Ergod. Th. & Dynam. Sys., (2022), **42**, 1446–1473
doi:10.1017/etds.2021.3

Prime number theorem for regular Toeplitz subshifts

KRZYSZTOF FRĄCZEK†, ADAM KANIGOWSKI‡ and
MARIUSZ LEMAŃCZYK§

§ *Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland*
(e-mail: fraczek@mat.umk.pl)
§ *Department of Mathematics, University of Maryland, 4176 Campus Drive, William E. Kirwan Hall, College Park, MD 20742-4015, USA*
(e-mail: akanigow@umd.edu)
§ *Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland*
(e-mail: mlem@mat.umk.pl)

(*Received* 27 *April* 2020 *and accepted in revised form* 6 *January* 2021)

Abstract. We prove that neither a prime nor an l-almost prime number theorem holds in the class of regular Toeplitz subshifts. But when a quantitative strengthening of the regularity with respect to the periodic structure involving Euler's totient function is assumed, then the two theorems hold.

Key words: almost prime numbers, polynomial ergodic theorems, prime number theorem, Toeplitz systems
2020 Mathematics Subject Classification: 37B10, 37A44 (Primary); 11N05, 11N13 (Secondary)

1. *Introduction*

Given a topological dynamical system (X, T), where T is a homeomorphism of a compact metric space X, one says that a prime number theorem (PNT) holds for (X, T) if the limit

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f(T^p x) \tag{1}$$

(p stands always for a prime number) exists for each $x \in X$, an arbitrary $f \in C(X)$ and $\pi(N)$ denotes the number of primes up to N. Then, via the Riesz theorem, for all

$f \in C(X)$, we have

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f(T^p x) = \int_X f\, d\nu_x \tag{2}$$

for a Borel probability measure ν_x on X, where ν_x depends only on $x \in X$.

Let us first consider the cyclic case $X = \mathbb{Z}/k\mathbb{Z}$ and $Tx = x + 1$. Fix $x \in X$ and notice that (1) indeed holds by the classical prime number theorem in arithmetic progressions, where ν_x is the uniform probability measure on the 'coset' $\{a < k : (a, k) = 1\} + x$. Hence, a PNT holds in cyclic (and therefore also in finite) systems.

Consider now the procyclic case, that is, assume we are given an odometer system (H, T) with

$$H = \operatorname{liminv}_{t\to\infty} \mathbb{Z}/n_t\mathbb{Z}, \quad Tx = x + (1, 1, \ldots)$$

(here $n_t|n_{t+1}$ for $t \geqslant 0$). In this case, a PNT still holds. Indeed, the space H has a sequence of natural partitions $D^t = (D^t_0, \ldots, D^t_{n_t-1})$, $t \geqslant 0$, consisting of clopen sets and such that $TD^t_i = D^t_{i+1 \bmod n_t}$. It follows that the sets D^t_i, $i \leqslant n_t - 1$, have the same diameter which goes to 0 as $t \to \infty$. Moreover, it is not hard to see that each character of the group H is constant on the levels of the towers D^t for t sufficiently large. Hence, each $f \in C(H)$ can be approximated *uniformly* by functions which are constant on the levels of the towers D^t and a PNT holds because it does in the finite case.

Our main results concern prime number theorems for extensions of odometers. Recall that odometers are zero-entropy topological systems which are minimal (all T-orbits are dense) and uniquely ergodic (there is only one T-invariant measure: Haar measure in this case). Before we describe our results, let us discuss a PNT in the class of uniquely ergodic systems. First, recall that for all such systems (1) holds almost everywhere with respect to the unique invariant measure [3, 23]. On the other hand, one can easily construct a counterexample to the validity of (1) for *all* $x \in X$. Indeed, denote by $\mathbb{P}$ the set of prime numbers and consider any subset $P \subset \mathbb{P}$ with no density in $\mathbb{P}$, the left shift S on $\{0, 1\}^{\mathbb{Z}}$ and the subshift $(X_{\mathbf{1}_{P\cup(-P)}}, S)$ obtained by the orbit closure of the characteristic function $\mathbf{1}_{P\cup(-P)}$. It has a unique invariant measure of zero entropy (which is the Dirac measure at the fixed point $\ldots 0.00 \ldots$) and a PNT fails in it (see, for example, [7] for details). Now, this particular uniquely ergodic model of the one-point system implies paradoxically that each ergodic dynamical system has a uniquely ergodic model (X, T) in which a PNT does not hold. (Recall that the Jewett–Kreiger theorem says the following. Suppose (Z, κ, R) is an ergodic measure-theoretic dynamical system. Then there exists a uniquely ergodic (even strictly ergodic, that is, additionally minimal) topological system (Y, S) with the unique invariant measure ν such that (Z, κ, R) and (Y, ν, S) are *measure-theoretically* isomorphic.) To see this, take any uniquely ergodic model (Y, ν, R) of the given measure-theoretic dynamical system. Since the one-point system is (Furstenberg) disjoint with any other system, the product system $(X_{\mathbf{1}_{P\cup(-P)}} \times Y, S \times R)$ is still uniquely ergodic, with the unique invariant measure $\delta_{\ldots 0.00\ldots} \otimes \nu$. It is not hard to see that the product system is still measure-theoretically isomorphic to the original system. Since the new system has $(X_{\mathbf{1}_{P\cup(-P)}}, S)$ as its *topological* factor, a PNT does not hold in $(X_{\mathbf{1}_{P\cup(-P)}} \times Y, S \times R)$. (To illustrate this, consider an irrational rotation R_α on $\mathbb{T}$ for which a PNT holds because of Vinogradov's theorem (prime 'orbits' are equidistributed). However, our observation shows that there is a uniquely ergodic model of R_α in which the eigenfunctions are still continuous but a PNT fails, that is, some of the prime 'orbits' are not equidistributed.) Hence, if we think about a necessary condition for a PNT to hold, it looks reasonable to add the minimality assumption to avoid a problem of 'exotic' orbits on which a PNT does not hold (we also recall that a uniquely ergodic system has a unique subsystem which is strictly ergodic). However, in this class one can still produce counterexamples to a PNT; see [18] for the first symbolic counterexamples (although their entropy is not determined in [18]), or [12] for non-symbolic counterexamples. On the other hand, we have quite a few classes in which a PNT holds, including systems of algebraic origin [10, 22], symbolic systems [4, 9, 15, 16] or recently [12] in the category of smooth systems, where a PNT has been proved in the class of analytic Anzai skew products. Finding a sufficient *dynamical* condition for a PNT to hold, postulated a few years ago by Sarnak [20] seems to be an important and difficult task in dynamics; however, we rather expect the following result.

Working Conjecture. Each ergodic and aperiodic (the set of periodic points has measure zero) measure-theoretic dynamical system has a strictly ergodic model in which a PNT fails.

If true, this makes Sarnak's postulate even harder to realize. The present paper should be viewed as introductory steps in trying to understand the conjecture.

A PNT can be reformulated as the existence of a limit of $(1/N)\sum_{n<N} f(T^n x)\Lambda(n)$, where Λ stands for the von Mangoldt function: $\Lambda(p^\ell) = \log p$ for $\ell \geqslant 1$ and 0 otherwise. Proving dynamical prime number theorems for zero-entropy systems is closely related to Sarnak's Möbius disjointness conjecture [19]:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(T^n x)\boldsymbol{\mu}(n) = 0 \tag{3}$$

for each $x \in X$, $f \in C(X)$ in each *zero-entropy* dynamical system (X, T) ($\boldsymbol{\mu}$ stands for the Möbius function: $\boldsymbol{\mu}(1) = 1$, $\boldsymbol{\mu}(p_1 \cdot \ldots \cdot p_k) = (-1)^k$ for different primes $p_1, \ldots, p_k$, and $\boldsymbol{\mu}(n) = 0$ for the remaining $n \in \mathbb{N}$). Here, the class of systems for which we expect the positive answer is precisely defined. In fact, in quite a few cases (see [4, 5, 8–10, 15, 16]) one can observe the following principle: once we can prove Sarnak's conjecture for (X, T) with a 'sufficient' speed of convergence to zero in (3) then a PNT holds in (X, T).

With all the above in mind, we return to extensions of odometers that we intend to study. We stay in the *zero-entropy* category of systems and we assume *minimality*. Further, we assume that the systems are *almost one-to-one extensions* of odometers. (If (H, T) is a factor of (X, S) via $\pi : X \to H$, then (X, S) is called an *almost one-to-one extension* of (H, T) if there is a point $h \in H$ such that $|\pi^{-1}(h)| = 1$; in fact, in this case the set of points with singleton fibers is G_δ and dense.) We also assume that our systems are *symbolic*. (We recall that each zero-entropy system has an extension which is symbolic [2], and clearly if a PNT holds for a system, it does for a factor.) All these natural assumptions determine, however, a very precise class of topological systems, namely Toeplitz subshifts (X_x, S), where x is a Toeplitz sequence over a finite alphabet $\mathcal{A}$; see Downarowicz's survey [6, §7]. That is, $x \in \mathcal{A}^{\mathbb{Z}}$ has the property that for every $a \in \mathbb{Z}$ there is $\ell \in \mathbb{N}$ such that $x(a) = x(a + k\ell)$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

for each $k \in \mathbb{Z}$, and X_x is the set of all $y \in \mathcal{A}^{\mathbb{Z}}$ with the property that all subblocks of y also appear in x. One shows then that there is a sequence $n_t | n_{t+1}$ such that if $\mathrm{Per}_{n_t}(x) := \{a \in \mathbb{Z} : x(a) = x(a + kn_t) \text{ for each } k \in \mathbb{Z}\}$ then

$$\bigcup_{t \geqslant 0} \mathrm{Per}_{n_t}(x) = \mathbb{Z}. \tag{4}$$

Moreover, there is a natural continuous factor map $\pi : X_x \to H$, where H stands for the odometer determined by (n_t). In fact, we will restrict our attention to so-called *regular* Toeplitz subshifts, whose formal definition is that the density of $\bigcup_{t=0}^{M} \mathrm{Per}_{n_t}(x)$ goes to 1. Regular Toeplitz subshifts are zero-entropy strictly ergodic systems, and measure-theoretically isomorphic to the rotations given by their maximal equicontinuous factors. Although in [**6**] there are four other equivalent conditions for regularity (see [**6**, Theorem 13.1]), we will choose a different path. Since $\pi : X_x \to H$ is a continuous and equivariant surjection,

$$E^t := \pi^{-1}(D^t) = (E_0^t, \ldots, E_{n_t - 1}^t) \quad \text{with } E_j^t = \pi^{-1}(D_j^t)$$

is an S-tower of height n_t whose levels are closed (hence clopen). By the minimality of (X_x, S) there is a unique tower with clopen levels and of fixed height. Let us consider a metric on $\mathcal{A}^{\mathbb{Z}}$ inducing the product topology given by

$$d(x, y) = 2^{-\inf\{|n| : x(n) \neq y(n)\}}.$$

The diameters of the levels of towers E^t *do not* converge to zero, unless x is periodic. Moreover, the diameters of different levels are in general different as the shift S is not an isometry. Let us consider the diameter of the tower E^t given by

$$\delta(E^t) := \sum_{0 \leqslant j < n_t} \mathrm{diam}(E_j^t).$$

It is not hard to see (see Appendix A) that the regularity of a Toeplitz sequence is equivalent to

$$\lim_{t \to \infty} \frac{\delta(E^t)}{n_t} = 0. \tag{5}$$

It is also not hard to see that this property does not depend on the choice of (n_t) satisfying (4). We recall that the Möbius disjointness of subshifts given by regular Toeplitz sequences has been proved in [**1**]. Here are two first results of the paper proved in §2 and §4, respectively.

THEOREM A. *A PNT does not hold in the class of minimal almost one-to-one symbolic extensions of odometers satisfying* (5). *That is, a PNT need not hold in a strictly ergodic subshift determined by a* regular *Toeplitz sequence.*

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

THEOREM B. *A PNT holds in the class of minimal almost one-to-one symbolic extensions of odometers in which* (5) *holds with a speed*

$$\lim_{t\to\infty} \frac{\delta(E^t)}{\varphi(n_t)} = 0, \tag{6}$$

where φ *denotes the Euler totient function.*

As for all Toeplitz dynamical systems constructed in the proof of Theorem A, we have

$$0 < \liminf_{t\to\infty} \frac{\delta(E^t)}{\varphi(n_t)} \leqslant \limsup_{t\to\infty} \frac{\delta(E^t)}{\varphi(n_t)} < +\infty,$$

which shows that the condition (6) in Theorem B is optimal to have a PNT. The systems in Theorem B are strictly ergodic and since they all have non-trivial cyclic factors, the measures ν_y, $y \in X_x$, in (2) are never S-invariant. (To be compared with the case of Sturmian systems (see Theorem B.1) in which ν_y, $y \in X_{\alpha,\beta}$, are equal to the unique S-invariant measure.)

We then turn our attention to an l-almost prime number theorem (P_lNT) which is much less explored than the PNT case and which, for the first time in dynamics, is studied in [**13**] (for some smooth Anzai skew products). Recall that for any $l \geqslant 1$ a natural number is called an l-almost prime if it is a product of l primes. We denote the set of l-almost prime numbers by $\mathbb{P}_l$. By $\mathbb{P}_l^N$ we denote the set of l-almost prime numbers less than or equal to N and we let $\pi_l(N)$ stand for the cardinality of $\mathbb{P}_l^N$. A classical result of Landau asserts that

$$\lim_{N\to\infty} \frac{\pi_l(N)}{(N/\log N)((\log\log N)^{l-1}/(l-1)!)} = 1; \tag{7}$$

see [**14**, §56].

Analogously to the PNT, we say that a topological dynamical system (X, T) satisfies a P_lNT if the limit

$$\lim_{N\to\infty} \frac{1}{\pi_l(N)} \sum_{n\in\mathbb{P}_l^N} f(T^n x)$$

exists for each $x \in X$ and each $f \in C(X)$.

In §3 and §5 we provide sketches of proofs of the exact analogues of Theorems A and B for a P_lNT for regular Toeplitz subshifts.

In §6.1 we prove a new polynomial ergodic theorem:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} f(S^{P(n)}x) \text{ exists}$$

for monic polynomials P with positive integer coefficients for all symbolic minimal almost one-to-one extensions of odometers with a modified condition (6). In §6.2 we provide a regular Toeplitz subshift which does not satisfy the polynomial ergodic theorem for squares but satisfies a PNT. We refer again to [**18**] for the first examples of strictly ergodic systems (of low complexity), where the Birkhoff ergodic averages along squares do not converge.

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

While Theorem A confirms our working conjecture for a subclass of odometers, we have been unable to confirm it for the whole class of odometers. Confirming our working conjecture for the class of automorphisms with discrete spectrum seems to be the first step toward a possible general statement. In Appendix B we provide a simple argument showing that a PNT holds for all symbolic models of irrational rotations given by Sturmian sequences. The Sturmian systems are strictly ergodic and are almost one-to-one extensions of irrational rotations.

2. *Regular Toeplitz subshifts which do not satisfy a PNT (proof of Theorem A)*

For all $K, n \in \mathbb{N}$ and $a \in \mathbb{Z}$, let

$$\pi(K; n, a) = \{1 \leqslant p \leqslant K : p \in \mathbb{P},\ p = a \bmod n\}.$$

THEOREM 2.1. (PNT in arithmetic progressions; see [21]) *For any natural n and any integer a with $(a, n) = 1$ we have*

$$\lim_{K\to\infty} \frac{\pi(K; n, a)}{\pi(K)} = \frac{1}{\varphi(n)}.$$

We construct a Toeplitz sequence $x \in \{0, 1\}^{\mathbb{Z}}$ with the period structure (n_t):

$$n_{t+1} = k_{t+1} n_t, \quad (k_{t+1}, n_t) = 1, \tag{8}$$

for each $t \geqslant 1$. We will show that for this x,

$$\lim_{t\to\infty} \frac{1}{\pi(n_t)} \sum_{p<n_t} F(S^p x) \text{ does not exist,}$$

where $F(y) = (-1)^{y(0)}$. At stage t, x is approximated by the infinite concatenation of $x_t[0, n_t - 1] \in \{0, 1, ?\}^{n_t}$ (i.e., we see a periodic sequence of 0, 1, ? with period n_t). Successive '?' will be filled in the next steps of construction of x. We require that

$$\frac{\varphi(n_t)}{n_t} \leqslant \frac{1}{2^t}, \tag{9}$$

$$\{0 \leqslant i < n_t : x_t(i) = ?\} \subset \{0 \leqslant j < n_t : (j, n_t) = 1\}, \tag{10}$$

$$\#\{0 \leqslant i < n_t : x_t(i) = ?\} \geqslant \bigg(1 - \sum_{l=1}^{t} \frac{1}{100^l}\bigg)\varphi(n_t), \tag{11}$$

$$\#\{p < n_t : x_t(p) = ?\} \geqslant \tfrac{1}{2}\pi(n_t). \tag{12}$$

We choose k_{t+1} satisfying (8), and

$$\frac{\varphi(k_{t+1})}{k_{t+1}} \leqslant \frac{1}{2}, \tag{13}$$

$$\varphi(k_{t+1}) \geqslant 100^{t+1}, \tag{14}$$

$$8 \log n_{t+1} \leqslant \pi(n_{t+1}), \quad 8\pi(n_t) \leqslant \pi(n_{t+1}) \tag{15}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

and *for each* $0 < a < n_t$, $(a, n_t) = 1$, we have

$$\#(\{a + jn_t : j = 0, \ldots, k_{t+1}\} \cap \mathbb{P}) = \pi(n_{t+1}; n_t, a) \leqslant 2\frac{\pi(n_{t+1})}{\varphi(n_t)}. \tag{16}$$

To see (13) note that if p_i stands for the ith prime then $\sum_{j\geqslant i} 1/p_j = +\infty$, whence, remembering that $\varphi(p_i p_{i+1} \ldots p_{i+s}) = p_i p_{i+1} \ldots p_{i+s}\prod_{j=0}^{s}(1 - 1/p_{i+j})$, we have $\prod_{j=0}^{s}(1 - 1/p_{i+j}) \to 0$, and therefore $\prod_{j=0}^{s}(1 - 1/p_{i+j}) < 1/2$ for s large enough. The latter we obtain from Theorem 2.1 (remembering that n_t is fixed, so the number of a is known, and we can obtain the accuracy as good as we want by taking k_{t+1} sufficiently large).

We need two simple observations. Firstly,

$$\{0 \leqslant k < n_{t+1} : (k, n_{t+1}) = 1\} \subset \bigcup_{\substack{0\leqslant a<n_t\\(a,n_t)=1}} \{a + jn_t : j = 0, \ldots, k_{t+1} - 1\}. \tag{17}$$

Then we have the following.

LEMMA 2.2. *For every* $0 \leqslant a < n_t$ *with* $(a, n_t) = 1$*, we have*

$$\#\{0 \leqslant j < k_{t+1} : (a + jn_t, n_{t+1}) = 1\} = \varphi(k_{t+1}).$$

Proof. First note that $(a + jn_t, n_{t+1}) = 1$ if and only if $(a + jn_t, k_{t+1}) = 1$. Indeed, assume that $(a + jn_t, k_{t+1}) = 1$. If for some prime p we have $p|(a + jn_t)$ and $p|n_{t+1} = n_t k_{t+1}$, then $p|k_{t+1}$. Otherwise, we have $p|n_t$, so $p|a$. As $(a, n_t) = 1$, this gives a contradiction. Thus $(a + jn_t, k_{t+1}) = 1$ implies $(a + jn_t, n_{t+1}) = 1$. The opposite implication is obvious. Thus

$$\{0 \leqslant j < k_{t+1} : (a + jn_t, n_{t+1}) = 1\} = \{0 \leqslant j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\}.$$

Let us consider the affine map

$$\mathbb{Z}/k_{t+1}\mathbb{Z} \ni j \overset{A}{\mapsto} a + jn_t \in \mathbb{Z}/k_{t+1}\mathbb{Z}.$$

If $J := \{0 \leqslant \ell < k_{t+1} : (\ell, k_{t+1}) = 1\}$ then

$$\{0 \leqslant j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\} = A^{-1}(J).$$

Since $(n_t, k_{t+1}) = 1$, the map A is a bijection. It follows that

$$\begin{aligned}\#\{0 \leqslant j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\} &= \#\{0 \leqslant \ell < k_{t+1} : (\ell, k_{t+1}) = 1\}\\ &= \varphi(k_{t+1}),\end{aligned}$$

which completes the proof. □

We now need to describe which '?' we fill in $x_{t+1}[0, n_{t+1} - 1]$ and how. This block is divided into k_{t+1} subblocks

$$\underbrace{x_t[0, n_t - 1]x_t[0, n_t - 1] \ldots x_t[0, n_t - 1]}_{k_{t+1}}.$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

We fill in *all* '?' in the first block $x_t[0, n_t - 1]$ in such a way as to 'destroy' a PNT for the time n_t, namely

$$\frac{1}{\pi(n_t)} \sum_{p<n_t} F(S^p x) = \frac{1}{\pi(n_t)} \sum_{\substack{p<n_t \\ p|n_t}} (-1)^{x(p)} + \frac{1}{\pi(n_t)} \Bigg(\sum_{\substack{p<n_t \\ (p,n_t)=1 \\ x_t(p)=0}} 1 - \sum_{\substack{p<n_t \\ (p,n_t)=1 \\ x_t(p)=1}} 1 + \sum_{\substack{p<n_t \\ (p,n_t)=1 \\ x_t(p)=?}} (-1)^{x(p)} \Bigg).$$

As the number of the primes dividing n_t is bounded by $\log n_t$, it is negligible compared to $\pi(n_t) = n_t/\log n_t$. It follows that

$$\Bigg| \frac{1}{\pi(n_t)} \sum_{\substack{p<n_t \\ p|n_t}} (-1)^{x(p)} \Bigg| \leqslant \frac{\log n_t}{\pi(n_t)} = o(1),$$

so the first summand does not affect the asymptotic of the averages in (1). Since the number of p in the last summand is at least $\frac{1}{2}\pi(n_t)$ in view of (12), we can fill in x_{t+1} at places $\{p < n_t : (p, n_t) = 1,\ x_t(p) =?\}$ to obtain a sum completely different than the known number which we had from stage $t - 1$.

We fill in (in an arbitrary way) all remaining places between 0 and $n_t - 1$ and all places $a + jn_t$ for $0 \leqslant j < k_{t+1}$ such that this number is not coprime with n_{t+1}, so that (10) will be satisfied at stage $t + 1$. We must remember that for certain $0 < a < n_t$ coprime to n_t, $x_t(a)$ was already defined at previous stages, so along the corresponding arithmetic progressions $a + jn_t$, $0 \leqslant j < k_{t+1}$, these places are also filled in previously. On the other hand, if $x_{t+1}(a + jn_t) \neq ?$ (i.e., $x_{t+1}(a + jn_t) = 0$ or $x_{t+1}(a + jn_t) = 1$) and $(a + jn_t, n_{t+1}) = 1$ for some $0 < j < k_{t+1}$ then $x_t(a) \neq ?$. In view of (17), this gives

$$\begin{aligned} &\{0 \leqslant i < n_{t+1} : (i, n_{t+1}) = 1, x_{t+1}(i) \neq ?\} \\ &\quad \subset \{0 < a < n_t : (a, n_t) = 1, x_{t+1}(a) \neq ?\} \\ &\quad\quad \cup \bigcup_{\substack{0\leqslant a<n_t \\ (a,n_t)=1 \\ x_t(a)\neq ?}} \{a + jn_t : 0 < j < k_{t+1},\ (a + jn_t, n_{t+1}) = 1\}. \end{aligned}$$

By (10), Lemma 2.2, (11) and (14), it follows that

$$\begin{aligned} &\#\{0 \leqslant i < n_{t+1} : (i, n_{t+1}) = 1, x_{t+1}(i) \neq ?\} \\ &\quad \leqslant \varphi(n_t) + \#\{0 \leqslant a < n_t : (a, n_t) = 1, x_t(a) \neq ?\}\varphi(k_{t+1}) \\ &\quad \leqslant \varphi(n_t) + \Bigg(\sum_{k=1}^{t} \frac{1}{100^k}\Bigg)\varphi(n_t)\varphi(k_{t+1}) = \Bigg(\frac{1}{\varphi(k_{t+1})} + \sum_{k=1}^{t} \frac{1}{100^k}\Bigg)\varphi(n_{t+1}) \\ &\quad \leqslant \sum_{k=1}^{t+1} \frac{1}{100^k}\varphi(n_{t+1}) \leqslant \frac{1}{99}\varphi(n_{t+1}). \end{aligned}$$

In particular, at stage $t + 1$, (11) is also satisfied.

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

Similar arguments show that

$$\{p < n_{t+1} : x_{t+1}(p) \neq ?\} \subset \{p < n_{t+1} : p | n_{t+1}\} \cup \{p < n_t : x_{t+1}(p) \neq ?\} \cup \bigcup_{\substack{0 \leqslant a < n_t \\ (a,n_t)=1 \\ x_t(a) \neq ?}} \{a + jn_t : 0 < j < k_{t+1},\ a + jn_t \in \mathbb{P}\}.$$

In view of (16), (11) and (15), it follows that

$$\begin{aligned}&\#\{p < n_{t+1} : x_{t+1}(p) \neq ?\}\\ &\quad \leqslant \log n_{t+1} + \pi(n_t) + 2\#\{0 \leqslant a < n_t : (a, n_t) = 1, x_t(a) \neq ?\}\frac{\pi(n_{t+1})}{\varphi(n_t)}\\ &\quad \leqslant \log n_{t+1} + \pi(n_t) + \frac{2}{99}\varphi(n_t)\frac{\pi(n_{t+1})}{\varphi(n_t)} \leqslant \frac{1}{2}\pi(n_{t+1}).\end{aligned}$$

Therefore, at stage $t + 1$, (12) is also satisfied.

Finally, note that

$$\frac{\varphi(n_{t+1})}{n_{t+1}} = \frac{\varphi(n_t)}{n_t}\frac{\varphi(k_{t+1})}{k_{t+1}} \leqslant \frac{\varphi(n_t)}{n_t}\frac{1}{2},$$

so (9) holds and the resulting Toeplitz sequence is regular.

3. *Toeplitz subshifts for which a P_lNT does not hold*

We now intend to give an example of a (regular) Toeplitz sequence x such that a P_lNT does not hold for the corresponding subshift. In fact,

$$\lim_{t\to\infty} \frac{1}{\pi_l(n_t)} \sum_{p^{(l)} \in \mathbb{P}_l^{n_t}} F(S^{p^{(l)}} x) \text{ does not exist.}$$

For any natural m and $0 \leqslant a < m$, let

$$\pi_l(N; m, a) := \#(\mathbb{P}_l^N \cap (a + m\mathbb{Z})).$$

LEMMA 3.1. *If $(a, m) > 1$ then*

$$\pi_l(N; m, a) = o(\pi_l(N)). \tag{18}$$

If $(a, m) = 1$ then

$$\lim_{N\to\infty} \frac{\pi_l(N; m, a)}{\pi_l(N)} = \frac{1}{\varphi(m)}. \tag{19}$$

Proof. The proof is by induction on l. If $l = 1$ and $(a, m) > 1$ then $\pi_l(N; m, a) \leqslant 1$, so (18) holds. If $(a, m) = 1$ then (19) is given by Theorem 2.1.

Suppose that (18) and (19) are satisfied for all parameters less than some natural number $l \geqslant 2$. Assume that $(a, m) \in \mathbb{P}_j$ for some $j \geqslant 1$. If $j > l$ then $\pi_l(N; m, a) = 0$. If $(a, m) \in \mathbb{P}_l$ then $\pi_l(N; m, a) \leqslant 1$, so (18) holds. If $p^{(j)} := (a, m) \in \mathbb{P}_j$ for some $1 \leqslant j < l$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

then

$$\pi_l(N; m, a) \leqslant \pi_{l-j}([N/p^{(j)}]; m/p^{(j)}, a/p^{(j)}) = O(\pi_{l-j}(N))$$
$$= O\bigg(\frac{N(\log\log N)^{l-j-1}}{\log N}\bigg) = o\bigg(\frac{N(\log\log N)^{l-1}}{\log N}\bigg) = o(\pi_l(N)).$$

Now suppose that $(a, m) = 1$. Assume that $p_1 \leqslant p_2 \leqslant \ldots \leqslant p_l$ are prime numbers such that $p^{(l)} = p_1 \cdots p_l \leqslant N$, $p^{(l)} = a \bmod m$. Then $p_1 \leqslant \sqrt[l]{N}$. Since $(p_1, m) = 1$, there exists a unique $0 \leqslant a(p_1) < m$ such that $p_1 \cdot a(p_1) = a \bmod m$ and $(a(p_1), m) = 1$. Then

$$\pi_l(N; m, a) = \sum_{p_1 \leqslant \sqrt[l]{N}} \pi_{l-1}([N/p_1]; m, a(p_1)).$$

As $p_1 \leqslant \sqrt[l]{N}$ implies $N/p_1 \geqslant N^{1-1/l}$, by assumption, for every $\varepsilon > 0$ there exists N_ε such that for all $N \geqslant N_\varepsilon$ and $p_1 \leqslant \sqrt[l]{N}$ with $(p_1, m) = 1$, we have

$$(1-\varepsilon)\frac{\pi_{l-1}([N/p_1])}{\varphi(m)} < \pi_{l-1}([N/p_1]; m, a(p_1)) < (1+\varepsilon)\frac{\pi_{l-1}([N/p_1])}{\varphi(m)}.$$

Since $\pi_l(N) = \sum_{p_1 \leqslant \sqrt[l]{N}} \pi_{l-1}([N/p_1])$, it follows that

$$(1-\varepsilon)\frac{\pi_l(N)}{\varphi(m)} < \pi_l(N; m, a) < (1+\varepsilon)\frac{\pi_l(N)}{\varphi(m)}$$

for every $N \geqslant N_\varepsilon$, so we have (19). □

LEMMA 3.2. *For every $l \geqslant 2$, we have*

$$\#\{p^{(l)} \in \mathbb{P}_l^N : (p^{(l)}, N) > 1\} = o(\pi_l(N)). \tag{20}$$

Proof. Notice that $\#\{p^{(l)} \in \mathbb{P}_l^N : (p^{(l)}, N) > 1\} \leqslant \sum_{p|N} \pi_{l-1}(N/p)$. Therefore, using (7),

$$\#\{p^{(l)} \in \mathbb{P}_l^N : (p^{(l)}, N) > 1\} = O\bigg(\sum_{p|N} \frac{N/p}{\log(N/p)} \frac{(\log\log(N/p))^{l-2}}{(l-2)!}\bigg)$$
$$= O\bigg(\frac{N}{\log N} \frac{(\log\log(N))^{l-1}}{(l-1)!} \frac{(l-1)}{\log\log N} \sum_{p|N} \frac{\log N}{p \log(N/p)}\bigg).$$

So again, by (7), the result will follow by showing that

$$\frac{1}{\log\log N} \sum_{p|N} \frac{\log N}{p \log(N/p)} = o(1).$$

Note that

$$\sum_{p|N} \frac{\log N}{p \log(N/p)} = \sum_{\substack{p|N \\ p \leqslant N^{1/2}}} \frac{\log N}{p \log(N/p)} + \sum_{\substack{p|N \\ p > N^{1/2}}} \frac{\log N}{p \log(N/p)},$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

and that the second term contains at most one prime p. Moreover, as $l \geqslant 2$ and $(p^{(l)}, N) > 1$, the number N is not prime, so $N/p \geqslant 2$. Using this, we get

$$\sum_{p|N} \frac{\log N}{p \log(N/p)} \leqslant 2 \sum_{p|N} \frac{1}{p} + \frac{1}{\log 2} \frac{\log N}{N^{1/2}} = O(\log \log \log N),$$

as $\sum_{p|N}(1/p) = O(\log \log \log N)$; see, for example, [11]. This finishes the proof. □

Now we repeat the scheme of the construction from §2 almost word for word, although we have to take care how we choose k_{t+1}.

First of all, we require that k_{t+1} is large enough so that

$$\pi_l(n_{t+1}; n_t, a) \leqslant 2\frac{\pi_l(n_{t+1})}{\varphi(n_t)} \quad \text{for every } 0 \leqslant a < n_t \text{ with } (a, n_t) = 1, \tag{21}$$

$$\sum_{0 \leqslant a < n_t, (a,n_t)>1} \pi_l(n_{t+1}; n_t, a) \leqslant \frac{\varepsilon}{8}\pi_l(n_{t+1}), \tag{22}$$

$$\#\{p^{(l)} \in \mathbb{P}_l^{n_{t+1}}, (p^{(l)}, n_{t+1}) > 1\} = o(\pi_l(n_{t+1})). \tag{23}$$

The existence of such k_{t+1} is guaranteed by Lemmas 3.1 and 3.2.

Next, we replace (12) by

$$\#\{p^{(l)} \in \mathbb{P}_l^{n_t} : x_t(p^{(l)}) = ?\} \geqslant \tfrac{1}{2}\pi_l(n_t)$$

and requiring (instead of (16)) that for $(a, n_t) = 1$ we have

$$\#(\{a + jn_t : 0 \leqslant j < k_{t+1}\} \cap \mathbb{P}_l) \leqslant 2\frac{\pi_l(n_{t+1})}{\varphi(n_t)};$$

cf. (21). Furthermore, we replace (15) by the requirement that

$$\#\{p^{(l)} \in \mathbb{P}_l^{n_t} : p^{(l)} \equiv a \bmod n_t \text{ with } (a, n_t) > 1\} \leqslant \tfrac{1}{8}\pi_l(n_{t+1});$$

cf. (22). To carry over the previous proof, it remains to show that

$$\frac{1}{\pi_l(n_t)} \sum_{p^{(l)} \in \mathbb{P}_l^{n_t}, (p^{(l)}, n_t) > 1} (-1)^{x(p^{(l)})} = o(1).$$

This follows from (23) applied in the previous step of the construction.

4. *Regular Toeplitz subshifts which satisfy a PNT (proof of Theorem B)*

Let $x \in \mathcal{A}^{\mathbb{Z}}$ be a regular Toeplitz sequence. Then, for every $k \in \mathbb{N}$, there is an n_k-periodic sequence $x_k \in (\mathcal{A} \cup \{?\})^{\mathbb{Z}}$ so that

$$x_k(j) \neq ? \text{ implies } x(j) = x_k(j) = x_l(j) \quad \text{for all } l \geqslant k,$$

and

$$?_k = ?_k(x) := \#\{0 \leqslant j < n_k : x_k(j) = ?\} = o(n_k).$$

For every Toeplitz sequence $x \in \mathcal{A}^{\mathbb{Z}}$ and natural m, let us consider a new Toeplitz sequence $x^{(m)} \in (\mathcal{A}^{2m+1})^{\mathbb{Z}}$ given by

$$x^{(m)}(j) = (x(j - m), \ldots, x(j + m)) \quad \text{for every } j \in \mathbb{Z}.$$

If $(n_t)_{t\geqslant 1}$ is a periodic structure of x, then it is also a periodic structure of $x^{(m)}$. Moreover,

$$?_k(x^{(m)}) \leqslant (2m+1)?_k(x) \quad \text{for every } k \geqslant 1. \tag{24}$$

Hence, the regularity of x implies the regularity of $x^{(m)}$.

Theorem B follows directly from Lemma A.1 and the following result.

THEOREM 4.1. *Suppose that (X_x, S) is a Toeplitz system such that*

$$?_k = o(\varphi(n_k)).$$

Then (X_x, S) satisfies a PNT.

Proof. To show a PNT for (X_x, S), it suffices to show that for every continuous $F : X_x \to \mathbb{C}$ and every $\varepsilon > 0$ there exists N_ε so that for every $N, M \geqslant N_\varepsilon$ and every $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{\pi(N)} \sum_{p\leqslant N} F(S^{p+r}x) - \frac{1}{\pi(M)} \sum_{p\leqslant M} F(S^{p+r}x) \right| < \varepsilon. \tag{25}$$

Note that the above is stronger than what is needed as it shows that the convergence in (1) is uniform in $y \in X_x$. We first assume that $F : X_x \to \mathbb{R}$ depends only on the zero coordinate, that is, $F(y) = f(y(0))$ for some $f : \mathcal{A} \to \mathbb{R}$.

Fix $\varepsilon > 0$. Fix also $k \geqslant 1$ so that

$$?_k < \frac{\varepsilon}{8}\varphi(n_k). \tag{26}$$

Next choose N_ε such that for every $N \geqslant N_\varepsilon$, we have

$$\left| \pi(N; n_k, a) - \frac{\pi(N)}{\varphi(n_k)} \right| < \frac{\varepsilon}{8} \frac{\pi(N)}{\varphi(n_k)} \quad \text{for all } a \in \mathbb{Z} \text{ with } (a, n_k) = 1, \tag{27}$$

$$\#\{p \leqslant N : p|n_k\} \leqslant \log n_k < \frac{\varepsilon}{8}\pi(N). \tag{28}$$

We will show that for all $N \geqslant N_\varepsilon$ and $r \in \mathbb{Z}$ we have

$$\left| \frac{1}{\pi(N)} \sum_{p\leqslant N} F(S^{p+r}x) - \frac{1}{\varphi(n_k)} \sum_{\substack{0\leqslant a<n_k \\ (a-r,n_k)=1 \\ x_k(a)\neq ?}} F(S^a x) \right| \leqslant \varepsilon \|F\|_{\sup}, \tag{29}$$

which implies (25).

Recall that $x_k \in (\mathcal{A} \cup \{?\})^{\mathbb{Z}}$ is an n_k-periodic sequence (used to construct x at stage k). If for some $a \in \mathbb{Z}$ we have

$$x_k(a) \neq ?,$$

then

$$x(a + j \cdot n_k) = x_k(a) \quad \text{for every } j \in \mathbb{Z}.$$

This implies that if $p \leqslant N$ and $x_k(p + r \bmod n_k) \neq ?$, then

$$F(S^{p+r}x) = F(S^{p+r \bmod n_k}x). \tag{30}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

Note that

$$\begin{aligned}\#\{p \leqslant N : x_k(p + r \bmod n_k) = ?\}\\ \leqslant \sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)=1\\x_k(a)=?}} \#\{p \leqslant N : p = a - r \bmod n_k\}\\ + \sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)>1}} \#\{p \leqslant N : p = a - r \bmod n_k\}.\end{aligned}$$

Assume that $N \geqslant N_\varepsilon$. By (27) and (28), for every integer v with $(v, n_k) = 1$, we have

$$\#\{p \leqslant N : p = v \bmod n_k\} = \pi(N; n_k, v) \leqslant \left(1 + \frac{\varepsilon}{8}\right) \frac{\pi(N)}{\varphi(n_k)}$$

and

$$\sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)>1}} \#\{p \leqslant N : p = a - r \bmod n_k\} \leqslant \#\{p \leqslant N : p|n_k\} < \frac{\varepsilon}{8}\pi(N), \tag{31}$$

where the left inequality follows from the fact that if $(a - r, n_k) > 1$ and $p_a = a - r \bmod n_k$ for a prime p_a, then $(a - r, n_k) = p_a$ and

$$\{p \leqslant N : \ p = a - r \bmod n_k\} = \{p_a\}.$$

It follows that (use also (26))

$$\begin{aligned}&\#\{p \leqslant N : x_k(p + r \bmod n_k) = ?\}\\ &\leqslant \#\{0 \leqslant a < n_k : (a - r, n_k) = 1, x_k(a) =?\} \left(1 + \frac{\varepsilon}{8}\right) \frac{\pi(N)}{\varphi(n_k)} + \frac{\varepsilon}{8}\pi(N)\\ &\leqslant ?_k \left(1 + \frac{\varepsilon}{8}\right) \frac{\pi(N)}{\varphi(n_k)} + \frac{\varepsilon}{8}\pi(N) \leqslant \frac{\varepsilon}{2}\pi(N).\end{aligned}$$

Let

$$P_N := \{p \leqslant N : x_k(p + r \bmod n_k) \neq ?\}.$$

Then by the above, for every $N \geqslant N_\varepsilon$,

$$\left|\frac{1}{\pi(N)} \sum_{p\leqslant N} F(S^{p+r}x) - \frac{1}{\pi(N)} \sum_{p\in P_N} F(S^{p+r}x)\right| \leqslant \frac{\varepsilon}{2}\|F\|_{\sup}. \tag{32}$$

But by (30),

$$\begin{aligned}\sum_{p\in P_N} F(S^{p+r}x) &= \sum_{\substack{0\leqslant a<n_k\\x_k(a)\neq ?}} \sum_{\substack{p\leqslant N\\p\equiv a-r \bmod n_k}} F(S^a x)\\ &= \sum_{\substack{0\leqslant a<n_k\\x_k(a)\neq ?}} F(S^a x)\#\{p \leqslant N, p = a - r \bmod n_k\}.\end{aligned}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

If $(a-r, n_k) = 1$, then again by (27), we have

$$\left|\#\{p \leqslant N, p = a - r \bmod n_k\} - \frac{\pi(N)}{\varphi(n_k)}\right| = \left|\pi(N; n_k, a-r) - \frac{\pi(N)}{\varphi(n_k)}\right| < \frac{\varepsilon}{8}\frac{\pi(N)}{\varphi(n_k)}.$$

In view of (31), it follows that

$$\begin{aligned}
&\left|\frac{1}{\pi(N)}\sum_{p\in P_N} F(S^{p+r}x) - \frac{1}{\varphi(n_k)}\sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)=1\\x_k(a)\neq ?}} F(S^a x)\right| \\
&= \left|\sum_{\substack{0\leqslant a<n_k\\x_k(a)\neq ?}} F(S^a x)\frac{\pi(N; n_k, a-r)}{\pi(N)} - \frac{1}{\varphi(n_k)}\sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)=1\\x_k(a)\neq ?}} F(S^a x)\right| \\
&\leqslant \frac{1}{\pi(N)}\sum_{\substack{0\leqslant a<n_k\\(a-r,n_k)=1\\x_k(a)\neq ?}} |F(S^a x)|\left|\pi(N; n_k, a-r) - \frac{\pi(N)}{\varphi(n_k)}\right| + \frac{\varepsilon}{8}\|F\|_{\sup} \\
&\leqslant \|F\|_{\sup}\left(\frac{\varepsilon}{8}\frac{\#\{0\leqslant a<n_k : x_k(a)\neq ?, (a-r,n_k)=1\}}{\varphi(n_k)} + \frac{\varepsilon}{8}\right) \leqslant \|F\|_{\sup}\frac{\varepsilon}{2}.
\end{aligned}$$

Together with (32), this gives (29), which completes the proof in the case of F depending only on the zero coordinate.

Now suppose that $F : X_x \to \mathbb{C}$ depends only on finitely many coordinates. Then there exist natural m and $f : \mathcal{A}^{2m+1} \to \mathbb{C}$ such that $F(y) = f(y(-m), \ldots, y(m))$ for every $y = (y(k))_{k\in\mathbb{Z}} \in X_x$. Denote by $X_{x^{(m)}} \subset (\mathcal{A}^{2m+1})^{\mathbb{Z}}$ the orbit closure of $x^{(m)} \in (\mathcal{A}^{2m+1})^{\mathbb{Z}}$. Then every $y^{(m)} \in X_{x^{(m)}}$ is of the form $y^{(m)}(k) = (y(k-m), \ldots, y(k+m))$ for some $y = (y(k))_{k\in\mathbb{Z}} \in X_x$.

In view of (24), $(X_{x^{(m)}}, S)$ is a regular Toeplitz shift with $?_k(x^{(m)}) = o(\varphi(n_k))$. Let us consider $\bar{F} : X_{x^{(m)}} \to \mathbb{C}$ given by $\bar{F}(y^{(m)}) = f(y^{(m)}(0)) = f(y(-m), \ldots, y(m))$ for $y^{(m)} \in X_{x^{(m)}}$. Since $\bar{F}$ depends only on the zero coordinate, by (25) applied to $x^{(m)}$ and the map $\bar{F}$, for every $\varepsilon > 0$ there exists N_ε such that for $N, M \geqslant N_\varepsilon$ we have

$$\begin{aligned}
&\left|\frac{1}{\pi(N)}\sum_{p\leqslant N} F(S^{p+r}x) - \frac{1}{\pi(M)}\sum_{p\leqslant M} F(S^{p+r}x)\right| \\
&= \left|\frac{1}{\pi(N)}\sum_{p\leqslant N} \bar{F}(S^{p+r}x^{(m)}) - \frac{1}{\pi(M)}\sum_{p\leqslant M} \bar{F}(S^{p+r}x^{(m)})\right| < \varepsilon.
\end{aligned}$$

Thus (25) holds for every $F : X_x \to \mathbb{C}$ depending only on finitely many coordinates. As the set of such functions is dense in $C(X_x)$, (25) also holds for every $F \in C(X_x)$, which completes the proof. □

As $\varphi(n) \to \infty$ when $n \to \infty$, we obtain the following result.

COROLLARY 4.2. *If x is Toeplitz for which the sequence $(?_k)$ is bounded then (X_x, S) satisfies a PNT.*

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

5. *Toeplitz subshifts for which a P_lNT holds*

THEOREM 5.1. *Suppose that (X_x, S) is a Toeplitz system such that*

$$?_k = o(\varphi(n_k)).$$

Then, for every $F \in C(X_x)$ and $y \in X_x$, the limit

$$\lim_{N\to\infty} \frac{1}{\pi_l(N)} \sum_{p^{(l)}\in\mathbb{P}_l^N} F(S^{p^{(l)}} y) \text{ exists.}$$

Proof. The proof proceeds along the same lines as the proof of Theorem 4.1. It relies on the following analogue of (29): for every $\varepsilon > 0$ there exists a natural N_ε such that for all $N \geqslant N_\varepsilon$ and $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{\pi_l(N)} \sum_{p^{(l)}\in\mathbb{P}_l^N} F(S^{p^{(l)}+r} x) - \frac{1}{\varphi(n_k)} \sum_{\substack{0\leqslant a<n_k \\ (a-r,n_k)=1 \\ x_k(a)\neq ?}} F(S^a x) \right| \leqslant \varepsilon \|F\|_{\sup}. \tag{33}$$

In turn, the proof of (29) is based on only two elements: (27) and (31). Their l-almost prime counterparts follow directly from (19) and (18), respectively. Now we repeat the arguments of the proof of (29) almost word for word, replacing (27) and (31) by their l-almost prime counterparts. □

Remark 5.2. In view of (29) and (33), under the assumption $?_k = o(\varphi(n_k))$, we have

$$\lim_{N\to\infty} \frac{1}{\pi_l(N)} \sum_{p^{(l)}\in\mathbb{P}_l^N} F(S^{p^{(l)}} y) = \lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} F(S^p y)$$

for every $F \in C(X_x)$ and $y \in X_x$, so a PNT and a P_lNT fully coincide for this class of regular Toeplitz systems.

6. *Ergodic averages along polynomial times*

Let P be a monic polynomial of degree $d > 1$ with non-negative integer coefficients. The leading coefficient of P equals 1. This assumption is only for simplicity. In fact, Theorem 6.8 below is true whenever the set of (non-zero) coefficients of $P - P(0)$ is coprime; see the proof of Corollary 6.3 and the assumptions of the Albis theorem in [17]. Note that, under these assumptions, $P(\cdot)$ is a strictly increasing function on $\mathbb{N}$. For every $n \in \mathbb{N}$, let

$$R_n^P := \{0 \leqslant a < n : a = P(m) \bmod n \text{ for some } m \in \mathbb{N}\} \quad \text{and} \quad \psi^P(n) := \#R_n^P.$$

For all $N, n \in \mathbb{N}$ and $a \in R_n^P$, let

$$\rho^P(N; n, a) = \#\{1 \leqslant m \leqslant N : P(m) = a \bmod n\}$$

and

$$\rho^P(n, a) := \rho^P(n; n, a), \quad \rho^P(n) := \max_{a\in R_n^P} \rho^P(n; n, a).$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

LEMMA 6.1. *The function ψ^P is multiplicative, that is, $\psi^P(n_1n_2) = \psi^P(n_1)\psi^P(n_2)$ if $(n_1, n_2) = 1$. If $a \in \mathbb{Z}/n\mathbb{Z}$, $n_1, \ldots, n_k$ are pairwise coprime and $n = n_1 \cdots n_k$ then $a \in R_n^P$ if and only if $a_i \in R_{n_i}^P$ for $i = 1, \ldots, k$, where $0 \leqslant a_i < n_i$ is the remainder of a when divided by n_i (i.e., $0 \leqslant a_i < n_i$ and $a_i = a \bmod n_i$). Moreover,*

$$\rho^P(n, a) = \prod_{i=1}^{k} \rho^P(n_i, a_i). \tag{34}$$

Proof. Note that the multiplicativity of ψ^P follows from the second part of the lemma.

Moreover, note that $a \in R_n^P$ if and only if $a = P(m) \bmod n$ for some $0 \leqslant m < n$. Indeed, if $a = P(m) \bmod n$ for some $m \in \mathbb{N}$, then $a = P(m') \bmod n$, where $0 \leqslant m' < n$ is the remainder of m when divided by n.

If $a \in R^P_{n_1 \cdots n_k}$, that is, $a = P(m) \bmod n_1 \cdots n_k$ for some $0 \leqslant m < n$, then $a_i = a = P(m) = P(m_i) \bmod n_i$ for every $i = 1, \ldots, k$, where $0 \leqslant m_i < n_i$ is the remainder of m when divided by n_i.

Now, suppose $a \in \mathbb{Z}/n\mathbb{Z}$, $a_i = a \bmod n_i$ and $a_i \in R^P_{n_i}$ for $i = 1, \ldots, k$. Then, for every $i = 1, \ldots, k$, there exists $0 \leqslant m_i < n_i$ such that $a_i = P(m_i) \bmod n_i$. By the Chinese remainder theorem, there exists a unique $0 \leqslant m < n$ such that $m = m_i \bmod n_i$ for $i = 1, \ldots, k$. It follows that

$$P(m) = P(m_i) = a_i = a \bmod n_i \quad \text{for all } i = 1, \ldots, k.$$

This yields $a = P(m) \bmod n_1 \cdots n_k$ and $a \in R_n^P$.

The argument above also shows (34). □

Remark 6.2. Note that in the argument above we used the fact that the a_i determine a as by the Chinese remainder theorem there exists only one $0 \leqslant a < n$ such that $a = a_i \bmod n_i$ for each $i = 1, \ldots, k$.

For any natural n, denote by $\omega(n)$ the number of its prime divisors (counted without multiplicities) and by $p(n)$ the product of its prime divisors.

COROLLARY 6.3. *The arithmetic function ρ^P is multiplicative and $\rho^P(n) \leqslant (d^{\omega(n)}/p(n))n$.*

Proof. The multiplicativity of ρ^P follows directly from (34). By the Albis theorem (see [**17**, Corollary 3 of Theorem 1.23]), for any prime number we have $\rho^P(p^n) \leqslant dp^{n-1}$. (Note that compared to the notation on [**17**], we have

$$\rho^P(n; n, a) = \lambda_{P-a}(n), \qquad \rho^P(n) = \max_{a \in R_n^P} \lambda_{P-a}(n);$$

the estimate on λ_P in [**17**] depends *only* on the degree of the polynomial.) This result combined with the multiplicativity of ρ^P gives the required bound of $\rho^P(n)$. □

LEMMA 6.4. *For all* $n \in \mathbb{N}$, $a \in R_n^P$ *and* $N \geqslant P(n)$, *we have*

$$\rho^P(n,a)\bigg(\frac{P^{-1}(N)}{n}-1\bigg) \leqslant \#\{m \in \mathbb{N} : 1 \leqslant P(m) \leqslant N,\ P(m) = a \bmod n\}$$
$$\leqslant \rho^P(n,a)\bigg(\frac{P^{-1}(N)}{n}+1\bigg).$$

Proof. Let $s := \rho^P(n,a)$ and let $1 \leqslant m_1 < \ldots < m_s \leqslant n$ be all numbers such that $P(m_i) = a \bmod n$. Note that a natural number m satisfies $P(m) \leqslant N$ and $P(m) = a \bmod n$ if and only if $m = jn + r$ with $0 \leqslant j \leqslant (P^{-1}(N) - r)/n$ and $0 < r \leqslant n$ satisfies $P(r) = a \bmod n$. Thus, $r = m_i$ for some $i = 1, \ldots, s$. It follows that

$$\rho := \#\{m \in \mathbb{N} : 1 \leqslant P(m) \leqslant N,\ P(m) = a \bmod n\}$$
$$= \sum_{i=1}^{s}\bigg(\bigg[\frac{P^{-1}(N)-m_i}{n}\bigg]+1\bigg).$$

Since

$$\frac{P^{-1}(N)}{n} - 1 \leqslant \frac{P^{-1}(N)-m_i}{n} < \bigg[\frac{P^{-1}(N)-m_i}{n}\bigg] + 1$$
$$\leqslant \frac{P^{-1}(N)-m_i}{n} + 1 < \frac{P^{-1}(N)}{n} + 1,$$

by summing up, this gives

$$s\bigg(\frac{P^{-1}(N)}{n}-1\bigg) \leqslant \rho \leqslant s\bigg(\frac{P^{-1}(N)}{n}+1\bigg).$$

□

Remark 6.5. As P is an increasing function, we can apply the above inequalities to $P(N)$ instead of N (as $P(N) \geqslant N$). Then $P(m) \leqslant P(N)$ if and only if $m \leqslant N$, and the result of the lemma implies

$$\rho^P(n,a)\bigg(\frac{N}{n}-1\bigg) \leqslant \rho^P(N; n, a) \leqslant \rho^P(n,a)\bigg(\frac{N}{n}+1\bigg).$$

We now focus on the simplest case when $P(n) = n^2$. We continue to write R for R^P, ψ for ψ^P and ρ for ρ^P. In view of [17, Theorems 1.27 and 1.30], we have the following result.

PROPOSITION 6.6. *For every prime number* $p > 2$, *for every* $a \in R_{p^N}$, *where* $N = 2n$ *or* $2n+1$, *we have*

$$\rho(p^N, a) = \begin{cases} 2 & \text{if } a = a' \bmod p \text{ for } a' \in R_p\setminus\{0\}, \\ 2p^r & \text{if } a = p^{2r}a' \text{ and } a' = a'' \bmod p \text{ for } a'' \in R_p\setminus\{0\}, \\ p^n & \text{if } a = 0. \end{cases}$$

Moreover, we have

$$\psi(p^{2n+1}) = \frac{p^{2n+2}+2p+1}{2(p+1)} \quad \text{and} \quad \psi(p^{2n}) = \frac{p^{2n+1}+p+2}{2(p+1)}. \tag{35}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

Furthermore, if $p = 2$ then

$$\rho(2, a) = 1 \ \textit{for all } a \in R_2, \qquad \rho(4, a) = 2 \ \textit{for all } a \in R_4,$$

and for any $N \geqslant 3$, where $N = 2n$ or $2n + 1$, for every $a \in R_{2^N}$, we have

$$\rho(2^N, a) = \begin{cases} 4 & \textit{if } a = 1 \bmod 8, \\ 4 \cdot 2^r & \textit{if } a = 2^{2r}a',\ 2r \leqslant N - 3,\ a' = 1 \bmod 8, \\ 2 \cdot 2^r & \textit{if } a = 2^{2r}a',\ 2r = N - 2,\ a' = 1 \bmod 4, \\ 2^r & \textit{if } a = 2^{2r}a',\ 2r = N - 1,\ a' = 1 \bmod 2, \\ 2^n & \textit{if } a = 0. \end{cases}$$

Moreover,

$$\psi(2^{2n}) = \frac{2^{2n-1} + 4}{3} \quad \textit{and} \quad \psi(2^{2n+1}) = \frac{2^{2n} + 5}{3}.$$

Proof. We obtain (35) by using the formulas for the values of ρ and counting the number of the possibilities in each row, so for $N = 2n$, we have

$$\begin{aligned} \psi(p^{2n}) &= \frac{p-1}{2} p^{2n-1} + \sum_{r=1}^{n-1} \frac{p-1}{2} p^{2n-2r-1} + 1 \\ &= 1 + p\frac{p-1}{2} \sum_{r=0}^{n-1} p^{2(n-r-1)} = 1 + \frac{p(p-1)}{2} \frac{(p^2)^n - 1}{p^2 - 1} = \frac{p^{2n+1} + p + 2}{2(p+1)}. \end{aligned}$$

□

COROLLARY 6.7. *For every natural $n \geqslant 2$, we have $\rho(n) \leqslant 4\sqrt{n}$. Moreover, if n is square-free, then $\rho(n) \leqslant 2^{\omega(n)}$.*

Proof. By a direct inspection of the formulas in Proposition 6.6, we obtain

$$\rho(2^N) \leqslant 2\sqrt{2^N}, \qquad \rho(3^N) \leqslant 2\sqrt{3^N},$$

but for all $p \geqslant 5$, we have

$$\rho(p^N) \leqslant \sqrt{p^N}.$$

Indeed, for the cases $a = a' \bmod p$ (for $a' \in R_p \backslash \{0\}$) and $a = 0$, this is direct. For the case $\rho(p^N, a) = 2p^r$, we have $a = p^{2r}a' < p^N$, so $2r \leqslant N - 1$ and then indeed $2p^r \leqslant p^{N/2}$.

The second inequality follows directly from $\rho(p) \leqslant 2$. □

For future purposes, we are interested in cases (in Proposition 6.6) which give possibly smallest values for the function ρ, hence, for every prime number p and any natural n, let

$$\widetilde{R}_{p^n} := \begin{cases} \{0 \leqslant a < p^n : a = a' \bmod p \text{ for } a' \in R_p \setminus \{0\}\} & \text{if } p > 2, \\ R_2 & \text{if } p^n = 2, \\ R_4 & \text{if } p^n = 4, \\ \{0 \leqslant a < 2^n : a = 1 \bmod 8\} & \text{if } n \geqslant 3. \end{cases}$$

By Proposition 6.6, $\widetilde{R}_{p^n} \subset R_{p^n}$.

Let $n = p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}$ be the canonical representation of n. Let

$$\Phi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k^{m_k}\mathbb{Z}$$

be the canonical ring isomorphism. Recall (cf. Lemma 6.1 and Remark 6.2) that Φ establishes a one-to-one correspondence between R_n and $R_{p_1^{m_1}} \times \cdots \times R_{p_k^{m_k}}$. Set

$$\widetilde{R}_n := \Phi^{-1}(\widetilde{R}_{p_1^{m_1}} \times \cdots \times \widetilde{R}_{p_k^{m_k}})$$

and

$$\widetilde{\psi}(n) := \#\widetilde{R}_n.$$

Then, clearly, $\widetilde{\psi}$ is a multiplicative function. Moreover, by Proposition 6.6, for each $a \in \widetilde{R}_{p^N}$, we have

$$\rho(p^N, a) = \begin{cases} 1 & \text{if } p^N = 2, \\ 2 & \text{if } p^N = 2 \text{ or } p > 2, \\ 4 & \text{if } p = 2 \text{ and } N \geqslant 3. \end{cases}$$

Hence, in view of (34), for every $a \in \widetilde{R}_n$, we have

$$\tfrac{1}{2} \cdot 2^{\omega(n)} \leqslant \rho(n, a) \leqslant 2 \cdot 2^{\omega(n)}. \tag{36}$$

Moreover, by definition,

$$\widetilde{\psi}(p^n) := \begin{cases} p^{n-1}(p-1)/2 & \text{if } p > 2, \\ 2 & \text{if } p^n = 2, \\ 2 & \text{if } p^n = 4, \\ 2^{n-3} & \text{if } p = 2 \text{ and } n \geqslant 3. \end{cases}$$

It follows that

$$\frac{1}{2} \prod_{p|n} \left(1 - \frac{1}{p}\right) \leqslant \frac{2^{\omega(n)} \widetilde{\psi}(n)}{n} \leqslant 4 \prod_{p|n} \left(1 - \frac{1}{p}\right). \tag{37}$$

To obtain these inequalities, for $n = p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}$, write

$$\frac{2^{\omega(n)} \widetilde{\psi}(n)}{n} = \prod_{i=1}^{k} \frac{2\tilde{\psi}(p_i^{m_i})}{p_i^{m_i}}$$

and apply the formula above.

6.1. *Polynomial ergodic theorem.* In the result below P is a monic polynomial of degree $d > 1$ with non-negative integer coefficients.

THEOREM 6.8. *Suppose that (X_x, S) is a Toeplitz system such that*

$$?_k = o(n_k/\rho^P(n_k)). \tag{38}$$

Then, for every continuous map $F : X_x \to \mathbb{C}$ and $y \in X_x$, the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m\leqslant N} F(S^{P(m)}y) \tag{39}$$

exists.

Proof. To show (39), we need to prove that for every $\varepsilon > 0$ there exists N_ε so that for every $N, M \geqslant N_\varepsilon$ and every $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{N} \sum_{m\leqslant N} F(S^{P(m)+r}x) - \frac{1}{M} \sum_{m\leqslant M} F(S^{P(m)+r}x) \right| < \varepsilon. \tag{40}$$

We first assume that $F : X_x \to \mathbb{R}$ depends only on the zero coordinate, that is, $F(y) = f(y(0))$ for some $f : \mathcal{A} \to \mathbb{R}$.

Fix $\varepsilon > 0$. Choose $k \geqslant 1$ so that

$$?_k < \frac{\varepsilon}{8} \frac{n_k}{\rho^P(n_k)}. \tag{41}$$

Next, choose $N_\varepsilon \geqslant 8n_k^2/\varepsilon$. Then, in view of Remark 6.5 (and the choice of N_ε), for every $N \geqslant N_\varepsilon$ and $a \in R^P_{n_k}$, we have

$$\left| \rho^P(N; n_k, a) - \rho^P(n_k, a)\frac{N}{n_k} \right| < \rho^P(n_k) \leqslant n_k \leqslant \frac{\varepsilon}{8} \frac{N}{n_k}. \tag{42}$$

From now on, we write that an integer number v belongs to $R^P_{n_k}$ if there exists $0 \leqslant v' < n_k$ such that $v' = v \bmod n_k$ and $v' \in R^P_{n_k}$. We will show that for all $N \geqslant N_\varepsilon$ and $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{N} \sum_{m\leqslant N} F(S^{P(m)+r}x) - \frac{1}{n_k} \sum_{\substack{0\leqslant a<n_k \\ a-r\in R^P_{n_k} \\ x_k(a)\neq ?}} \rho^P(n_k, a-r)F(S^a x) \right| \leqslant \varepsilon \|F\|_{\sup}, \tag{43}$$

and this implies (40).

Recall that $x_k \in (\mathcal{A} \cup \{?\})^{\mathbb{Z}}$ is an n_k-periodic sequence (used to construct x at stage k). Note that for every $a \in \mathbb{Z}$, we have

$$x_k(a) \neq ? \Rightarrow x(a + j \cdot n_k) = x_k(a) \quad \text{for every } j \in \mathbb{Z}.$$

This implies that if $m \leqslant N$ and $x_k(P(m) + r \bmod n_k) \neq ?$, then

$$F(S^{P(m)+r}x) = F(S^{P(m)+r \bmod n_k}x). \tag{44}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

Therefore,

$$\begin{aligned}&\#\{m \leqslant N : x_k(P(m)+r \bmod n_k) = ?\}\\ &\quad= \sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)=?}} \#\{m \leqslant N : P(m) = a-r \bmod n_k\} = \sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)=?}} \rho^P(N; n_k, a-r).\end{aligned}$$

Assume that $N \geqslant N_\varepsilon$. By (42), for every integer $v \in R^P_{n_k}$, we have

$$\rho^P(N; n_k, v) \leqslant 2\rho^P(n_k)\frac{N}{n_k}.$$

In view of (41), it follows that

$$\begin{aligned}&\#\{m \leqslant N : x_k(P(m)+r \bmod n_k) = ?\}\\ &\quad\leqslant \#\{0 \leqslant a < n_k : a-r \in R^P_{n_k}, x_k(a) =?\}2\rho^P(n_k)\frac{N}{n_k}\\ &\quad\leqslant 2?_k\rho^P(n_k)\frac{N}{n_k} \leqslant \frac{\varepsilon}{4}N.\end{aligned}$$

Let

$$U_N := \{m \leqslant N : x_k(P(m) + \bmod\ n_k) \neq ?\}.$$

Then by the above, for every $N \geqslant N_\varepsilon$,

$$\left|\frac{1}{N}\sum_{m\leqslant N} F(S^{P(m)+r}x) - \frac{1}{N}\sum_{m\in U_N} F(S^{P(m)+r}x)\right| \leqslant \frac{\varepsilon}{4}\|F\|_{\sup}. \tag{45}$$

But by (44),

$$\begin{aligned}\sum_{m\in U_N} F(S^{P(m)+r}x) &= \sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} \sum_{\substack{m\leqslant N\\ P(m)=a-r \bmod n_k}} F(S^a x)\\ &= \sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} F(S^a x)\#\{m \leqslant N : P(m) = a-r \bmod n_k\}\\ &= \sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} F(S^a x)\rho^P(N; n_k, a-r).\end{aligned}$$

By (42), we have

$$\left|\rho^P(N; n_k, a-r) - \rho^P(n_k, a-r)\frac{N}{n_k}\right| < \frac{\varepsilon}{8}\frac{N}{n_k}.$$

It follows that

$$\begin{aligned}
&\bigg|\frac{1}{N}\sum_{m\in U_N} F(S^{P(m)+r}x) - \frac{1}{n_k}\sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} \rho^P(n_k, a-r)F(S^a x)\bigg|\\
&= \bigg|\frac{1}{N}\sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} F(S^a x)\rho^P(N; n_k, a-r) - \frac{1}{n_k}\sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} \rho^P(n_k, a-r)F(S^a x)\bigg|\\
&\leqslant \frac{1}{N}\sum_{\substack{0\leqslant a<n_k\\ a-r\in R^P_{n_k}\\ x_k(a)\neq ?}} |F(S^a x)|\bigg|\rho^P(N; n_k, a-r) - \rho^P(n_k, a-r)\frac{N}{n_k}\bigg|\\
&\leqslant \|F\|_{\sup}\frac{\varepsilon}{8}\frac{\#\{0\leqslant a<n_k : x_k(a)\neq ?, a-r\in R^P_{n_k}\}}{n_k} \leqslant \|F\|_{\sup}\frac{\varepsilon}{8}.
\end{aligned}$$

Together with (45), this gives (43), which completes the proof in the case of F depending only on the zero coordinate. The rest of the proof runs as in the proof of Theorem 4.1, by passing to the Toeplitz sequences $x^{(m)} \in (\mathcal{A}^{2m+1})^{\mathbb{Z}}$ for $m \geqslant 1$. □

Remark 6.9. Denote by $\mathbb{P}_{(n_t)}$ the set of all prime divisors of elements of the sequence $(n_t)_{t\geqslant 1}$. In view of Corollary 6.3, $?_t = o(p(n_t)/d^{\omega(n_t)})$ implies (38). Unfortunately, if $\mathbb{P}_{(n_t)}$ is finite then the sequence $(p(n_t)/d^{\omega(n_t)})_{t\geqslant 1}$ is bounded, so the little "o" argument does not work and Theorem 6.8 is not applicable. Fortunately, if $\mathbb{P}_{(n_t)}$ is infinite then $p(n_t)/d^{\omega(n_t)} \to +\infty$ as $t \to +\infty$, so Theorem 6.8 applies to a non-trivial class of regular Toeplitz shifts; in particular, it applies when the periodic sequences x_t defining x have a bounded number of '?'.

However, Theorem 6.8 applies to a much wider class of regular Toeplitz shifts when $P(n) = n^2$. Then, by Corollary 6.7, $?_t = o(\sqrt{n_t})$ implies (38). Here the finiteness or infinity of the set $\mathbb{P}_{(n_t)}$ does not matter.

The assumption (38) about the growth of the sequence $(?_t)_{t\geqslant 1}$ is the least restrictive when all n_t are square-free. Then, by the second part of Corollary 6.7, $?_t = o(n_t/2^{\omega(n_t)})$ implies (38). Therefore, $?_t = O({n_t}^{1-(1/\log_2\log_2\log_2 n_t)})$ also implies (38). Indeed, it suffices to show that $2^{\omega(n)} = o(n^{1/\log_2\log_2\log_2 n})$ for square-free numbers $n \to +\infty$. Suppose that $\omega(n) = k$ and denote by $(p_l)_{l\geqslant 1}$ the increasing sequence of all prime numbers. Since

$$\ln n \geqslant \sum_{l=1}^{k} \ln p_l \geqslant k \ln k,$$

we have

$$\begin{aligned}
\frac{2^{\omega(n)}}{n^{(1/\log_2\log_2\log_2 n)}} &= \frac{2^k}{2^{(\log_2 n/\log_2\log_2\log_2 n)}} \leqslant \frac{2^k}{2^{(k\log_2 k/\log_2\log_2(k\log_2 k))}}\\
&= \frac{1}{2^{k(\log_2 k/(\log_2\log_2(k\log_2 k))-1)}}.
\end{aligned}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

As $\log_2 k/(\log_2 \log_2(k \log_2 k)) \to +\infty$ when $k \to +\infty$, this gives $2^{\omega(n)} = o(n^{(1/\log_2 \log_2 \log_2 n)})$.

6.2. *Counter-examples.* We will show that there exists a regular Toeplitz sequence $x \in \{0, 1\}^{\mathbb{Z}}$ with the period structure $(n_t)_{t\geqslant 1}$ satisfying

$$n_{t+1} = k_{t+1}n_k \text{ with } (k_{t+1}, n_t) = 1,\ n_{t+1} \geqslant 2^4 n_t^2 \quad \text{and} \quad \sum_{p\in\mathbb{P}_{(n_t)}} \frac{1}{p} < +\infty \tag{46}$$

and such that

$$\lim_{t\to\infty} \frac{1}{\sqrt{n_t}} \sum_{0\leqslant m<\sqrt{n_t}} F(S^{m^2} x) \text{ does not exist,}$$

where $F(y) = (-1)^{y(0)}$. Let

$$0 < \beta := \frac{1}{16} \prod_{p\in\mathbb{P}_{(n_t)}} \frac{p-1}{p}.$$

By (37), for every $t \geqslant 1$, we have

$$\frac{2^{\omega(n_t)}\widetilde{\psi}(n_t)}{n_t} \geqslant 8\beta. \tag{47}$$

Passing to a subsequence of $(n_t)_{t\geqslant 1}$ (and remembering that $\widetilde{\psi}(m) \to \infty$ when $m \to \infty$), we can assume that

$$\sum_{t\geqslant 1} \frac{1}{\widetilde{\psi}(k_t)} \leqslant \frac{1}{2}.$$

Set

$$\gamma_t := \sum_{l=1}^{t} \frac{1}{\widetilde{\psi}(k_l)} \left(\leqslant \frac{1}{2}\right).$$

At stage t, x is approximated by the infinite concatenation of $x_t[0, n_t - 1] \in \{0, 1, ?\}^{n_t}$ (i.e., we see a periodic sequence of 0, 1, ? with period n_t). Successive '?' will be filled in the next steps of construction of x. We require that:

$$\{0 \leqslant i < n_t : x_t(i) = ?\} \subset R_{n_t}; \tag{48}$$

$$\#\{a \in \widetilde{R}_{n_t} : x_t(a) = ?\} \geqslant (1-\gamma_t)\widetilde{\psi}(n_t); \tag{49}$$

$$\#\{0 \leqslant m < \sqrt{n_t} : x_t(m^2) = ?\} \geqslant \beta\sqrt{n_t}. \tag{50}$$

Recall that, in view of Lemma 6.4 (remembering that $P^{-1}(n_{t+1}) = \sqrt{n_{t+1}}$), (36) and (46), for each $a \in \widetilde{R}_{n_t}$, we have ($\mathbb{N}^2$ stands for $\{m^2 : m \geqslant 0\}$)

$$\begin{aligned}
&\#(\{a + jn_t : 0 < j < k_{t+1}\} \cap \mathbb{N}^2) \geqslant \#(\{m^2 = a \bmod n_t : m^2 < n_{t+1}\}) - 1 \\
&\geqslant \left(\frac{\sqrt{n_{t+1}}}{n_t} - 1\right)\rho(n_t, a) - 1 \geqslant \left(\frac{\sqrt{n_{t+1}}}{n_t} - 1\right)\frac{1}{2}2^{\omega(n_t)} - 1 \\
&\geqslant \frac{1}{2}2^{\omega(n_t)}\left(\frac{\sqrt{n_{t+1}}}{n_t} - 2\right) \geqslant \frac{1}{4}2^{\omega(n_t)}\frac{\sqrt{n_{t+1}}}{n_t},
\end{aligned}$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

so

$$\#(\{a + jn_t : 0 < j < k_{t+1}\} \cap \mathbb{N}^2) \geqslant \frac{2^{\omega(n_t)}}{4} \frac{\sqrt{n_{t+1}}}{n_t}. \tag{51}$$

By the definition of the sets R_n and $\widetilde{R}_n$, we have

$$R_{n_{t+1}} \subset \bigcup_{a \in R_{n_t}} \{a + jn_t : 0 \leqslant j < k_{t+1}\}, \tag{52}$$

$$\widetilde{R}_{n_{t+1}} \subset \bigcup_{a \in \widetilde{R}_{n_t}} \{a + jn_t : 0 \leqslant j < k_{t+1}\}. \tag{53}$$

Moreover, by Lemma 6.1, for every $a \in \widetilde{R}_{n_t}$, we have

$$\#\{i \in \widetilde{R}_{n_{t+1}} : i = a \bmod n_t\} = \#\widetilde{R}_{k_{t+1}} = \widetilde{\psi}(k_{t+1}). \tag{54}$$

We need to describe now which '?' we fill in $x_{t+1}[0, n_{t+1} - 1]$ and how. This block is divided into k_{t+1} subblocks

$$\underbrace{x_t[0, n_t - 1]x_t[0, n_t - 1] \ldots x_t[0, n_t - 1]}_{k_{t+1}}.$$

We fill in *all* '?' in the first block $x_t[0, n_t - 1]$ in such a way as to 'destroy' the convergence of averages in (46) for the time n_t, namely,

$$\frac{1}{\sqrt{n_t}} \sum_{0 \leqslant m < \sqrt{n_t}} F(S^{m^2} x) = \frac{1}{\sqrt{n_t}} \Bigg(\sum_{\substack{m < \sqrt{n_t} \\ x_t(m^2)=0}} 1 - \sum_{\substack{m < \sqrt{n_t} \\ x_t(m^2)=1}} 1 + \sum_{\substack{m < \sqrt{n_t} \\ x_t(m^2)=?}} (-1)^{x(m^2)} \Bigg).$$

And, since the number of m in the last summand is at least $\beta\sqrt{n_t}$ in view of (50), we can fill in these places at stage $t + 1$ to obtain a sum completely different than the known number which we had from stage t. We also fill in (in an arbitrary way) the remaining places in $\{0, \ldots, n_t - 1\}$.

We fill in (in an arbitrary way) all places in $\{n_t, \ldots, n_{t+1} - 1\} \setminus R_{n_{t+1}}$ and only these places, so that (48) will be satisfied at stage $t + 1$.

We must remember that for any $a \in R_{n_t}$ if $x_t(a) \neq ?$ then for every $0 \leqslant j < k_{t+1}$, we have $x_{t+1}(a + jn_t) = x_t(a + jn_t) = x_t(a) \neq ?$. Moreover, for any $a \in \widetilde{R}_{n_t}$ if $x_t(a) = ?$ then for every $0 < j < k_{t+1}$ with $a + jn_t \in \widetilde{R}_{n_{t+1}}$ we have $x_{t+1}(a + jn_t) = ?$. In view of (53), this gives

$$\#\{i \in \widetilde{R}_{n_{t+1}} : x_{t+1}(i) \neq ?\} \leqslant \widetilde{\psi}(n_t) + \sum_{a \in \widetilde{R}_{n_t} : x_t(a) \neq ?} \#\{a + jn_t \in \widetilde{R}_{n_{t+1}} : 0 < j < k_{t+1}\}.$$

In view of (54) and (49), it follows that

$$\#\{i \in \widetilde{R}_{n_{t+1}} : x_{t+1}(i) \neq ?\} \leqslant \widetilde{\psi}(n_t) + (\widetilde{\psi}(k_{t+1}) - 1)\#\{a \in \widetilde{R}_{n_t} : x_t(a) \neq ?\}$$
$$\leqslant \widetilde{\psi}(n_t) + (\widetilde{\psi}(k_{t+1}) - 1)\gamma_t \widetilde{\psi}(n_t) = \left(\gamma_t + \frac{1 - \gamma_t}{\widetilde{\psi}(k_{t+1})}\right) \widetilde{\psi}(n_{t+1}) \leqslant \gamma_{t+1} \widetilde{\psi}(n_{t+1}).$$

Therefore, at stage $t + 1$, (49) is also satisfied.

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

A similar argument combined with (51), (49) and (47) shows that

$$\begin{aligned}\#\{0 \leqslant m^2 < n_{t+1} : x_{t+1}(m^2) = ?\} &= \#\{i \in R_{n_{t+1}} \cap \mathbb{N}^2 : x_{t+1}(i) = ?\} \\ &\geqslant \sum_{a \in R_{n_t} : x_t(a) = ?} \#\{a + jn_t \in R_{n_{t+1}} \cap \mathbb{N}^2 : 0 < j < k_{t+1}\} \\ &\geqslant \sum_{a \in \widetilde{R}_{n_t} : x_t(a) = ?} \frac{2^{\omega(n_t)}}{4} \frac{\sqrt{n_{t+1}}}{n_t} = \frac{\sqrt{n_{t+1}}}{4n_t} 2^{\omega(n_t)} \#\{a \in \widetilde{R}_{n_t} : x_t(a) = ?\} \\ &= (1 - \gamma_t) \frac{\sqrt{n_{t+1}}}{4n_t} 2^{\omega(n_t)} \widetilde{\psi}(n_t) \geqslant \beta \sqrt{n_{t+1}}.\end{aligned}$$

Therefore, at stage $t + 1$, (50) is also satisfied. This completes the construction.

Remark 6.10. In view of (48), in the constructed example of Toeplitz system (X_x, S) we have $?_t \leqslant \psi(n_t)$. Moreover, $\psi(n_t) = o(\varphi(n_t))$. Indeed, by Proposition 6.6, for every prime number p we have $\psi(p^n) \leqslant p^{n-1}(p + 2)/2$. It follows that

$$\frac{\psi(p^n)}{\varphi(p^n)} \leqslant \frac{1}{2} \cdot \frac{p+2}{p-1} \leqslant \frac{3}{4}$$

for all prime $p \geqslant 7$. It follows that

$$\frac{\psi(n_t)}{\varphi(n_t)} = O\bigg(\bigg(\frac{3}{4}\bigg)^{\omega(n_t)}\bigg) = o(1).$$

Consequently, we have $?_t = o(\varphi(n_t))$. Therefore, in view of Theorem 4.1, (X_x, S) satisfies a PNT.

A. *Appendix. The diameter of a tower*

Let $x \in \mathcal{A}^{\mathbb{Z}}$ be a Toeplitz sequence with the periodic structure given by $(n_t)_{t \geqslant 1}$. Recall that

$$\mathrm{Per}_{n_t}(x) = \{a \in \mathbb{Z} : x(a + jn_t) = x(a) \text{ for all } j \in \mathbb{Z}\}.$$

Let $\mathrm{Aper}_{n_t}(x) := \mathbb{Z} \setminus \mathrm{Per}_{n_t}(x)$. Then we define the periodic sequence $x_t \in (\mathcal{A} \cup \{?\})^{\mathbb{Z}}$ by $x_t(k) = x(k)$ if $k \in \mathrm{Per}_{n_t}(x)$ and $x_t(k) = ?$ if $k \in \mathrm{Aper}_{n_t}(x)$. Note that the density of the set $\mathrm{Aper}_{n_t}(x)$ is equal to $?_t/n_t$, where

$$?_t = \#\{0 \leqslant k < n_t : x_t(k) = ?\} = \#(\mathrm{Aper}_{n_t}(x) \cap \{0, 1, \ldots, n_t - 1\}).$$

It follows that the regularity of (X_x, S) is equivalent to $?_t = o(n_t)$.

LEMMA A.1. *For any Toeplitz sequence $x \in \mathcal{A}^{\mathbb{Z}}$ we have*

$$?_t \leqslant \delta(E^t) \leqslant 3?_t \quad \textit{for every } t \geqslant 1.$$

Proof. Note that for every $0 \leqslant j < n_t$ we have

$$E^t_j = \{y \in X_x : y(k - j) = x(k) = x_t(k) \text{ for all } k \in \mathrm{Per}_{n_t}\}.$$

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

Moreover, if $k \in \mathrm{Aper}_{n_t}(x)$ then we can find $y, z \in E^t_j$ so that $y(k-j) \neq z(k-j)$. It follows that

$$\mathrm{diam}(E^t_j) = 2^{-\inf\{|n|:n\in\mathrm{Aper}_{n_t}(x)-\{j\}\}}.$$

Suppose that

$$\mathrm{Aper}_{n_t}(x) \cap \{0, 1, \ldots, n_t - 1\} = \{l_1, l_2, \ldots, l_s\}$$

with $1 \leqslant l_1 < \cdots < l_s \leqslant n_t$ and $s = ?_t$. Thus, $\mathrm{diam}(E^t_{l_i}) = 1$ and if $l_{i-1} < j < l_i$ ($l_0 = l_s - n_t$ and $l_{s+1} = l_1 + n_t$) then $\mathrm{diam}(E^t_j) = 2^{-\min\{j-l_{i-1}, l_i-j\}}$. Therefore,

$$\delta(E^t) = \sum_{0\leqslant j<n_t} \mathrm{diam}(E^t_j) \geqslant \sum_{i=1}^{s} \mathrm{diam}(E^t_{l_i}) = s$$

and

$$\begin{aligned}\delta(E^t) &= \sum_{0\leqslant j<n_t} \mathrm{diam}(E^t_j) = \sum_{i=1}^{s} \sum_{(l_{i-1}+l_i)/2\leqslant j<(l_i+l_{i+1})/2} \mathrm{diam}(E^t_j)\\ &= \sum_{i=1}^{s}\Bigg(1 + \sum_{1\leqslant j<(l_{i+1}-l_i)/2} 2^{-j} + \sum_{1\leqslant j\leqslant(l_i-l_{i-1})/2} 2^{-j}\Bigg) \leqslant 3s,\end{aligned}$$

which completes the proof. □

As the regularity of x is equivalent to $?_t = o(n_t)$, we have the following conclusion.

COROLLARY A.2. *A Toeplitz sequence is regular if and only if* $\delta(E^t) = o(n_t)$.

B. *Appendix. Sturmian dynamical systems satisfy a PNT*

Let $T : \mathbb{T} \to \mathbb{T}$ ($\mathbb{T} := \mathbb{R}/\mathbb{Z}$) be an irrational rotation on $\mathbb{T}$ by α. For every non-zero $\beta \in \mathbb{T}$, let $\{A_0, A_1\}$ be the partition given by the intervals $A_0 = [0, \beta)$ and $A_1 = [\beta, 1)$. For every $x \in \mathbb{T}$, denote by $\bar{x} \in \{0, 1\}^{\mathbb{Z}}$ the code of x defined by $\bar{x}(k) = i$ if and only if $T^k x \in A_i$. Finally, denote by $X_{\alpha,\beta} \subset \{0, 1\}^{\mathbb{Z}}$ the closure of the set $\{\bar{x} \in \{0, 1\}^{\mathbb{Z}} : x \in \mathbb{T}\}$. Since $X_{\alpha,\beta}$ is an invariant subset for the left shift S on $\{0, 1\}^{\mathbb{Z}}$, we can focus on the topological dynamical system $S : X_{\alpha,\beta} \to X_{\alpha,\beta}$.

THEOREM B.1. *For the topological dynamical system* $S : X_{\alpha,\beta} \to X_{\alpha,\beta}$ *a PNT holds.*

Proof. For every $y = (y(n))_{n\in\mathbb{Z}} \in X_{\alpha,\beta}$ the set $\bigcap_{n\in\mathbb{Z}} \overline{A}_{y(n)} \subset \mathbb{T}$ has exactly one element $\pi(y) \in \mathbb{T}$. Moreover, $\pi : X_{\alpha,\beta} \to \mathbb{T}$ is a continuous map intertwining S and T and there exists a unique S-invariant probability measure μ on $X_{\alpha,\beta}$. The π-image of μ coincides with Lebesgue measure on $\mathbb{T}$.

By Vinogradov's theorem, for any character $f(x) = e^{2\pi i n x}$, $n \in \mathbb{Z}$, we have

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f(T^p x) = \int_{\mathbb{T}} f(x)\, dx \quad \text{for every } x \in \mathbb{T}. \tag{55}$$

Since every continuous function $f : \mathbb{T} \to \mathbb{C}$ is uniformly approximated by trigonometric polynomials, (55) holds also for any continuous f. Moreover, (55) holds for any Riemann

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press

integrable $f : \mathbb{T} \to \mathbb{R}$. Indeed, for every $\varepsilon > 0$ there are two continuous functions $f_-, f_+ : \mathbb{T} \to \mathbb{R}$ such that $f_-(x) \leqslant f(x) \leqslant f_+(x)$ for every $x \in \mathbb{T}$ and $\int_{\mathbb{T}}(f_+(x) - f_-(x))dx < \varepsilon$. It follows that

$$\limsup_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f(T^p x) \leqslant \lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f_+(T^p x) = \int_{\mathbb{T}} f_+(x)\,dx < \int_{\mathbb{T}} f(x)\,dx + \varepsilon$$

and

$$\liminf_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f(T^p x) \geqslant \lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p<N} f_-(T^p x) = \int_{\mathbb{T}} f_-(x)\,dx > \int_{\mathbb{T}} f(x)\,dx - \varepsilon.$$

As $\varepsilon > 0$ can be chosen freely, this gives (55).

Suppose that $f : X_{\alpha,\beta} \to \mathbb{R}$ depends only on finitely many coordinates. More precisely, assume that $f(y) = g(y(-n), \ldots, y(n))$ for some $g : \{0, 1\}^{2n+1} \to \mathbb{R}$. Then there exists $F : \mathbb{T} \to \mathbb{R}$ such that $f = F \circ \pi$ and F is constant on the atoms of the partition $\bigvee_{i=-n}^{n} T^{-i}\{A_0, A_1\}$ (e.g., if $n = 0$ and f is the characteristic function of $\{y \in X_{\alpha,\beta} : y(0) = 0\}$ then F is $\mathbf{1}_{A_0}$). It follows that F is Riemann integrable. Therefore, for every $y \in X_{\alpha,\beta}$, we have

$$\frac{1}{\pi(N)} \sum_{p<N} f(S^p y) = \frac{1}{\pi(N)} \sum_{p<N} F(T^p \pi(y)) \to \int_{\mathbb{T}} F(x)\,dx = \int_{X_{\alpha,\beta}} f\,d\mu.$$

Since every continuous function $f : X_{\alpha,\beta} \to \mathbb{R}$ is uniformly approximated by functions depending on finitely many coordinates,

$$\frac{1}{\pi(N)} \sum_{p<N} f(S^p y) \to \int_{X_{\alpha,\beta}} f\,d\mu \quad \text{for any } y \in X_{\alpha,\beta},$$

holds for every continuous f. □

Acknowledgements. Research supported by Narodowe Centrum Nauki grant 2019/33/B/ST1/00364. A.K. was partially supported by the NSF grant DMS-1956310. We would like to thank the anonymous referee for suggestions which improved the paper.

REFERENCES

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

[2] M. Boyle, D. Fiebig and U. Fiebig. Residual entropy, conditional entropy and subshift covers. *Forum Math.* **14** (2002), 713–757.

[3] J. Bourgain. An approach to pointwise ergodic theorems. *Geometric Aspects of Functional Analysis (1986/87) (Lecture Notes in Mathematics, 1317)*. Springer, Berlin, 1988, pp. 204–223.

[4] J. Bourgain. Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat. *J. Anal. Math.* **119** (2013), 147–163.

[5] J. Bourgain. On the correlation of the Möbius function with rank-one systems. *J. Anal. Math.* **120** (2013), 105–130.

[6] T. Downarowicz. Survey of odometers and Toeplitz flows. *Algebraic and Topological Dynamics (Contemporary Mathematics, 385)*. American Mathematical Society, Providence, RI, 2005, pp. 7–37.

[7] S. Ferenczi, J. Kułaga-Przymus and M. Lemańczyk. Sarnak's conjecture – what's new. *Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, CIRM Jean-Morlet Chair, Fall 2016 (Lecture Notes in Mathematics, 2213)*. Ed. S. Ferenczi, J. Kułaga-Przymus and M. Lemańczyk. Springer International Publishing, Cham, 2018.

[8] S. Ferenczi and C. Mauduit. On Sarnak's conjecture and Veech's question for interval exchanges. *J. Anal. Math.* **134** (2018), 545–573.

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

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

[11] A. Kanigowski, M. Lemańczyk and M. Radziwiłł. Rigidity in dynamics and Möbius disjointness. *Preprint*, 2019, arXiv:1905.13256.

[12] A. Kanigowski, M. Lemańczyk and M. Radziwiłł. Prime number theorem for analytic skew products. To appear in *Fundamenta Math. Preprint*, 2020, arXiv:2004.01125.

[13] A. Kanigowski, M. Lemańczyk and M. Radziwiłł. Semiprime number theorem for smooth Anzai skew products, in preparation.

[14] E. Landau. *Handbuch der Lehre von der Verteilung der Primzahlen* (2 volumes, in German), 2nd edn. Chelsea Publishing Co., New York, 1953. With an appendix by P. T. Bateman.

[15] C. Mauduit and J. Rivat. Prime numbers along Rudin–Shapiro sequences. *J. Eur. Math. Soc.* **17** (2015), 2595–2642.

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

[17] W. Narkiewicz. *Number Theory*. World Scientific Publishing Co., Singapore, 1983. Translated from the Polish by S. Kanemitsu.

[18] R. Pavlov. Some counterexamples in topological dynamics. *Ergod. Th. & Dynam. Sys.* **28** (2008), 1291–1322.

[19] P. Sarnak. Three lectures on the Möbius function, randomness and dynamics. *IAS Lecture Notes*, 2011, http://publications.ias.edu/sarnak/paper/506.

[20] P. Sarnak. Möbius randomness and dynamics six years later at CIRM at 1h 08 minute, 2017, https://library.cirm-math.fr/Record.htm?idlist=1&record=19282918124910001909.

[21] A. Selberg. An elementary proof of the prime-number theorem for arithmetic progressions. *Canad. J. Math.* **2** (1950), 66–78.

[22] I. M. Vinogradov. The method of trigonometrical sums in the theory of numbers. *Trav. Inst. Math. Stekloff* **23** (1947), 109 (in Russian).

[23] M. Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.* **64** (1989), 315–336.

https://doi.org/10.1017/etds.2021.3 Published online by Cambridge University Press