+(b-1)P}{b}\right\rfloor,\left\lfloor\frac{j+(b-1)Q}{b}\right\rfloor\right).$$

Lemma 2. *Let $\mathcal{C}$ denote the strongly connected component of the di-graph D that contains $(0,0)$. Then $\mathcal{C}$ contains precisely $P+Q-1$ elements that can be also represented by*

$$\mathcal{C} = \{(\lfloor tP\rfloor,\lfloor tQ\rfloor) : 0\le t<1\}.$$

In particular if $(i,j)\in\mathcal{C}$ and $(i,j)\ne(P-1,Q-1)$ then either $(i+1,j)\in\mathcal{C}$ or $(i,j+1)\in\mathcal{C}$.

Proof. Clearly we have $(0,0)\in\{(\lfloor tP\rfloor,\lfloor tQ\rfloor) : 0\le t<1\}$; we just have to set $t=0$.

Next we show that for $0\le t<1$ and for integers $0\le r\le b-1$

$$\left\lfloor\frac{\lfloor tP\rfloor+rP}{b}\right\rfloor = \left\lfloor\frac{t+r}{b}P\right\rfloor. \tag{9}$$

For this purpose we write $tP=\lfloor tP\rfloor+y$ with $0\le y<1$. This gives

$$\frac{\lfloor tP\rfloor+rP}{b} = \frac{t+r}{b}P-\frac{y}{b}.$$

Suppose now that for some integer m

$$m\le\frac{t+r}{b}P<m+\frac{1}{b}. \tag{10}$$

Equivalently this means that

$$bm-rP\le t<bm-rP+1.$$

Since $0\le t<1$ this implies that $bm=rP$. If $1\le r\le b-1$ this is impossible since b and P are coprime. Thus we either have $r=m=0$, that is, $tP<1$ or (10) does not hold. In the first case we have $\lfloor tP\rfloor=0$ (or $y=\lfloor tP\rfloor$) and consequently (9) is just the trivial identity $0=0$ (note that $r=0$). In the second case (where (10) does not hold) we clearly have

$$\left\lfloor\frac{t+r}{b}P-\frac{y}{b}\right\rfloor = \left\lfloor\frac{t+r}{b}P\right\rfloor$$

so that (9) holds, too.

Clearly (9) remains true if we replace P by Q. Hence, if (i,j) is represented by $(i,j) = (\lfloor tP \rfloor, \lfloor tQ \rfloor)$ then for every $0 \le r \le b-1$ we have

$$\left(\left\lfloor \frac{i+rP}{b} \right\rfloor, \left\lfloor \frac{j+rQ}{b} \right\rfloor \right) = \left(\left\lfloor \frac{t+r}{b} P \right\rfloor, \left\lfloor \frac{t+r}{b} Q \right\rfloor \right).$$

Note that $0 \le \frac{t+r}{b} < 1$ so that we stay in the same set. Furthermore if we start with $(0,0)$ represented by $t=0$ then by repeated application it follows that we can reach any pair of the kind

$$(i,j) = \left(\left\lfloor \frac{r_L + r_{L-1}b + \cdots r_1 b^{L-1}}{b^L} P \right\rfloor, \left\lfloor \frac{r_L + r_{L-1}b + \cdots r_1 b^{L-1}}{b^L} Q \right\rfloor \right).$$

Actually this is sufficient to reach all elements of $\{(\lfloor tP \rfloor, \lfloor tQ \rfloor) : 0 \le t < 1\}$. Since P and Q are coprime the line $\{(tP, tQ) : 0 \le t < 1\}$ does not meet a lattice point different from $(0,0)$. Consequently line is cut into $P+Q-1$ intervals that correspond to its $P+Q-1$ elements. In each of this interval we could restrict ourselves to b-adic rational numbers t. This means that starting with $(0,0)$ we can reach every element of $\{(\lfloor tP \rfloor, \lfloor tQ \rfloor) : 0 \le t < 1\}$. Conversely if we start with the pair $(\lfloor tP \rfloor, \lfloor tQ \rfloor)$ and if L is large enough then

$$\left(\left\lfloor \frac{t}{b^L} P \right\rfloor, \left\lfloor \frac{t}{b^L} Q \right\rfloor \right) = (0,0).$$

Summing up this means that we have actually described the strongly connected component of the di-graph D that contains $(0,0)$. ☐

As a corollary we obtain the following property that will be crucial for the proof of a non-trivial upper bound of $F_\lambda^{0,0}(t)$.

Corollary 1. *Let $\mathcal{C}$ be as above and assume that $b < P < Q$. Then there exists $i_0 < b$ such that $\{(i_0, b-1), (i_0, b)\} \subseteq \mathcal{C}$.*

Proof. By Lemma 2, $\mathcal{C}$ can be considered as a lattice *path* from $(0,0)$ to $(P-1, Q-1)$ that is close to the diagonal and has only steps of the form $(i,j) \to (i+1, j)$ and $(i,j) \to (i, j+1)$. Thus, there is a unique step of the form $(i_0, b-1) \to (i_0, b)$. Since $P < Q$ it follows that $i_0 \le b-1$. ☐

Next we use the relation (8) to obtain proper vector recurrences for $F_\lambda^{i,j}(t)$. Set

$$\mathbf{F}_\lambda(t) = \left(F_\lambda^{i,j}(t) \right)_{(i,j) \in \mathcal{C}}$$

and

$$\mathbf{A}(t) = \left(a_{(i,j),(i',j')}(t) \right)_{(i,j),(i',j') \in \mathcal{C}}$$

where

$$a_{(i,j),(i',j')}(t) = \begin{cases} \frac{1}{b} g(Pr + i \bmod b) \overline{g(Qr + j \bmod b)}\, \mathrm{e}(-rt) & \text{for } (i',j') = \left(\left\lfloor \frac{i+rP}{b} \right\rfloor, \left\lfloor \frac{j+rQ}{b} \right\rfloor \right), \\ 0 & \text{else.} \end{cases}$$

Then (8) rewrites to

$$\mathbf{F}_\lambda(t) = \mathbf{A}(t) \cdot \mathbf{F}_{\lambda-1}(bt).$$

Thus, we are led to study the product of matrices $\mathbf{A}(t) \cdot \mathbf{A}(bt) \cdots \mathbf{A}(b^L t)$.

Let $\| \cdot \|$ denote that row-sum-norm of a matrix. Then we have the following property.

Lemma 3. *Suppose that $g(n)$ is non-periodic. There exist $L > 0$ and $\delta > 0$ such that*

$$\sup_{t \in \mathbb{R}} \| \mathbf{A}(t) \cdot \mathbf{A}(bt) \cdots \mathbf{A}(b^L t) \| \le 1 - \delta. \tag{11}$$

Proof. We interpret the entries of the matrix $\mathbf{A}(t)\cdot\mathbf{A}(bt)\cdots\mathbf{A}(b^L t)$ in the following way. Let $D_C(t)$ be the strongly connected subgraph of D corresponding to the vertex set $\mathcal{C}$, where the edges of $D_C(t)$

$$(i,j) \to \left(\left\lfloor \frac{i+rP}{b} \right\rfloor, \left\lfloor \frac{j+rQ}{b} \right\rfloor \right), \qquad 0 \le r \le b-1,$$

are labelled by

$$a_{(i,j),(\langle(i+rP)/b\rangle,\langle(j+rQ)/b\rangle)}(t) = \frac{1}{b} g(Pr+i \bmod b)\overline{g(Qr+j \bmod b)}\, \mathrm{e}(-rt).$$

If $(e_0, e_1, \ldots, e_L)$ be a directed path in D such that e_j is actually an edge in $D_C(b^j t)$, $0 \le j \le L$, then we define the weight w of this path by

$$w(e_0, e_1, \ldots, e_L) = a_{e_0}(t) a_{e_1}(bt) \cdots a_{e_L}(b^L t).$$

Note that $|w(e_0, e_1, \ldots, e_L)| = b^{-L-1}$. It the entries of $\mathbf{A}(t)\cdot\mathbf{A}(bt)\cdots\mathbf{A}(b^L t)$ are denoted by $b_{L+1;(i,j),(i'j')}(t)$ then we have by definition

$$b_{L+1;(i,j),(i'j')}(t) = \sum w(e_0, e_1, \ldots, e_L),$$

where the sum is taken over all directed paths $(e_0, e_1, \ldots, e_L)$ in D that connect (i,j) and (i',j') such that e_j is an edge in $D_C(b^j t)$, $0 \le j \le L$.

For $(i,j),(i',j') \in \mathcal{C}$ let $B_{L+1}(i,j),(i'j')$ denote the number of different paths from (i,j) to (i',j'). Clearly we have

$$\sum_{(i',j')\in\mathcal{C}} B_{L+1}((i,j),(i'j')) = b^{L+1}.$$

Hence

$$\sum_{(i',j')\in\mathcal{C}} \left| b_{(i,j),(i'j')}(t) \right| \le b^{L-1} \sum_{(i',j')\in\mathcal{C}} B_{L+1}((i,j),(i'j')) = 1.$$

Note that this just says that $\|\mathbf{A}(t)\cdot\mathbf{A}(bt)\cdots\mathbf{A}(b^L t)\| \le 1$ Furthermore, in order to prove (11) we just have to show that for every $(i,j)\in\mathcal{C}$ there exists $(i',j')\in\mathcal{C}$ with

$$\left| b_{L+1;(i,j),(i'j')}(t) \right| < b^{-L-1} B_{L+1}((i,j),(i'j')). \tag{12}$$

In order to prove (12) we proceed in two steps. We first show that there exist $L \ge 1$ such that

$$\left| b_{L+1;(i_0,b-1),(0,0)}(t) \right| < b^{-L-1} B_{L+1}((i_0,b-1),(0,0)) \tag{13}$$

or

$$\left| b_{(i_0,b),(0,0)}(t) \right| < b^{-L-1} B_{L+1}((i_0,b),(0,0)). \tag{14}$$

Since $D_C(t)$ is strongly connected it is clear that these properties imply (12) for some $L > 0$. We first define define L_1 the minimal n such that for every pair $((i,j),(i',j'))\in\mathcal{C}$ there exists a path of length L_1 that connects (i,j) and (i',j'). (Since there is s loop from $(0,0)$ to itself, there are such n.) Second we define L_2 as the smallest L such that the above construction works. Then for every $(i,j)\in\mathcal{C}$ there are two paths p_1,p_2 of length L_1 that connect (i,j) to $(0,1)$ and (i,j) to $(0,2)$, respectively. This shows that $B_{L_1}((i,j),(0,1)) > 0$ and $B_{L_1}((i,j),(0,2)) > 0$. Consequently we have

$$\begin{aligned}\left|b_{L_1+L_2+1;(i,j),(0,0)}(t)\right| &= \left|\sum_{(i',j')\in\mathcal{C}} b_{L_1;(i,j),(i',j')}(t) b_{L_2+1;(i',j'),(0,0)}(b^{L_1}t)\right| \\ &< b^{-L_1-L_2-1} \sum_{(i',j')\in\mathcal{C}} B_{L_1}((i,j),(i',j'))B_{L_2+1}((i',j'),(0,0)) \\ &= b^{-L_1-L_2-1} B_{L_1+L_1+1}((i,j),(0,1))\end{aligned}$$

since $(i',j') = (i_0, b-1)$ or $(i',j') = (i_0,b)$ appears in this sum with a non-zero contribution, and so (12) follows.

We fix some $1 \le r \le b-1$ and consider two paths from $(i_0, b-1)$ to $(0,0)$ and two from (i_0,b) to $(0,0)$, respectively:

$$\begin{aligned}&(i_0, b-1) \to \left(\left\lfloor\frac{i_0+rP}{b}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b}\right\rfloor\right) \to \left(\left\lfloor\frac{i_0+rP}{b^2}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b^2}\right\rfloor\right) \to \cdots \\ &\qquad \to \left(\left\lfloor\frac{i_0+rP}{b^L}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b^L}\right\rfloor\right) = (0,0), \\ &(i_0, b-1) \to (0,0) \to \cdots \to (0,0), \\ &(i_0, b) \to \left(\left\lfloor\frac{i_0+rP}{b}\right\rfloor, \left\lfloor\frac{rQ+b}{b}\right\rfloor\right) = \left(\left\lfloor\frac{i_0+rP}{b}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b}\right\rfloor\right) \to \\ &\qquad \left(\left\lfloor\frac{i_0+rP}{b^2}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b^2}\right\rfloor\right) \to \cdots \to \left(\left\lfloor\frac{i_0+rP}{b^L}\right\rfloor, \left\lfloor\frac{rQ+b-1}{b^L}\right\rfloor\right) = (0,0), \\ &(i_0, b) \to (0,1) \to (0,0) \to \cdots \to (0,0),\end{aligned}$$

where L is chosen in a way that $b^{L+1} > \max\{P+1, Q+1\}$. Here we have used the facts that (since by assumption rQ/b is not an integer)

$$\left\lfloor\frac{rQ+b-1}{b}\right\rfloor = \left\lfloor\frac{rQ+b}{b}\right\rfloor$$

and that for all non-negative integers a

$$\left\lfloor\frac{\lfloor a/b\rfloor}{b}\right\rfloor = \left\lfloor\frac{a}{b^2}\right\rfloor.$$

The weights of these paths are given by

$$v_1(r) := g(i_0 + rP \bmod b)\overline{g(rQ-1 \bmod b)}\,\mathrm{e}(-rt)A, \quad w_1 := g(i_0)\overline{g(b-1)}$$

and by

$$v_2(r) := g(i_0 + rP \bmod b)\overline{g(rQ \bmod b)}\,\mathrm{e}(-rt)A, \quad w_2 := g(i_0)\overline{g(1)}$$

where

$$A = \prod_{j=1}^{L+1} g\left(\lfloor (i_0+rP)/b^j\rfloor\right)\overline{g\left(\lfloor (rQ+b-1)/b^j\rfloor\right)}.$$

Thus we have

$$\left|b_{(i_0,b-1),(0,0)}(t)\right| \le b^{-L-1}\left(B_{L+1}((i_0,b-1),(0,0)) - 2 + |v_1(r)+w_1|\right)$$

and

$$\left|b_{(i_0,1),(0,0)}(t)\right| \le b^{-L-1}\left(B_{L+1}((i_0,b),(0,0)) - 2 + |v_2(r)+w_2|\right)$$

Thus, in order to prove Lemma 3 we just have to check that there exist $1 \le r \le b-1$ with

$$\min\{|v_1(r)+w_1|, |v_2(r)+w_2|\} < 2. \tag{15}$$

Suppose that the converse statement holds, that is, for all $1 \le r \le b-1$ we have

$$|v_1(r)+w_1| = |v_2(r)+w_2| = 2.$$

Then we would have (for all $1 \le r \le b-1$)

$$v_1(r)/w_1 = v_2(r)/w_2$$

or equivalently

$$g(rQ \bmod b) = g(rQ-1 \bmod b)g(1)\overline{g(b-1)}.$$

Since $rQ \bmod b$, $1 \le r \le b-1$, runs precisely through the residue classes $1 \le \ell \le b-1$ this implies

$$g(\ell) = g(\ell-1)g(1)\overline{g(b-1)}.$$

By setting $\ell = 1$ it follows that $g(b-1) = 1$ (since $g(0) = 1$) and consequently we have

$$g(\ell) = g(1)^\ell, \qquad 1 \le \ell \le b-1.$$

Recall that $g(b-1) = 1$. Hence, we have $g(\ell) = \mathrm{e}(\ell j_0/(b-1))$ for some j_0 and consequently $g(n)$ is periodic. This is of course a contradiction and so (15) (and consequently Lemma 3) follows. □

This finally implies the main result of this section.

Proposition 2. *There exist constants $C > 0$ and $\eta > 0$ such that for all $\lambda \ge 0$*

$$\sup_{t\in\mathbb{R}} |F_\lambda(t)| \le C\, e^{-\eta\lambda}.$$

Proof. It follows from Lemma 3 that

$$\|\mathbf{A}(t)\cdot\mathbf{A}(bt)\cdots\mathbf{A}(b^\lambda t)\| \le (1-\delta)^{\lfloor \lambda/(L+1)\rfloor}$$

This implies that

$$|F_\lambda(t)| \le \|\mathbf{F}_\lambda(t)\| \le (1-\delta)^{\lfloor \lambda/(L+1)\rfloor}\|\mathbf{F}_0(t)\| \le C\, e^{-\eta\lambda}$$

holds uniformly for all $t \in \mathbb{R}$. □

3. Proof of Theorem 2

The essential step in the proof of Theorem 2 is the application of a theorem by the second and third authors [27, Theorem 1].

Assume that g is a non-periodic strongly b-multiplicative function of modulus 1. By our Proposition 2, the function f defined by $f(n) = g(p^2n)\overline{g(q^2n)}$ belongs to the set $\mathcal{F}_{\gamma,c}$ defined in [27, Definition 4], where $c > 0$ is arbitrary and $\gamma(\lambda)$ is maximal such that $Cb^{-\eta\lambda} \le b^{-\gamma(\lambda)}$ for all $\lambda \ge 0$. (here C and η are as in Proposition 2). Clearly, $\gamma(\lambda) \gg \eta\lambda$.

In order to apply Theorem 1 from [27], it is therefore sufficient to verify a *carry property* [27, Definition 3] for the function f. For this, we define, for any function $h : \mathbb{N} \to \mathbb{C}$ and $\lambda \ge 0$, the *truncation* h_λ as the b^λ-periodic continuation of $h \mid [0, b^\lambda)$. This function only takes into account the digits with indices below λ.

Lemma 4. *Assume that g is a non-periodic strongly b-multiplicative function of modulus 1. Define $f(n) = g(p^2n)\overline{g(q^2n)}$. There exists $C > 0$ such that for all nonnegative integers λ, κ, ρ satisfying $\rho < \lambda$, the number of integers $0 \le \ell < b^\lambda$ such that*

$$f(\ell b^\kappa + k_1 + k_2)\overline{f(\ell b^\kappa + k_1)} \ne f_{\kappa+\rho}(\ell b^\kappa + k_1 + k_2)\overline{f_{\kappa+\rho}(\ell b^\kappa + k_1)} \tag{16}$$

for some $(k_1, k_2) \in \{0, \dots, b^\kappa - 1\}^2$ is bounded by $Cb^{\lambda-\rho}$.

Proof. Separating the factors corresponding to p and q, it is sufficient to verify this property for the function $f(n) = g(an)$, where $a \ge 0$. We need to investigate the carry propagation occurring in the addition $s_1 + s_2$, where $s_1 = a\ell b^\kappa + ak_1$ and $s_2 = ak_2$. If $s_1 \in [0, b^{\kappa+\rho} - ab^\kappa) + b^{\kappa+\rho}\mathbb{N}$,the addition of s_2 does not change the base-b digits of s_1 above $\kappa + \rho$; it is therefore sufficient to demand that $a\ell \in [0, b^\rho - 2a) + b^\rho\mathbb{N}$ in order to obtain equality in (16) for all k_1, k_2. For ρ large enough, this condition is violated for $\mathcal{O}(b^\lambda a/b^\rho)$ many $\ell < b^\lambda$, which implies the statement. □

Applying Mauduit and Rivat's Theorem 1 [27], we obtain

$$\sum_{0 \le n < N} g(p^2n^2)\overline{g(q^2n^2)} = o(N)$$

for all strongly b-multiplicative functions g of modulus 1 and coprime p and q that are also coprime to b. As a final step, we apply the Daboussi–Kátai criterion [4, 22]

Lemma 5 (Daboussi–Kátai/Bourgain–Sarnak–Ziegler). *Let $f : \mathbb{N} \to \mathbb{C}$ be bounded and such that*

$$\sum_{n \le x} f(pn)\overline{f(qn)} = o(x) \tag{17}$$

for all distinct primes p and q. Then

$$\sum_{n \le x} \mu(n) f(n) = o(x).$$

In fact it is sufficient to restrict the condition (17) to large enough primes p and q — Bourgain–Sarnak–Ziegler [4, page 80] note that their proof only involves primes larger than an arbitrary bound. Applying this lemma to $f(n) = g(n^2)$, we obtain

$$\sum_{0 \le n < N} \mu(n) g(n^2) = o(N)$$

and therefore our Theorem 2.

Acknowledgements. The authors wish to thank Mariusz Lemańczyk for pointing out valuable references to the literature concerning Möbius orthogonality of positive entropy dynamical systems, and helping us place our result into the context of related research works. Moreover, we thank Clemens Müllner for several fruitful discussions.

References

[1] J.-P. Allouche and J. Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations.
[2] J. Bourgain, *Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat*, J. Anal. Math., 119 (2013), pp. 147–163.
[3] ——, *On the correlation of the Moebius function with rank-one systems*, J. Anal. Math., 120 (2013), pp. 105–130.
[4] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Moebius from horocycle flows*, in From Fourier analysis and number theory to Radon transforms and geometry, vol. 28 of Dev. Math., Springer, New York, 2013, pp. 67–83.
[5] H. Davenport, *On some infinite series involving arithmetical functions (II)*, The Quarterly Journal of Mathematics, os-8 (1937), pp. 313–320.
[6] T. Downarowicz and S. a. Kasjan, *Odometers and Toeplitz systems revisited in the context of Sarnak's conjecture*, Studia Math., 229 (2015), pp. 45–72.
[7] T. Downarowicz and J. Serafin, *Almost full entropy subshifts uncorrelated to the Möbius function*, Int. Math. Res. Not. IMRN, (2019), pp. 3459–3472.
[8] ——, *A strictly ergodic, positive entropy subshift uniformly uncorrelated to the Möbius function*, Studia Math., 251 (2020), pp. 195–206.

[9] M. Drmota, C. Mauduit, and J. Rivat, *Normality along squares*, J. Eur. Math. Soc. (JEMS), 21 (2019), pp. 507–548.

[10] A. Dymek, S. a. Kasjan, J. Kuł aga Przymus, and M. Lemańczyk, *$\mathcal{B}$-free sets and dynamics*, Trans. Amer. Math. Soc., 370 (2018), pp. 5425–5489.

[11] T. Eisner, *A polynomial version of Sarnak's conjecture*, C. R. Math. Acad. Sci. Paris, 353 (2015), pp. 569–572.

[12] E. H. El Abdalaoui, J. Kuł aga Przymus, M. Lemańczyk, and T. de la Rue, *The Chowla and the Sarnak conjectures from ergodic theory point of view*, Discrete Contin. Dyn. Syst., 37 (2017), pp. 2899–2944.

[13] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, *A dynamical point of view on the set of $\mathcal{B}$-free integers*, Int. Math. Res. Not. IMRN, (2015), pp. 7258–7286.

[14] E. H. el Abdalaoui, M. Lemańczyk, and T. de la Rue, *Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals*, Int. Math. Res. Not. IMRN, (2017), pp. 4350–4368.

[15] ——, *Erratum to "Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals" [MR3674173]*, Int. Math. Res. Not. IMRN, (2017), p. 4493.

[16] S. Ferenczi and C. Mauduit, *On Sarnak's conjecture and Veech's question for interval exchanges*, 2018.

[17] B. Green, *On (not) computing the Möbius function using bounded depth circuits*, Combin. Probab. Comput., 21 (2012), pp. 942–951.

[18] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2), 175 (2012), pp. 541–566.

[19] E. Houcein El Abdalaoui, S. Kasjan, and M. Lemańczyk, *0-1 sequences of the Thue-Morse type and Sarnak's conjecture*, Proc. Amer. Math. Soc., 144 (2016), pp. 161–176.

[20] E. Houcein El Abdalaoui, M. Lemańczyk, and T. de la Rue, *On spectral disjointness of powers for rank-one transformations and Möbius orthogonality*, J. Funct. Anal., 266 (2014), pp. 284–317.

[21] D. Karagulyan, *On Möbius orthogonality for interval maps of zero entropy and orientation-preserving circle homeomorphisms*, Ark. Mat., 53 (2015), pp. 317–327.

[22] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar., 47 (1986), pp. 223–225.

[23] J. Kułaga-Przymus and M. Lemańczyk, *The Möbius function and continuous extensions of rotations*, Monatsh. Math., 178 (2015), pp. 553–582.

[24] J. Liu and P. Sarnak, *The Möbius function and distal flows*, Duke Math. J., 164 (2015), pp. 1353–1399.

[25] C. Mauduit and J. Rivat, *La somme des chiffres des carrés*, Acta Math., 203 (2009), pp. 107–148.

[26] C. Mauduit and J. Rivat, *Sur un problème de Gelfond : la somme des chiffres des nombres premiers*, Ann. of Math. (2), 171 (2010), pp. 1591–1646.

[27] C. Mauduit and J. Rivat, *Rudin-Shapiro sequences along squares*, Trans. Amer. Math. Soc., 370 (2018), pp. 7899–7921.

[28] Y. Moshe, *On the subword complexity of Thue-Morse polynomial extractions*, Theoret. Comput. Sci., 389 (2007), pp. 318–329.

[29] C. Müllner, *Automatic sequences fulfill the Sarnak conjecture*, Duke Math. J., 166 (2017), pp. 3219–3290.

[30] ——, *The Rudin-Shapiro sequence and similar sequences are normal along squares*, Canad. J. Math., 70 (2018), pp. 1096–1129.

[31] R. Peckner, *Uniqueness of the measure of maximal entropy for the squarefree flow*, Israel J. Math., 210 (2015), pp. 335–357.

[32] ——, *Möbius disjointness for homogeneous dynamics*, Duke Math. J., 167 (2018), pp. 2745–2792.

[33] P. Sarnak, *Three lectures on the Mobius function randomness and dynamics.* Available from https://www.math.ias.edu/files/wam/2011/PSMobius.pdf.

[34] P. Sarnak, *Mobius randomness and dynamics*, Not. S. Afr. Math. Soc., 43 (2012), pp. 89–97.

[35] P. Sarnak and A. Ubis, *The horocycle flow at prime times*, J. Math. Pures Appl. (9), 103 (2015), pp. 575–618.

[36] W. A. Veech, *Möbius orthogonality for generalized Morse-Kakutani flows*, American Journal of Mathematics, (2016). (to appear).

Email address: `michael.drmota@tuwien.ac.at`

Institut für Diskrete Mathematik und Geometrie TU Wien, Wiedner Hauptstr. 8–10, 1040 Wien, Austria

Email address: `joel.rivat@univ-amu.fr`

Université d'Aix-Marseille, Institut de Mathématiques de Marseille, CNRS UMR 7373, 163, avenue de Luminy, Case 907, 13288 MARSEILLE Cedex 9, France

Email address: lukas.spiegelhofer@tuwien.ac.at

Institut für Diskrete Mathematik und Geometrie TU Wien, Wiedner Hauptstr. 8–10, 1040 Wien, Austria