LIMIT THEOREMS FOR NUMBERS OF RETURNS IN ARRAYS UNDER $\phi$-MIXING

YURI KIFER

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL

Abstract. We consider a $\phi$-mixing shift $T$ on a sequence space $\Omega$ and study the number $N_N$ of returns $\{T^{q_N(n)}\omega \in A_n^a\}$ at times $q_N(n)$ to a cylinder $A_n^a$ constructed by a sequence $a \in \Omega$ where $n$ runs either until a fixed integer $N$ or until a time $\tau_N$ of the first return $\{T^{q_N(N)}\omega \in A_m^b\}$ to another cylinder $A_m^b$ constructed by $b \in \Omega$. Here $q_N(n)$ are certain functions of $n$ taking on nonnegative integer values when $n$ runs from 0 to $N$ and the dependence on $N$ is the main generalization here in comparison to [16]. Still, the dependence on $N$ requires certain conditions under which we obtain Poisson distributions limits of $N_N$ when counting is until $N$ as $N \to \infty$ and geometric distributions limits when counting is until $\tau_N$ as $N \to \infty$. The results and the setup are similar to [13] where multiple returns are considered but under the stronger $\psi$-mixing assumption.

1. Introduction

The study of returns to (hits of) shrinking targets by a dynamical system, started in [17], [9] and [6], has already about 30 years history. These works were extended in various directions, in particular, to returns to shrinking geometric balls by uniformly and non-uniformly hyperbolic dynamical systems (see, for instance, [11] and references there), to multiple returns to shrinking cylinder sets under $\psi$-mixing (see, for instance, [14] and others. More recently, motivated by the research on open dynamical systems (see, for instance, [5]) the asymptotic behaviour of numbers of returns to a shrinking target until the first arrival to another shrinking target was investigated in [15] and [16] where the first work dealt with the $\psi$-mixing case while the second one dealt with a $\phi$-mixing situation which allowed applications to a wider class of dynamical systems. Another generalization started in Ch.3 of [10] and continued in [13] and [7], dealt with returns at prescribed times which depended also on the total observation time where additional peculiarities appeared.

In this paper we consider two related types of limit theorems for numbers of returns which are represented by the sums

$$S_N = \sum_{k=1}^{N} \mathbb{1}_{A_n^a \circ T^{q_N(k)}}$$

and

$$\Sigma_N = \sum_{k=1}^{\tau_N} \mathbb{1}_{A_m^b \circ T^{q_N(k)}}$$

Date: December 13, 2021.

2000 Mathematics Subject Classification. Primary: 60F05 Secondary: 37D35, 60J05.

Key words and phrases. Geometric distribution, Poisson distribution, numbers of returns, $\phi$-mixing, shifts.
The probability $P$ by a sequence and $k$ first $k$ertain functions taking on nonnegative integer values when $k$ runs from 1 to $N$ for $N = 1, 2, ...$. In probability such sums where summands themselves depend on the number of summands are usually called (triangular) arrays. We will provide conditions on functions $q_N(k)$ such that as $N \to \infty$ the sum $S_N$ converges in distribution to a Poisson random variable while the sum $\Sigma_N$ converges in distribution to a geometric random variable. It is easy to see that without certain conditions such results do not hold, in general. Indeed, taking $q_N(k) = k(N - k)$ we obtain that the above sums may converge only to a random variable taking on just even values, and so the limits cannot have Poisson or geometric distributions.

Our results remain valid for dynamical systems possessing appropriate symbolic representations such as Axiom A diffeomorphisms (see [3]), expanding transformations and some maps having symbolic representations with an infinite alphabet and a $\psi$-mixing invariant measure such as the Gauss map with its Gauss invariant measure and more general $f$-expansions (see [2]). A direct application of the above results in the symbolic setup yields the corresponding results for arrivals to elements of Markov partitions but employing additional technique (see, for instance, [16]) it is not difficult to extend these results for arrivals to shrinking geometric balls. Since we assume only $\phi$-mixing, rather than $\psi$-mixing, our results remain valid for some classes of nonuniformly expanding maps of the interval such as Gibbs-Markov maps and some others (cf. [16]). In the probability direction we can consider Markov chains with countable state spaces satisfying the Doeblin condition which are known to be exponentially fast $\phi$-mixing (see [4]), and so our results are applicable to the corresponding shifts in the path spaces.

2. Preliminaries and main results

Our setup consists of a finite or countable set $A$ which is not a singleton, the sequence space $\Omega = A^\mathbb{N}$, the $\sigma$-algebra $\mathcal{F}$ on $\Omega$ generated by cylinder sets, the left shift $T : \Omega \to \Omega$, and a $T$-invariant probability measure $P$ on $(\Omega, \mathcal{F})$ which is assumed to be $\phi$-mixing with respect to the $\sigma$-algebras $\mathcal{F}_{mn}$, $n \geq m$ generated by the cylinder sets of the form $\{\omega = (\omega_0, \omega_1, \ldots) \in \Omega : \omega_i = a_i \text{ for } m \leq i \leq n\}$ for some $a_m, a_{m+1}, \ldots, a_n \in A$. Observe also that $\mathcal{F}_{mn} = T^{-m} \mathcal{F}_{0, n-m}$ for $n \geq m$.

Recall, that the $\phi$-dependence (mixing) coefficient between two $\sigma$-algebras $\mathcal{G}$ and $\mathcal{H}$ can be written in the form (see [4]),

$$\phi(\mathcal{G}, \mathcal{H}) = \sup_{\Gamma \in \mathcal{G}, \Delta \in \mathcal{H}} \left\{ \frac{|P(\Gamma \cap \Delta) - P(\Gamma)\cdot P(\Delta)|}{P(\Gamma)}, P(\Gamma) \neq 0 \right\}$$

$$= \frac{1}{2} \sup \{\|E(g|\mathcal{G}) - E(g)\|_{L^\infty} : g \text{ is } \mathcal{H} \text{-measurable and } \|g\|_{L^\infty} \leq 1\}.$$ 

Set also

$$\phi(n) = \sup_{m \geq 0} \phi(\mathcal{F}_{0,m}, \mathcal{F}_{m+n,\infty}).$$

The probability $P$ is called $\phi$-mixing if $\phi(n) \to 0$ as $n \to \infty$.

We will need also the $\alpha$-dependence (mixing) coefficient between two $\sigma$-algebras $\mathcal{G}$ and $\mathcal{H}$ which can be written in the form (see [4]),

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup_{\Gamma \in \mathcal{G}, \Delta \in \mathcal{H}} \left\{ |P(\Gamma \cap \Delta) - P(\Gamma)\cdot P(\Delta)|\right\}$$

$$= \frac{1}{2} \sup \{\|E(g|\mathcal{G}) - E(g)\|_{L^1} : g \text{ is } \mathcal{H} \text{-measurable and } \|g\|_{L^1} \leq 1\}.$$
Set also
\[ \alpha(n) = \sup_{m \geq 0} \alpha(\mathcal{F}_{0,m}, \mathcal{F}_{m+n,\infty}). \]

For each word \( a = (a_0, a_1, \ldots, a_{n-1}) \in \mathcal{A}^n \) we will use the notation \([a] = \{ \omega = (\omega_0, \omega_1, \ldots) : \omega_i = a_i, \ i = 0, 1, \ldots, n-1 \}\) for the corresponding cylinder set. Without loss of generality we assume that the probability of each 1-cylinder set is positive, i.e., \( P([a]) > 0 \) for every \( a \in \mathcal{A} \), and since \( \mathcal{A} \) is not a singleton we have also \( \sup_{a \in \mathcal{A}} P([a]) < 1 \). Write \( \Omega_P \) for the support of \( P \), i.e.,
\[ \Omega_P = \{ \omega \in \Omega : P[\omega_0, \ldots, \omega_n] > 0 \text{ for all } n \geq 0 \}. \]

For \( n \geq 1 \) set \( \mathcal{C}_n = \{ [w] : w \in \mathcal{A}^n \} \). Then \( \mathcal{F}_{0,n} \) consists of \( \emptyset \) and all unions of disjoint elements from \( \mathcal{C}_{n+1} \). If the probability \( P \) is \( \phi \)-mixing then by Lemma 3.1 from [10] there exists \( \nu > 0 \) such that
\[ P(A) \leq e^{-\nu n} \text{ for all } n \geq 1 \text{ and } A \in \mathcal{C}_n. \]

Next, for any \( U \in \mathcal{F}_{0,n-1}, U \neq \emptyset \) define
\[ \pi(U) = \min \{ k \geq 1 : U \cap T^{-k}U \neq \emptyset \} \]
and observe that \( \pi(U) \leq n \). We will be counting the returns to \( U \) at times \( q_N(k) \) considering the sum
\[ S_N^U = \sum_{k=1}^{N} I_U \circ T^{q_N(k)}. \]

Our goal will be to show that if \( U \) is replaced by a sequence of Borel sets \( U_N \subset \Omega \) such that \( N P(U_N) \) converges as \( N \to \infty \) then \( S_N^{U_N} \) converges in distribution to a Poisson random variable and, as an example in Introduction, in order to achieve this some assumptions on functions \( q_N(k) \) are necessary.

2.1. **Assumption.** \( q_N(n) \) is a function taking on nonnegative integer values on integers \( n, N \geq 0 \), defined arbitrarily when \( n > N \) and such that for some constant \( K > 0 \) and all \( N \geq 1 \) the following properties hold true:

(i) For all \( k \) the number of integers \( n, \ 0 \leq n \leq N \) satisfying the equation
\[ q_N(n) = k \]
does not exceed \( K \);

(ii) The number of pairs \( m \neq n \) satisfying \( 0 \leq m, n \leq N \) and solving the equation
\[ q_N(n) - q_N(m) = 0 \]
does not exceed \( K \);

First, note that the example of \( q_N(k) = k(N - k) \) from Introduction does not satisfy Assumption 2.1(ii) since \( q_N(n) = q_N(N - n) \), and so at least \( [N/2] - 1 \) pairs \( n \neq m = N - n \) solve the equation in (ii). Next, observe that if there exist \( n_0, N_0 \geq 1 \) such that for all \( N \geq N_0 \) the function \( q_N(n) \) of \( n \) is strictly increasing when \( n_0 \leq n \leq N \), then the whole Assumption 2.1 is satisfied. Indeed, at most one \( n \geq n_0 \) can solve the equation \( q_N(n) = k \) when \( N \) and \( k \) are fixed, and so the number of solutions in (i) cannot exceed \( n_0 + 1 \). Next, if \( N \geq n, m \geq n_0 \) then \( q_N(n) = q_N(m) \) will hold true only if \( n = m \). If, say, \( m < n_0 \) then for such \( m \) there could be at most one \( n \geq n_0 \) satisfying \( q_N(n) = q_N(m) \). It follows that there exist at most \( n_0^2 + 2n_0 \) pairs \( 0 \leq m, n \leq N \) such that \( q_N(n) = q_N(m) \). In particular, if \( q_N(n) = r(n) + g(N) \) where \( r \) is a nonconstant polynomial in \( n \) and \( g \) is a function
of \( N \), both nonnegative for \( n, N \geq 0 \) and taking on integer values on integers, then \( q_N \) satisfies Assumption 2.1. Indeed, the number of solutions in Assumption 2.1 is bounded by the degree of \( r \) and there exists an integer \( n_0 \geq 1 \) such that the polynomial \( r \) is strictly increasing on \([n_0, \infty)\).

For any two random variables or random vectors \( Y \) and \( Z \) of the same dimension denote by \( \mathcal{L}(Y) \) and \( \mathcal{L}(Z) \) their distribution and by

\[
d_{TV}(\mathcal{L}(Y), \mathcal{L}(Z)) = \sup_G |\mathcal{L}(Y)(G) - \mathcal{L}(Z)(G)|
\]

the total variation distance between \( \mathcal{L}(Y) \) and \( \mathcal{L}(Z) \) where the supremum is taken over all Borel sets. Denote by \( \text{Pois}(\lambda) \) the Poisson distribution with a parameter \( \lambda > 0 \), i.e. \( \text{Pois}(\lambda)(k) = e^{-\lambda} \lambda^k / k! \) for each \( k = 0, 1, 2, \ldots \). Our first result is the following.

2.2. Theorem. Suppose that Assumption 2.1 is satisfied. Then there exists a constant \( C \geq 1 \) such that for any \( n, V \in \mathcal{F}_{0,n-1} \), \( N \) and \( R \),

\[
d_{TV}(\mathcal{L}(S_N^V), \text{Pois}(\lambda_N)) \leq CN(R(P(V))^2 + P(V) \sum_{r=\pi(V)}^{R} (\phi([r/2]+1) + P(T^{n-\tau/2}V) + \phi(R-n))
\]

where \( \lambda_N = N(P(V))^\xi \).

We observe that related estimates under \( \phi \)-mixing were obtained in [1] but it is difficult to obtain definitive convergence results from there.

2.3. Corollary. Suppose that Assumption 2.1 is satisfied and the \( \phi \)-mixing coefficient is summable, i.e.

\[
\sum_{k=1}^{\infty} \phi(k) < \infty.
\]

Let \( V_L \in \mathcal{F}_{0,n_L-1} \), \( L = 1, 2, \ldots \) be a sequence of sets such that \( n_L P(V_L) \to 0 \) and \( \sum_{r=\pi(V_L)}^{n_L-1} P(T^{r}V_L) \to 0 \) as \( L \to \infty \). Let \( N_L \to \infty \) and \( L \to \infty \) be a sequence of integers such that \( 0 < C^{-1} \leq \lambda_L = N_L P(V_L) \leq C < \infty \) for some constant \( C \) and all \( L \geq 1 \). Then

\[
d_{TV}(\mathcal{L}(S_{N_L}^{V_L}), \text{Pois}(\lambda_L)) \to 0 \quad \text{as} \quad L \to \infty
\]

and if \( \lim_{L \to \infty} \lambda_L = \lambda \) then the distribution of \( S_{N_L}^{V_L} \) converges in total variation as \( L \to \infty \) to the Poisson distribution with the parameter \( \lambda \). In particular, if \( V_L = A_{n_L}^L = [\eta_0, \ldots, \eta_{n_L-1}] = \{ \omega \in \Omega : \omega_0 = \eta_0, \ldots, \omega_{n_L-1} = \eta_{n_L-1} \} \) with \( n_L \to \infty \) as \( L \to \infty \) and \( \eta \in \Omega \) is nonperiodic then \( \pi(A_{n_L}^L) \to \infty \) as \( L \to \infty \) and the above statements hold true for such \( V_L \)'s provided the above conditions on \( \lambda_L \) are satisfied.

Next, for any \( V \in \mathcal{F}_{0,n-1} \), \( V \neq \emptyset \) and \( W \in \mathcal{F}_{0,m-1} \), \( W \neq \emptyset \) define

\[
\pi(V, W) = \min\{ k \geq 1 : V \cap T^{-k}W \neq \emptyset \text{ or } W \cap T^{-k}V \neq \emptyset \}.
\]

It is clear that \( \pi(V, W) \leq m \land n \), and so

\[
\kappa_{V,W} = \min\{ \pi(V, W), \pi(V), \pi(W) \} \leq m \land n
\]

where, as usual, for \( n, m \geq 1 \) we denote \( m \lor n = \max\{ m, n \} \) and \( m \land n = \min\{ m, n \} \). Set

\[
\tau_W(\omega) = \min\{ k \geq 1 : T^{q_n(k)} \omega \in W \}
\]
with \( \tau_W(\omega) = \infty \) if the event in braces does not occur and define

\[
\Sigma_N^{V,W} = \sum_{k=1}^{\tau_W} \mathbb{I}_V \circ T^{q_N(k)}.
\]

Denote by \( \text{Geo}(\rho) \) the geometric distribution with a parameter \( \rho \in (0,1) \), i.e. \( \text{Geo}(\rho)(k) = \rho(1-\rho)^k \) for each \( k = 0, 1, 2, \ldots \).

**2.4. Theorem.** Assume that Assumption 2.1 is satisfied. Then there exists a constant \( C > 0 \) such that for any disjoint sets \( V \in \mathcal{F}_{0,n-1} \) and \( W \in \mathcal{F}_{0,m-1} \) with \( P(V), P(W) > 0 \) and all integers \( n, m, N, R \geq 1 \),

\[
d_{TV}(\mathcal{L}(\Sigma_N^{V,W}), \text{Geo}(\rho)) \leq C \left( (1 - P(W))^N + (n \lor m)(P(V) + P(W)) + RN(P(V) + P(W))^2 + N\phi(R - n \lor m) + N(P(V) + P(W)) \sum_{r=\kappa_{V,W}^l}^{n \lor m - 1} (\phi(r) + P(T^n \lor m - r W)) \right)
\]

where \( \rho = \frac{P(W)}{P(V) + P(W)} \).

**2.5. Corollary.** Suppose that Assumption 2.1 holds true and the \( \phi \)-mixing coefficient is summable. Let \( V_L \in \mathcal{F}_{0,n_L-1} \) and \( W_L \in \mathcal{F}_{0,m_L-1}, L = 1, 2, \ldots \) be two sequences of sets such that

\[
(n_L \lor m_L)(P(V_L) + P(W_L)) \to 0, \kappa_{V_L,W_L} \to \infty \quad \text{as} \quad L \to \infty,
\]

\[
\alpha_L = \sum_{r=\kappa_{V_L,W_L}^l}^{n_L \lor m_L - 1} (P(T^{n_L \lor m_L - r} V_L) + P(T^{n_L \lor m_L - r} W_L)) \to 0 \quad \text{as} \quad L \to \infty
\]

and for some constant \( C \) and all \( L \geq 1 \),

\[
0 < C^{-1} \leq \frac{P(V_L)}{P(W_L)} \leq C < \infty.
\]

Let \( N_L, L = 1, 2, \ldots \) be a sequence satisfying

\[
N_L P(W_L) \to \infty \quad \text{and} \quad N_L (M_{N_L} + n_L \lor m_L + \alpha_L) (P(W_L))^2 \to 0 \quad \text{as} \quad L \to \infty
\]

where \( M_N = M_N^{(c)} = \min\{n \geq 1 : \frac{n^\gamma}{\gamma^r(n)} \geq N\} \) for some \( 0 < \varepsilon < 1 \) and \( \gamma(n) = n\phi(n) \). Then

\[
d_{TV}(\mathcal{L}(\Sigma_{N_L}^{V_L,W_L}), \text{Geo}(\rho_L)) \to 0 \quad \text{as} \quad L \to \infty
\]

where \( \rho_L = P(W_L)/(P(W_L) + P(V_L))^{-1} \). In particular, if \( \lim_{L \to \infty} \rho_L = \rho \), then \( \Sigma_{N_L}^{V_L,W_L} \) converges in total variation as \( L \to \infty \) to the geometric distribution with the parameter \( \rho \). Furthermore, let \( V_L = A_{n_L}^\xi = [\xi_0, \ldots, \xi_{n_L-1}] \in \mathcal{C}_{n_L} \) and \( W_L = A_{m_L}^\eta = [\eta_0, \ldots, \eta_{m_L-1}] \in \mathcal{C}_{m_L} \) with \( n_L, m_L \to \infty \) as \( L \to \infty \) and suppose that \( \xi, \eta \) are not periodic and do not shift each other. Then

\[
\kappa_{A_{n_L}^\xi, A_{m_L}^\eta} \to \infty \quad \text{as} \quad L \to \infty
\]

and if also

\[
n_L \land m_L + \kappa_{A_{n_L}^\xi, A_{m_L}^\eta} - n_L \lor m_L \to \infty \quad \text{as} \quad L \to \infty
\]
then (2.8) holds true. In fact, (2.15) is satisfied for $P \times P$-almost all $(\xi, \eta) \in \Omega \times \Omega$ provided
\begin{equation}
2 n_L \wedge m_L - n_L \lor m_L - 3 v \ln(n_L \wedge m_L) \to \infty \text{ as } L \to \infty
\end{equation}
where $v$ is from (2.3).

Observe that when $q_N(n)$ does not depend on $N$ then $\Sigma_N^W$ does not depend on $N$ either and in order to obtain (2.11) relying on (2.6) we have only to pick up some sequence $N_L$ satisfying (2.10) which is always possible provided (2.7)–(2.9) hold true.

3. Poisson distribution limits

We will need the following semi-metrics between positive integers $k, l > 0$,
\[\delta_N(k, l) = |q_N(k) - q_N(l)|.\]

It follows from Assumption (2.1i) that for any integers $n \in \{1, ..., N\}$ and $k \geq 0$,
\begin{equation}
\#\{m : \delta_N(n, m) = k\} \leq 2K.
\end{equation}

For any integers $M, R \geq 1$ and $1 \leq n \leq N$ introduce the sets
\[B_{n,N}^M, R = \{l : 1 \leq l \leq M, \delta_N(l, n) < R\} \quad \text{and} \quad B_{n,N}^R = B_{n,N}^{N,R}.\]

By (3.2), for any $n$,
\begin{equation}
\# B_{n,N}^M, R \leq \min(M, 2KR).
\end{equation}

Let $V \in F_{0,n-1}$ and set $X_{k,N} = X_{k,N}^V = I_V \circ T^{q_N(k)}$. Then $S_N = S_N^V = \sum_{k=1}^N X_{k,N}$. Set $p_{k,N} = P\{X_{k,N} = 1\}$ and $p_{k,N} = P\{X_{k,N} = 1, X_{k,N} = 1\}$. Since $T$ is $P$-preserving $p_{k,N} = E(I_V \circ T^{q_N(k)}) = P(V)$ and $p_{k,N} = P(V \cap T^{-q_N(k)} \circ q_N(l)V)$ provided $q_N(l) \leq q_N(k)$. By Theorem 1 from [2] we obtain
\begin{equation}
d_{TV}(\mathcal{L}(S_N), \text{Pois}(\lambda_N)) \leq b_1 + b_2 + b_3
\end{equation}
where $b_1$, $b_2$ and $b_3$ are defined by
\begin{equation}
b_1 = \sum_{n=1}^N \sum_{l \in B_{n,N}^R} p_{n,N} p_{l,N}, \quad b_2 = \sum_{n=1}^N \sum_{l \in B_{n,N}^R} p_{n,l,N}
\end{equation}
and
\begin{equation}
b_3 = \sum_{n=1}^N s_{n,N} \text{ with } s_{n,N} = E[E(X_{n,N} - p_{n,N}|\sigma\{X_{l,N} : l \in \{1, ..., N\} \setminus B_{n,N}^R\})].
\end{equation}

By (3.2) and (3.4) we conclude that
\begin{equation}
b_1 = \sum_{k=1}^N \sum_{l \in B_{k,N}^R} p_{k,N} p_{l,N} \leq 2KR N(P(V))^2.
\end{equation}

In order to estimate $p_{k,l,N}$ we observe that if $|i - j| < \pi(V)$ then $\langle I_V \circ T^i \rangle(I_V \circ T^j) = 0$. Hence, $p_{k,l,N} = 0$ if $\delta_N(k, l) < \pi(V)$. Now suppose that $\delta_N(k, l) = d$ with $\pi(V) \leq d < n$. Then
\begin{equation}
either q_N(l) \leq q_N(k) - d \text{ or } q_N(l) \geq q_N(k) + d.
\end{equation}
Assume, for instance, that the first inequality in (3.7) holds true and let \( r = q_N(k) - q_N(l) \). Then \( r \geq d \geq \pi(V) \). If \( r \geq n \) then since \( V \in \mathcal{F}_{0,n-1} \), we obtain by the definition of the \( \phi \)-mixing coefficient that
\[
(3.8) \quad p_{k,l,N} = P(V \cap T^{-r}V) \leq (\phi(r - n + 1) + P(V))P(V).
\]

Suppose that \( \pi(V) \leq r < n \) and assume that \( V \cap T^{-r}V \neq \emptyset \). Let \( s \geq n - r \), set \( V_s = T^sV \) and observe that \( T^{-s}V_s \supset V \). Then by the definition of the \( \phi \)-mixing coefficient,
\[
(3.9) \quad p_{k,l,N} = P(V \cap T^{-r}V) \leq P(V \cap T^{-(r+s)}V_s) \leq (\phi(r + s - n + 1) + P(T^sV))P(V) \leq (\phi([r/2] + 1) + P(T^{[r/2]}V))P(V)
\]
taking \( s = n - \lfloor r/2 \rfloor \). If the second inequality in (3.7) holds true then we obtain (3.8) if \( r = q_N(l) - q_N(k) \geq n \), while if \( \pi(V) \leq r < n \) then we arrive at (3.9).

Observe that by Assumption 2.1(i) for any \( N \geq 1 \) and integers \( k \geq 0 \) and \( r \),
\[
(3.10) \quad \# \{ l \geq 0 : q_N(k) - q_N(l) = r \} \leq K.
\]
Now, it follows from (3.7), (3.10) and (3.10) that
\[
(3.11) \quad b_2 = \sum_{k=1}^{N} \sum_{k \neq 1} p_{k,l,N} \leq 4KNP(V) \sum_{r=\pi(V)}^{R} (\phi([r/2] + 1) + P(T^{[r/2]}V)).
\]

Next, we estimate \( s_{k,N} \) and \( b_3 \) defined by (3.5). Since \( \delta_{N}(k,l) \geq R \) for \( l \notin \mathcal{B}_{k,N}^R \) and \( V \in \mathcal{F}_{0,n-1} \), we derive from Lemma 3.3 in [16] and the definition of the \( \alpha \)-mixing coefficient that for \( n < R < N \),
\[
(3.12) \quad s_{k,N} \leq \alpha(\mathcal{F}_{q_N(k),q_N(k)+n},\sigma(\mathcal{F}_{q_N(k)-R+n},\mathcal{F}_{q_N(k)+R-n,\infty})) \leq 3\phi(R - n).
\]
Hence, by (3.5) and (3.12),
\[
(3.13) \quad b_3 = \sum_{k=1}^{N} s_{k,N} \leq 3N\phi(R - n).
\]

Finally, collecting (3.3), (3.6), (3.11) and (3.13) we derive (2.4) completing the proof of Theorem 2.2.

Next, we will derive Corollary 2.3 from the estimate (2.4). The first part of Corollary 2.3 would hold if we find an integer valued sequence \( R_L, L = 1,2,... \) such that \( R_L \to \infty, \frac{R_L}{N_L} \to 0 \) and \( N_L \phi(R_L - n_L) \to 0 \) as \( L \to \infty \). In order to do this we observe first that since \( \phi(k) \) is summable and nonincreasing,
\[
[N/2]\phi(N) \leq \sum_{k=[N/2]}^{N} \phi(k) \leq \sum_{k=[N/2]}^{k} \phi(k) \to 0 \quad \text{as} \quad N \to \infty
\]

which means that \( \gamma(N) = N\phi(N) \to 0 \) as \( N \to \infty \). Observe that if \( \phi(k) = 0 \) for some \( k \geq 1 \) then by monotonicity \( \phi(n) = 0 \) for all \( n \geq k \). In this case there is nothing to prove taking, say, \( R_L = 2n_L \to \infty \) as \( L \to \infty \). Hence, we can and will assume that \( \phi(n) > 0 \) for all \( n \geq 1 \). For some \( 0 < \varepsilon < 1 \) set
\[
M_N = M_N^{(\varepsilon)} = \min \{ n \geq 1 : \frac{n}{\gamma^\varepsilon(n)} \geq N \} \to \infty \quad \text{as} \quad N \to \infty.
\]
Then
\[
\frac{M_N}{\gamma^\varepsilon(M_N)} \geq N \quad \text{and} \quad N\phi(M_N) = \frac{N}{M_N} \gamma(M_N) \leq \gamma^{1-\varepsilon}(M_N) \to 0 \quad \text{as} \quad N \to \infty.
\]
Since \( \frac{M_{N-1}}{N} < N \) then \( \frac{N}{N} > \frac{1}{\tau(M_N-1)} \) as \( N \to \infty \), and so \( \frac{M_N}{N} \to 0 \) as \( N \to \infty \). Hence, taking \( R_L = M_{N_L} + n_L \) we conclude the proof of the first part of Corollary 2.3.

In the second part of Corollary 2.3 we set \( V_L = A^\eta_{n_L} \) where \( \eta \) is a nonperiodic sequence and observe that \( P(A^\eta_{n_L}) \leq e^{-v n_L} \) by (2.3). Hence, the conditions of the first part of Corollary 2.3 would hold true provided

\[
\pi(A^\eta_{n}) \to \infty \quad \text{as} \quad n \to \infty
\]

whenever \( \eta \) is a nonperiodic sequence. To see this note that \( \pi(A^\eta_{n}) \) is, clearly, nondecreasing in \( n \), and so \( \lim_{n \to \infty} \pi(A^\eta_{n}) = r \) exists. If \( r < \infty \) then there exists \( n_0 \geq 1 \) such that \( \pi(A^\eta_{n}) = r \) for all \( n \geq n_0 \) which means that \( \eta \) is periodic with the period \( r \). Hence, \( r = \infty \) since \( \eta \) is not periodic completing the proof of Corollary 2.3. \( \square \)

4. Geometric distribution limits

It will be convenient to set \( V^{(0)} = V \in \mathcal{F}_{0,n-1}, V^{(1)} = W \in \mathcal{F}_{0,m-1} \) and

\[
X^{(\alpha)}_{k,N} = I_{V^{(\alpha)}} \circ T^{q_N(k)} , \quad \alpha = 0, 1
\]

so that

\[
\tau = \tau_{V^{(1)}} = \min\{k \geq 1 : X^{(1)}_{k,N} = 1\} \quad \text{and} \quad \Sigma^{V^{(0)}, V^{(1)}}_{N} = \sum_{k=1}^{\tau} X^{(0)}_{k,N}.
\]

Set also \( S_L = \sum_{k=1}^{L} X^{(0)}_{k,N} \), so that \( \tau = \Sigma^{V^{(0)}, V^{(1)}}_{N} \), and denote \( \tau_N = \min(\tau, N) \). Let \( \{Y^{(\alpha)}_{k,N} : k \geq 1, \alpha = 0, 1\} \) be a sequence of independent Bernoulli random variables such that \( Y^{(\alpha)}_{k,N} \) has the same distribution as \( X^{(\alpha)}_{k,N} \). Since \( P \) is \( T \)-invariant

\[
E(X^{(\alpha)}_{k,N}) = P(X^{(\alpha)}_{k,N} = 1) = E(Y^{(\alpha)}_{k,N}) = P(Y^{(\alpha)}_{k,N} = 1) = P(V^{(\alpha)}).
\]

Set

\[
S_L^{*} = \sum_{k=1}^{L} Y^{(0)}_{k,N} , \quad \tau^{*} = \min\{k \geq 1 : Y^{(1)}_{k,N} = 1\} \quad \text{and} \quad \tau^{*}_N = \min(\tau^{*}, N).
\]

We can and will assume that all above random variables are defined on the same (sufficiently large) probability space. By Lemma 3.1 from [14] the sum \( S^{*}_\tau \) has the geometric distribution with the parameter

\[
(4.1) \quad \theta = \frac{P(V^{(1)})}{P(V^{(1)}) + P(V^{(0)})(1 - P(V^{(1)}))} \geq \rho
\]

where \( \rho = P(V^{(1)})(P(V^{(1)}) + P(V^{(0)}))^{-1} \).

Next, we can write

\[
(4.2) \quad d_{TV}(\mathcal{L}(S_\tau), \mathcal{L}(\tau_N)) \leq A_1 + A_2 + A_3 + A_4
\]

where \( A_1 = d_{TV}(\mathcal{L}(S_\tau), \mathcal{L}(S^{*}_N)) \), \( A_2 = d_{TV}(\mathcal{L}(S^{*}_N), \mathcal{L}(S^{**}_N)) \), \( A_3 = d_{TV}(\mathcal{L}(S^{**}_N), \mathcal{L}(S^{**}_N)) \) and \( A_4 = d_{TV}(\mathcal{L}(\tau_N), \mathcal{L}(\tau_N)) \).

Introduce random vectors \( X^{(\alpha)}_N = \{X^{(\alpha)}_{k,N}, 1 \leq k \leq N\}, \alpha = 0, 1 \), \( X_N = \{X^{(0)}_N, X^{(1)}_N\}, \ Y^{(\alpha)}_N = \{Y^{(\alpha)}_{n,N}, 1 \leq k \leq N\}, \alpha = 0, 1 \) and \( Y_N = \{Y^{(0)}_N, Y^{(1)}_N\} \). Observe that the event \( \{S_\tau \neq S^{*}_N\} \) can occur only if \( \tau > N \). Also, we can write
The estimate of \( A_1 \) is also easy
\[
A_1 \leq P\{ \tau > N \} = P\{ \tau^* > N \} + P\{ X_{n,N}^{(1)} = 0 \text{ for } n = 1, \ldots, N \}
\]
\[
- \frac{1}{2} \sum_{k=0}^{\infty} |x(1-x)k - \rho(1-\rho)k| \leq C \sum_{k=1}^{\infty} ((1-\rho)k - (1-x)k)
\]
Next, we observe that by Theorem 3 in [2],
\[
B_{k,l,N}^r = \{ (l,0), (l,1) : 1 \leq l \leq N, \delta(k,l) < R \}, \quad \rho_{k,l}^{(\alpha,\beta)} = E(X_{k,N}\cdot V_{l,N})
\]
and \( I_N = \{ (k,\alpha) : 1 \leq k \leq N, \alpha = 0,1 \} \) we have
\[
b_1 = \sum_{(k,\alpha)\in I_N} \sum_{(l,\beta)\in B_{k,l,N}^r} \rho_{k,N}^{(\alpha)} \rho_{l,N}^{(\beta)}
\]
\[
b_2 = \sum_{(k,\alpha)\in I_N} \sum_{(l,\beta)\in B_{k,l,N}^r} \rho_{k,l}^{(\alpha,\beta)}
\]
\[
b_3 = \sum_{(k,\alpha)\in I_N} s_{k,N}^{(\alpha)}
\]
where
\[
\begin{align*}
& s_{k,N}^{(\alpha)} = E\{ E(X_{k,N}\cdot V_{l,N}) - \rho_{k,N}^{(\alpha)} | \sigma\{ X_{l,N}^{(\beta)} : (l,\beta) \in I_N \setminus B_{k,l,N}^r \} \}\}
\end{align*}
\]
Since \( \rho_{k,l}^{(\alpha,\beta)} = P(V(\alpha)) \), it follows taking into account [3.1] and [3.2] that
\[
b_1 \leq 6KN \left( (P(V(0))^2 + (P(V(1)))^2 \right).
\]
In order to estimate \( p_{k,l,N}^{(\alpha,\beta)} \) (and, eventually, \( b_2 \)) we observe that
\[
(I_{V(0)} \circ T^i) (I_{V(1)} \circ T^j) = 0 \text{ if } |i - j| < \kappa_{V(0),V(1)}.
\]
Hence, \( \rho_{k,l}^{(\alpha,\beta)} = 0 \) if \( \delta_N(k,l) < \kappa_{V(0),V(1)} \). Now suppose that \( \delta_N(k,l) = d \geq \kappa_{V(0),V(1)} \).
Then we have to deal with two alternatives from [3.7]. Assume, for instance, that the first inequality in [3.7] holds true and let \( r = q_{N}(k) - q_{N}(l) \). Then \( r \geq d \geq \kappa_{V(0),V(1)} \). If \( r \geq n \) then since \( V(0) \in \mathcal{F}_{0,n-1} \) and \( V(1) \in \mathcal{F}_{0,m-1} \) we obtain by the definition of the \( \phi \)-mixing coefficient that
\[
p_{k,l,N}^{(\alpha,\beta)} = P(V(\beta) \cap T^{-r}V(\alpha)) \leq (\phi(r-m+1) + P(V(\alpha))P(V(\beta)),
\]
Suppose that \( \kappa_{V^{(0)},V^{(1)}} \leq r < n \) and assume that \( V^{(\beta)} \cap T^{-r}V^{(\alpha)} \neq \emptyset \). Let \( r \geq n - r \), set \( V_s = T^sV \) and observe that \( T^{-s}V_s \supset V \). Then by the definition of the \( \phi \)-mixing coefficient,

\[
P_{k,l,N}^{\alpha,\beta} = P(V^{(\beta)} \cap T^{-r}V^{(\alpha)}) \leq P(V^{(\beta)} \cap T^{-r+s}V_s^{(\alpha)})
\]

\[
\leq (\phi(r + s - n + 1) + P(T^sV^{(\alpha)}))P(V^{(\beta)})
\]

\[
\leq (\phi([r/2] + 1) + P(T^{n-[r/2]}V^{(\alpha)}))P(V^{(\beta)})
\]

taking \( s = n - \lfloor r/2 \rfloor \). If the second inequality in (4.7) holds true then we obtain (4.12) if \( r = q_N(l) - q_N(k) \geq n \), while if \( \kappa_{V^{(0)},V^{(1)}} \leq r < n \) then we arrive at (4.13), and integers \( k \geq 0 \) and \( r \).

Now, it follows from (3.1), (3.2), (3.10), (4.12) and (4.13) that

\[
b_2 = \sum_{k=1}^{N} \sum_{j \neq b} p_{k,l,N}^{\alpha,\beta} \leq 4K N (P(V^{(0)}) + P(V^{(1)}))
\]

\[
\times \sum_{r=\kappa_{V^{(0)},V^{(1)}}}^{R} \phi([r/2] + 1) + P(T^{n-[r/2]}V^{(0)}) + P(T^{m-[r/2]}V^{(1)}).
\]

Similarly to (3.13) we obtain also that

\[
b_3 = \sum_{1 \leq k \leq N, \alpha = 0,1} s_{k,N}^{(\alpha)} \leq 6N \phi(R - n \lor m).
\]

These provide the estimate of \( A_2 \) by (4.7), (4.8), (4.11), (4.14) and (4.15). Finally, combining (4.2), (4.11), (4.14) and (4.15) we derive (2.6) completing the proof of Theorem 2.4. □

Corollary 2.5 follows from the estimate (2.6) choosing \( R = R_L \) as in Corollary 2.3 and if \( V_L = A_\xi^n L \) and \( W_L = A_\eta^n L \) it remains only to verify the assertion that \( \kappa_{A_\xi^n L, A_\eta^n L} \to \infty \) as \( n, m \to \infty \) provided that \( \xi, \eta \in \Omega_P \) are not periodic and not shifts of each other. Indeed, \( \pi(A_\xi^n L) \), \( \pi(A_\eta^n L) \) and \( \pi(A_\xi^n A_\eta^n L) \) are nondecreasing in \( n \) and \( m \), and so does \( \pi(A_\xi^n L, A_\eta^n L) \). Hence, the limit \( r = \lim_{m,n \to \infty} \kappa_{A_\xi^n L, A_\eta^n L} \) exists. If \( r < \infty \) then, at least, one of the limits \( r_1 = \lim_{m,n \to \infty} \pi(A_\xi^n L) \), \( r_2 = \lim_{m,n \to \infty} \pi(A_\eta^n L) \) or \( r_3 = \lim_{m,n \to \infty} \pi(A_\xi^n A_\eta^n L) \) is finite. If \( r_1 < \infty \) then \( \xi \) is periodic with the period \( r_1 \), if \( r_2 < \infty \) then \( \eta \) is periodic with the period \( r_2 \) and if \( r_3 < \infty \) then either \( T^{r_3} \xi = \eta \) or \( T^{r_3} \eta = \xi \). Finally, it follows from Lemma 3.2 from [10] that (2.12) holds true for \( P \times P \)-almost all \((\xi, \eta)\), completing the proof. □

**References**

[1] M. Abadi and N. Vergne, *Sharp errors for point-wise Poisson approximations in mixing processes*, Nonlinearity 21 (2008), 2871–2885.

[2] R. Arratia, L. Goldstein and L. Gordon, *Two moments suffice for Poisson approximations: the Chen–Stein method*, Ann. Probab. 17 (1989), 9–25.

[3] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Math. 470, Springer–Verlag, Berlin, 1975.

[4] R.C. Bradley, *Introduction to Strong Mixing Conditions*, Kendrick Press, Heber City, 2007.

[5] M. Demers, P.Wright and L.-S. Young, *Entropy, Lyapunov exponents and escape rates in open systems*, Ergod. Th.& Dynam. Sys. 30 (2012), 1270–1301.

[6] M. Denker, *Remarks on weak laws for fractal sets*, in: Fractal Geometry and Stochastics (eds. C.Bandt, S.Graf and M.Zähle), p.p.167–178, Birkhäuser, Basel, 1995.

[7] Ye. Hafouta, *A functional CLT for nonconventional polynomial arrays*, arXiv: 1907.03303.
[8] L. Heinrich, *Mixing properties and central limit theorem for a class of non-identical piecewise monotonic $C^2$-transformations*, Mathematische Nachricht. 181 (1996), 185–214.

[9] M. Hirata, *Poisson law for Axiom A diffeomorphisms*, Ergod. Th. Dynam. Sys. 13 (1993), 533–556.

[10] Ye. Hafouta and Yu. Kifer, *Nonconventional Limit Theorems and Random Dynamics*, World Scientific, Singapore, 2018.

[11] N.T.A. Haydn and K. Wasilewska, *Limiting distribution and error terms for the number of visits to balls in non-uniform hyperbolic dynamical systems*, Discr. Cont. Dyn.Syst. 36 (2016), 2586–2611.

[12] Yu. Kifer, *Nonconventional Poisson limit theorems*, Israel J. Math. 195 (2013), 373–392.

[13] Yu. Kifer, *Limit theorems for numbers of multiple returns in nonconventional arrays*, arXiv: 1910.01439.

[14] Yu. Kifer and A. Rapaport, *Poisson and compound Poisson approximations in conventional and nonconventional setups*, Probab. Th. Relat. Fields 160 (2014), 797–831.

[15] Yu. Kifer and A. Rapaport, *Geometric law for multiple returns until a hazard*, Nonlinearity 32 (2019), 1525–1545.

[16] Yu. Kifer and F. Yang, *Geometric law for numbers of returns until a hazard under $\phi$-mixing*, arXiv: 1812.09927.

[17] B. Pitskel, *Poisson limit law for Markov chains*, Ergod. Th. Dynam. Sys. 11 (1991), 501–513.

Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
E-mail address: kifer@math.huji.ac.il