Ann. Inst. Statist. Math.
Vol. 56, No. 2, 317–349 (2004)

WAITING TIME DISTRIBUTIONS OF RUNS IN HIGHER ORDER MARKOV CHAINS

ANISH SARKAR[1], KANWAR SEN[2] AND ANURADHA[2]

[1]*Statistics & Mathematics Unit, Indian Statistical Institute, 7 S.J.S. Sansanwal Marg, New Delhi 110016, India,* e-mail: anish@isid.ac.in
[2]*Department of Statistics, University of Delhi, Delhi 110007, India,* e-mail: ksen@ndb.vsnl.net.in

(Received December 26, 2001; revised June 26, 2003)

Abstract. We consider a $\{\mathbf{0},\mathbf{1}\}$-valued m-th order stationary Markov chain. We study the occurrences of runs where two $\mathbf{1}$'s are separated by *at most/exactly/at least* k $\mathbf{0}$'s under the overlapping enumeration scheme where $k \geq 0$ and occurrences of scans (at least k_1 successes in a window of length at most k, $1 \leq k_1 \leq k$) under both non-overlapping and overlapping enumeration schemes. We derive the generating function of first two types of runs. Under the conditions, (1) strong tendency towards success and (2) strong tendency towards reversing the state, we establish the convergence of waiting times of the r-th occurrence of runs and scans to Poisson type distributions. We establish the central limit theorem and law of the iterated logarithm for the number of runs and scans up to time n.

Key words and phrases: m-th order Markov chain, generating function, scans, Poisson distribution, central limit theorem, law of the iterated logarithm, α-mixing, strong Markov property.

1. Introduction

Let $X_1, X_2, \dots$ be a sequence of $\{\mathbf{0},\mathbf{1}\}$-valued random variables. One may think of X_n as the outcome of an experiment at the n-th time point and occurrence of $\mathbf{1}/\mathbf{0}$ as the success (S)/failure (F) of the experiment. Throughout this paper we will assume that the sequence of outcomes constitutes a stationary m-th order Markov chain. For any non-negative integer k, we define a run of *type at most k/type exactly k/type at least k* as the occurrence of two successes separated by at most k/ exactly k/ at least k failures respectively. Koutras (1996) investigated the waiting time distribution of the r-th ($r \geq 1$) occurrence of run of type at most k under i.i.d. as well as Markov chain set up and obtained the generating function and Poisson type convergence of the waiting time variable. Koutras (1996) employed a non-overlapping scheme for counting runs, in the sense that a success can only be part of one possible run. However, in this paper, we employ an overlapping counting scheme for all the three types of runs, in the sense that a success may contribute towards counting of two possible runs—one which ends with the occurrence of the success and the next one which is started by it. We associate random variables $T_r^{(M)}$, $T_r^{(E)}$ and $T_r^{(L)}$ with the waiting time for the r-th occurrence of run of type at most k, type exactly k and type at least k respectively. Further, $M_n^{(M)}$, $M_n^{(E)}$ and $M_n^{(L)}$ represent the number of occurrences, up to time n, of runs of type at most k, type exactly k and type at least k respectively.

A natural generalization of the concept of runs of successes of length k as well as runs of type at most k is achieved through the study of scans. For non-negative integers $1 \leq k_1 \leq k$, a scan refers to the occurrence of at least k_1 successes in a window of length at most k. Clearly, when $k = k_1$, the scan is equivalent to the run of successes of length k. For the non-overlapping counting of scans, we can proceed in two ways: in the first scheme we count from scratch every time a scan has been observed, while in the second scheme we start counting afresh only after the window of length k, in which the scan has been observed, is completed. In the overlapping scheme for counting scans, we count the number of windows (not necessarily disjoint) of length k, each of which contains a scan. It is evident that if runs of type at most k are counted using a non-overlapping scheme of counting runs, then they can alternatively be viewed as the occurrence of scans (in a window of length $k+2$ with $k_1 = 2$) under the first non-overlapping scheme of counting of scans. We define $T_{r,k}^{(k_1,j)}$ and $S_{k_1,k}^{(j)}(n)$ as the waiting time for the r-th occurrence of scan and the number of occurrences of scan up to time n respectively under the j-th scheme of counting scans ($j = I$, II and III). Here, I and II refer to the first and second schemes of non-overlapping counting of scans, while III represents the overlapping counting of scans. In this article, we investigate the exact distribution of the waiting time variables $T_r^{(M)}$ and $T_r^{(E)}$. Also, we establish several Poisson type convergence results for the waiting time variables $T_r^{(M)}$, $T_r^{(E)}$, $T_r^{(L)}$ and $T_{r,k}^{(k_1,j)}$ for $j = I$, II and III. Further, we derive the asymptotic results for the enumerating variables $M_n^{(M)}$, $M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(j)}(n)$ for $j = I$, II and III.

To make the definitions more transparent, we quote the same example from Koutras (1996). Consider a string of length 20 of symbols $-\mathbf{1}$ and $\mathbf{0}$, with $k = 1$

$$\mathbf{10010110011101001111}.$$

We can see from the above example that for runs of type at most k, $T_1^{(M)} = 6$, $T_2^{(M)} = 7$, $T_3^{(M)} = 11$, $T_4^{(M)} = 12$, $T_5^{(M)} = 14$, $T_6^{(M)} = 18$, $T_7^{(M)} = 19$ and $T_8^{(M)} = 20$. It is important to note that the success at trial 6 is the end point of the first run while it is the beginning for the second run. Similarly, the successes at trials 11, 12, 18 and 19 also contribute to two runs. For runs of type exactly k, we have $T_1^{(E)} = 6$, $T_2^{(E)} = 14$ while for runs of type at least k, we have $T_1^{(L)} = 4$, $T_2^{(L)} = 6$, $T_3^{(L)} = 10$, $T_4^{(L)} = 14$ and $T_5^{(L)} = 17$. Also we have that $M_{20}^{(M)} = 8$, $M_{20}^{(E)} = 2$ and $M_{20}^{(L)} = 5$. For scans, we consider the window of length 3 with $k_1 = 2$. Then we have, $T_{1,3}^{(2,I)} = 6$, $T_{2,3}^{(2,I)} = 11$, $T_{3,3}^{(2,I)} = 14$, $T_{4,3}^{(2,I)} = 18$, $T_{5,3}^{(2,I)} = 20$, $T_{1,3}^{(2,II)} = 6$, $T_{2,3}^{(2,II)} = 12$, $T_{3,3}^{(2,II)} = 19$ and $T_{1,3}^{(2,III)} = 6$, $T_{2,3}^{(2,III)} = 7$, $T_{3,3}^{(2,III)} = 8$, $T_{4,3}^{(2,III)} = 11$, $T_{5,3}^{(2,III)} = 12$, $T_{6,3}^{(2,III)} = 13$, $T_{7,3}^{(2,III)} = 14$, $T_{8,3}^{(2,III)} = 18$, $T_{9,3}^{(2,III)} = 18$, $T_{10,3}^{(2,III)} = 20$. It is obvious that the waiting time variable $T_{r,3}^{(2,I)}$ can be viewed as the waiting time for the r-th occurrence of the run of type at most 1 when the runs are obtained using the non-overlapping counting scheme; however for the overlapping counting of scans, $T_{r,3}^{(2,III)}$ does not match with the corresponding waiting time variable $T_r^{(M)}$.

The importance of scan statistics and scan waiting time distributions arise from its applications in diverse scientific fields such as reliability, queueing models, molecular biology, statistical quality control, signal detection, computer networking etc. For a detailed discussions on applications of scan statistics and related scan waiting time, we refer the reader to Glaz and Balakrishnan (1999) and Balakrishnan and Koutras (2002).

Feller (1968) initiated the systematic study of the generating functions of distributions of runs of successes of the non-overlapping kind using the renewal theory. In the last 15 years, there has been a major thrust towards finding the exact distribution of different run related statistics. These distributions have been derived under various assumptions on the sequence of underlying random variables such as i.i.d. or independent but not identically distributed or first order Markov dependent as well as higher order Markov dependency. Several authors have contributed to the development of the theory of runs (see Aki (1985, 1992), Philippou (1986), Hirano (1986), Ling (1988), Philippou and Makri (1986), Hirano and Aki (1993), Aki and Hirano (1995), Fu and Koutras (1994), Koutras (1996), Uchida and Aki (1995), Uchida (1998) and references therein). The method of conditional p.g.f. (introduced by Ebneshahrashoob and Sobel (1990)) and the method of Markov embedding techniques (introduced by Fu and Koutras (1994)) have been effectively used to study such distributions. The waiting time distributions for the occurrence of runs of specific type has also been studied extensively by several authors (see Aki *et al.* (1996), Balasubramanian *et al.* (1993) and references therein). Koutras (1996) studied the waiting time distributions for non-overlapping runs of type at most k under the independent as well as Markov dependent set up. Uchida (1998) has also investigated the waiting time problems for patterns under m-th order Markov set up. Several authors have studied the scan waiting time distribution and the scan statistics (see, for example, Koutras and Alexandrou (1995), Koutras (1996), Chadjiconstantinidis *et al.* (2000), Chadjiconstantinidis and Koutras (2001) and references therein). Chen and Glaz (1999) proposed a Poisson approximation and gave an asymptotic expression for tail probabilities of the scan waiting time distribution. Boutsikas and Koutras (2001) has given an approximation scheme for $S_{k_1,k}^{(III)}(n)$. We refer to Balakrishnan and Koutras (2002) for a detailed and thorough account of the development and recent results on scan statistics and scan waiting time distributions. For the theory and applications of the continuous scan statistics, we refer the reader to Glaz *et al.* (2001). In this article, we study the scan statistics and the scan waiting time distribution under a m-th order Markov chain set up.

In the next section, we introduce the necessary mathematical definitions and notations. In Section 3, we derive the generating functions of the waiting time distributions of $T_r^{(M)}$ and $T_r^{(E)}$. For both the waiting time variables, we develop a system of linear equations using the method of conditional p.g.f.s. In Section 4, various asymptotic results on the convergence of waiting time distributions have been derived under two broad set-ups. In the first set-up, we assume that the system has a strong tendency towards success while in the second set-up we assume that the system has a strong tendency towards reversing its states, i.e., from failure it would like to switch to success and vice versa. Under the first set-up, we show that $T_r^{(M)}$ converges to a Poisson type distribution (Theorem 4.1) and the waiting time variables for scans $T_{r,k}^{(k_1,I)}$, $T_{r,k}^{(k_1,II)}$ and $T_{r,k}^{(k_1,III)}$ exhibit similar Poisson type convergence (Theorem 4.2). Under the second set-up, we show that the waiting time variables $T_r^{(M)}$, $T_r^{(E)}$ and $T_r^{(L)}$ converge to sum of independent Poisson type random variables (Theorems 4.3, 4.4 and 4.5). In this case, similar results have been established for $T_{r,k}^{(k_1,III)}$ corresponding to the cases: (a) k even and $2k_1 \leq k$ (Theorem 4.6) and (b) k odd and $2k_1 - 1 \leq k$ (Theorem 4.7). In the final section, we derive the central limit theorems for $M_n^{(M)}$, $M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(j)}(n)$ for $j = I$, II and III. Further, law of the iterated logarithm has been obtained for $M_n^{(M)}$,

$M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(III)}(n)$. To obtain the above results, we define a sequence of new random variables, in terms of the original random variables, so that the sequence of new random variables is a stationary Markov chain, assuming values in a finite set. We derive a meta central limit theorem and law of the iterated logarithm, involving functions of the newly defined random variables (Theorem 5.1). This theorem yields the results for $M_n^{(M)}$, $M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(III)}(n)$ as special cases (see Theorems 5.2, 5.3, 5.4 and 5.5). For the two non-overlapping scheme of counting scans, we define a sequence of stopping times so that conditioned on these stopping times, the number of occurrences of scans become independent. Hence $S_{k_1,k}^{(j)}(n)$ can be approximated by a random sum of i.i.d. random variables, for both $j = I$ and II. This result has been used to establish the asymptotic normality of $S_{k_1,k}^{(j)}(n)$, $j = I$ and II (see Lemma 5.1 and Theorem 5.6).

2. Definitions & notations

Let $X_{-m+1}, X_{-m+2}, \ldots, X_0, X_1, X_2, \ldots,$ be a sequence of stationary m-th order $\{0,1\}$-valued Markov chain. It is assumed that the states of $X_{-m+1}, X_{-m+2}, \ldots, X_0$ are known, i.e., we are given the initial condition $\{X_0 = x_0, X_{-1} = x_1, \ldots, X_{-m+1} = x_{m-1}\}$. For any $i \geq 1$, define $N_i = \{0, 1, \ldots, 2^i - 1\}$. The initial condition can be represented by $x = \sum_{j=0}^{m-1} 2^j x_j$. Then, $x \in N_m = \{0, 1, \ldots, 2^m - 1\}$. Clearly, for any $x \in N_m$ we will have a unique initial condition which is given by the binary representation of x (written in the reverse order). We define, for any $n \geq 0$,

$$p_x = P(X_{n+1} = 1 \mid X_n = x_0, X_{n-1} = x_1, \ldots, X_{n-m+1} = x_{m-1}). \tag{2.1}$$

Consequently, $q_x = P(X_{n+1} = 0 \mid X_n = x_0, X_{n-1} = x_1, \ldots, X_{n-m+1} = x_{m-1}) = 1 - p_x$. We denote the probability measure governing the system with initial condition $x \in N_m$ by P_x. Further, we define two functions, $f_0, f_1 : N_m \to N_m$ as

$$f_0(x) = 2x (\text{mod } 2^m) \qquad \text{and} \qquad f_1(x) = (2x+1)(\text{mod } 2^m).$$

Note that $f_0(x)/f_1(x)$ stands for the initial condition derived from x if we obtain a failure/success in the next trial.

The probability generating functions of $T_r^{(M)}$, $T_r^{(E)}$ and $T_r^{(L)}$ are denoted by $\phi_r^{(M)}(x,s)$, $\phi_r^{(E)}(x,s)$ and $\phi_r^{(L)}(x,s)$ respectively. In other words, we have $\phi_r^{(j)}(x,s) = \sum_{n=0}^{\infty} P_x(T_r^{(j)} = n) s^n$ for $j = M$, E, or L. Further, let us define, $\Phi^{(M)}(x,z)$, $\Phi^{(E)}(x,z)$ and $\Phi^{(L)}(x,z)$ as the generating functions of $\{\phi_r^{(M)}(x,s) : r \geq 1\}$, $\{\phi_r^{(E)}(x,s) : r \geq 1\}$ and $\{\phi_r^{(L)}(x,s) : r \geq 1\}$ respectively, i.e., $\Phi^{(j)}(x,z) = \sum_{r=1}^{\infty} \phi_r^{(j)}(x,s) z^r$ for $j = M$, E, or L.

The two non-overlapping counting schemes of scan can be interpreted as follows: the first scheme counts the number of disjoint scans while the second scheme counts the number of disjoint windows of length k each of which contains a scan. To facilitate our study, while considering the second non-overlapping scheme for counting scans, we set the end point of the window of length k, containing the scan, as the end point of the scan itself. More precisely, we define the counting random variables for scans as follows:

$$R_{k_1,k}^{(I)}(n) = 1_{\{X_n + \sum_{j=n-k+1}^{n-1} X_j \prod_{t=j}^{n-1} (1 - R_{k_1,k}^{(I)}(t)) \geq k_1\}}$$

$$R_{k_1,k}^{(II)}(n) = 1_{\{\sum_{j=n-k+1}^{n} X_j \geq k_1\}} \prod_{j=n-k+1}^{n-1} (1 - R_{k_1,k}^{(II)}(j))$$

$$R_{k_1,k}^{(III)}(n) = 1_{\{\sum_{j=n-k+1}^{n} X_j \geq k_1\}}$$

with the convention that $R_{k_1,k}^{(I)}(n) = 0$ for $n < k_1$, $R_{k_1,k}^{(II)}(n) = R_{k_1,k}^{(III)}(n) = 0$ for $n < k$. It should be noted that $R_{k_1,k}^{(I)}(n) = 1$ if and only if n is the end point of a scan when counted under the first non-overlapping scheme. Similarly, it is evident that $R_{k_1,k}^{(II)}(n)$ and $R_{k_1,k}^{(III)}(n)$ counts the scans under the second non-overlapping scheme and the overlapping scheme respectively. Thus, we may define the scan waiting time variables, for $r \geq 1$, as

$$T_{r,k}^{(k_1,j)} = \inf\left\{n : \sum_{i=0}^{n} R_{k_1,k}^{(j)}(i) = r\right\}$$

where j is I, II or III. Finally, the scan statistics (the number of scans up to time n) $S_{k_1,k}^{(j)}(n)$, under the j-th scheme of counting, is defined by $S_{k_1,k}^{(j)}(n) = \sum_{i=1}^{n} R_{k_1,k}^{(j)}(i)$ where j is either I, II or III.

3. Generating functions

In this section, we obtain the generating functions of the probability distribution of the waiting time variables $T_r^{(M)}$ and $T_r^{(E)}$. We define a new event in the following way: suppose that we are given an initial condition x which is odd; this represents the event that $X_0 = 1$. However, when we start looking for a run (of any type), we ignore the value of X_0. We define the new event which will consider this case and take into account the value of X_0. Formally speaking, given x odd, we say that an *associated run of type at most* k occurs at time n if we observe $n-1$ failures followed by a success for $n \leq k+1$ and for $n > k+1$, if we see more than k failures at the beginning and then observe a run of type at most k at time n. It implies that when we take X_0 into account, we get a run of type at most k which ends at time n. For runs of type exactly k, we define an *associated run* in the similar way: given an initial condition that x is odd, we say that an associated run of type exactly k occurs at time n if a run of type exactly k occurs at time n when we start looking from time 0.

Define $N'_{m-1} = \{y : y = 2x+1, x \in N_{m-1}\}$ as the set of all odd numbers in N_m. For $x \in N'_{m-1}$, let $S_r^{(M)}$ and $S_r^{(E)}$ be the waiting times of the r-th occurrence of the associated run of type at most k and type exactly k respectively. Define, for $r \geq 1$, the probability generating function of the waiting times $S_r^{(M)}$ and $S_r^{(E)}$ by $\psi_r^{(M)}(x,s)$ and $\psi_r^{(E)}(x,s)$ respectively. Further, define $\Psi^{(M)}(x,z)$ and $\Psi^{(E)}(x,z)$ as the generating function of the sequences $\{\psi_r^{(M)}(x,s) : r \geq 1\}$ and $\{\psi_r^{(E)}(x,s) : r \geq 1\}$ respectively.

Define the sequence of events,

$$A_i = \{X_j = 0 \text{ for } 1 \leq j \leq i \text{ and } X_{i+1} = 1\}$$

for $i = 0, 1, \ldots, k$ and

$$A_{k+1} = \{X_j = 0 \text{ for } 1 \leq j \leq k+1\}.$$

Clearly, the events $\{A_i : i = 0, 1, \ldots, k+1\}$ form a partition of the sample space.

Now, conditioning on the outcome observed at the first time point, the following equation is easy to derive: for any $r \geq 1$ and $n \geq 1$,

$$P_x(T_r^{(M)} = n) = q_x P_{f_0(x)}(T_r^{(M)} = n-1) + p_x P_{f_1(x)}(S_r^{(M)} = n-1). \tag{3.1}$$

This equation can now be used to derive the following relation between their respective probability generating functions:

$$\phi_r^{(M)}(x,s) = q_x s \phi_r^{(M)}(f_0(x),s) + p_x s \psi_r^{(M)}(f_1(x),s). \tag{3.2}$$

This, in turn, gives us a linear relation between $\Phi^{(M)}(x,z)$ and $\Psi^{(M)}(x,z)$.

$$\Phi^{(M)}(x,z) = q_x s \Phi^{(M)}(f_0(x),z) + p_x s \Psi^{(M)}(f_1(x),z). \tag{3.3}$$

We define for $x, y \in N_m$,

$$a_{11}^{(M)}(x,y) = \begin{cases} 1 & \text{if } \ y = x \\ -q_x s & \text{if } \ y = f_0(x) \\ 0 & \text{otherwise} \end{cases}$$

and for $x \in N_m$ and $y \in N'_{m-1}$,

$$a_{12}^{(M)}(x,y) = \begin{cases} -p_x s & \text{if } \ y = f_1(x) \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$\boldsymbol{A}_{11}^{(M)} = (a_{11}^{(M)}(x,y))_{x,y \in N_m} \quad \text{and} \quad \boldsymbol{A}_{12}^{(M)} = (a_{12}^{(M)}(x,y))_{x \in N_m,\, y \in N'_{m-1}}$$

be the associated matrices. Hence, we can express the set of above equations as

$$\boldsymbol{A}_{11}^{(M)} \boldsymbol{\Phi}^{(M)} + \boldsymbol{A}_{12}^{(M)} \boldsymbol{\Psi}^{(M)} = \boldsymbol{0} \tag{3.4}$$

where $\boldsymbol{\Phi}^{(M)} = (\Phi^{(M)}(x,z))_{x \in N_m}$ and $\boldsymbol{\Psi}^{(M)} = (\Psi^{(M)}(x,z))_{x \in N'_{m-1}}$.

A similar argument holds also for runs of type exactly k. As earlier, conditioning on the result of the first trial, we obtain exactly the same relation between $P(T_r^{(E)} = n)$ and $P(S_r^{(E)} = n)$ which, in turn, yields the equations:

$$\boldsymbol{A}_{11}^{(E)} \boldsymbol{\Phi}^{(E)} + \boldsymbol{A}_{12}^{(E)} \boldsymbol{\Psi}^{(E)} = \boldsymbol{0} \tag{3.5}$$

where $\boldsymbol{A}_{11}^{(E)} = \boldsymbol{A}_{11}^{(M)}$ and $\boldsymbol{A}_{12}^{(E)} = \boldsymbol{A}_{12}^{(M)}$.

3.1 *Runs of type at most k*

Now, we can easily derive the following relation between $P_x(T_r^{(M)} = n)$ and $P_x(S_r^{(M)} = n)$. For $r = 1$, we have

$$P_x(S_1^{(M)} = n) = \begin{cases} P_x(A_{n-1}) & \text{for } \ n = 1, 2, \ldots, k+1 \\ P_x(A_{k+1}) P_{f_0^{k+1}(x)}(T_1^{(M)} = n-k-1) & \text{for } \ n > k+1. \end{cases}$$

For $r \geq 2$, we can find the following relation by conditioning on the partition $\{A_i : i = 0, 1, \ldots, k+1\}$. When $m \geq k+1$, we have:

$$P_x(S_r^{(M)} = n) = \sum_{i=0}^{k+1} P_x(S_r^{(M)} = n, A_i)$$

$$
\begin{aligned}
&= \sum_{i=0}^{k+1} P_x(S_r^{(M)} = n \mid A_i) P_x(A_i) \\
&= \sum_{i=0}^{k} P_{f_1(f_0^i(x))}(S_{r-1}^{(M)} = n - i - 1) P_x(A_i) \\
&\quad + P_x(A_{k+1}) P_{f_0^{k+1}(x)}(T_r^{(M)} = n - k - 1)
\end{aligned}
$$

where $f_0^0(x) = x$ and $f_0^{i+1}(x) = f_0(f_0^i(x))$.

When $m < k+1$, we have,

$$
\begin{aligned}
&P_x(S_r^{(M)} = n) \\
&\quad = \sum_{i=0}^{k+1} P_x(S_r^{(M)} = n \mid A_i) P_x(A_i) \\
&\quad = \sum_{i=0}^{m-2} P_{f_1(f_0^i(x))}(S_{r-1}^{(M)} = n - i - 1) P_x(A_i) \\
&\qquad + \sum_{i=m-1}^{k} P_1(S_{r-1}^{(M)} = n - i - 1) P_x(A_i) + P_x(A_{k+1}) P_0(T_r^{(M)} = n - k - 1).
\end{aligned}
$$

From the definition of functions f_0 and f_1, it is clear that they satisfy the following relations:

$$
(3.6) \qquad f_0^{m+j} = 0 \quad \text{and} \quad f_1(f_0^{m-1+j}(x)) = 1 \quad \text{for} \quad j \geq 0, x \in N_m.
$$

Using these relations, both the cases, namely $m \geq k+1$ and $m < k+1$, can be combined to yield,

$$
\psi_r^{(M)}(x, s) = \begin{cases}
\displaystyle\sum_{i=0}^{k} P_x(A_i) s^{i+1} & \\
\qquad + P_x(A_{k+1}) s^{k+1} \phi_1^{(M)}(f_0^{k+1}(x), s) & \text{for} \quad r = 1 \\
\displaystyle\sum_{i=0}^{k} P_x(A_i) s^{i+1} \psi_{r-1}^{(M)}(f_1(f_0^i(x)), s) & \\
\qquad + P_x(A_{k+1}) s^{k+1} \phi_r^{(M)}(f_0^{k+1}(x), s) & \text{for} \quad r > 1.
\end{cases}
$$

These will give us another set of equations

$$
\begin{aligned}
(3.7) \qquad \Psi^{(M)}(x, z) &= \sum_{i=0}^{k} P_x(A_i) s^{i+1} z \Psi^{(M)}(f_1(f_0^i(x)), z) \\
&\quad + P_x(A_{k+1}) s^{k+1} \Phi^{(M)}(f_0^{k+1}(x), z) \\
&\quad + \sum_{i=0}^{k} P_x(A_i) s^{i+1} z.
\end{aligned}
$$

The probabilities $P_x(A_i)$ can easily be computed. Indeed,

$$P_x(A_i) = \begin{cases} p_{f_0^i(x)} \prod_{j=0}^{i-1} q_{f_0^j(x)} & \text{for} \quad 0 \le i \le k \\ \prod_{j=0}^{k} q_{f_0^j(x)} & \text{for} \quad i = k+1. \end{cases} \tag{3.8}$$

We define for $x, y \in N'_{m-1}$,

$$a_{22}^{(M)}(x,y) = \begin{cases} 1 & \text{if} \quad y = x \\ -P_x(A_i)s^{i+1}z & \text{if} \quad y = f_1(f_0^i(x)) \quad \text{for some } i = 0, 1, \dots, k \\ 0 & \text{otherwise} \end{cases}$$

and for $x \in N'_{m-1}$ and $y \in N_m$,

$$a_{21}^{(M)}(x,y) = \begin{cases} -P_x(A_{k+1})s^{k+1} & \text{if} \quad y = f_0^{k+1}(x) \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$\boldsymbol{A}_{21}^{(M)} = (a_{21}^{(M)}(x,y))_{x \in N'_{m-1}, y \in N_m} \quad \text{and} \quad \boldsymbol{A}_{22}^{(M)} = (a_{22}^{(M)}(x,y))_{x,y \in N'_{m-1}}$$

and $\boldsymbol{b}^{(M)} = \sum_{i=0}^{k} P_x(A_i)s^{i+1}z$. Hence, we can write the set of equations in (3.7) as

$$\boldsymbol{A}_{21}^{(M)} \boldsymbol{\Phi}^{(M)} + \boldsymbol{A}_{22}^{(M)} \boldsymbol{\Psi}^{(M)} = \boldsymbol{b}^{(M)} \tag{3.9}$$

where $\boldsymbol{\Phi}^{(M)}$ and $\boldsymbol{\Psi}^{(M)}$ are as defined earlier. Therefore, we have

THEOREM 3.1. *The generating function of* $\{\Phi_r^{(M)}(x,s) : r \ge 1\}$ *is given by*

$$\boldsymbol{\Phi}^{(M)} = (\boldsymbol{A}_{12}^{(M)}(\boldsymbol{A}_{22}^{(M)})^{-1}\boldsymbol{A}_{21}^{(M)} - \boldsymbol{A}_{11}^{(M)})^{-1}\boldsymbol{A}_{12}^{(M)}(\boldsymbol{A}_{22}^{(M)})^{-1}\boldsymbol{b}^{(M)}. \tag{3.10}$$

3.2 *Runs of type exactly k*

We derive a relation between $P_x(T_r^{(E)} = n)$ and $P_x(S_r^{(E)} = n)$ by using a similar technique. For $r = 1$, we have

$$P_x(S_1^{(E)} = n) = \begin{cases} P_x(A_n) & \text{for} \quad n = k \\ \sum_{i=0}^{k-1} P_x(A_i) P_{f_1(f_0^i(x))}(S_1^{(E)} = n - i - 1) & \\ \qquad + P_x(A_{k+1}) P_{f_0^{k+1}(x)}(T_1^{(E)} = n - k - 1) & \text{otherwise.} \end{cases}$$

For $r \ge 2$, using the relations in (3.6), we can combine the two cases, namely, $k+1 > m$ and $k+1 \le m$, to obtain the following equations:

$$P_x(S_r^{(E)} = n) = \sum_{i=0}^{k+1} P_x(S_r^{(E)} = n \mid A_i) P_x(A_i)$$

$$\begin{aligned} &= \sum_{i=0}^{k-1} P_{f_1(f_0^i(x))}(S_r^{(E)} = n-i-1)P_x(A_i) \\ &\quad + P_x(A_k)P_{f_1(f_0^k(x))}(S_{r-1}^{(E)} = n-k-1) \\ &\quad + P_x(A_{k+1})P_{f_0^{k+1}(x)}(T_r^{(E)} = n-k-1). \end{aligned}$$

Therefore, we have

$$\psi_r^{(E)}(x,s) = \begin{cases} \displaystyle\sum_{i=0}^{k-1} P_x(A_i)s^{i+1}\psi_r^{(E)}(f_1(f_0^i(x)),s) & \\ \qquad + P_x(A_k)s^{k+1} + P_x(A_{k+1})s^{k+1}\phi_r^{(E)}(f_0^{k+1}(x),s) & \text{for } \ r=1 \\ \displaystyle\sum_{i=0}^{k-1} P_x(A_i)s^{i+1}\psi_r^{(E)}(f_1(f_0^i(x)),s) & \\ \qquad + P_x(A_k)s^{k+1}\psi_{r-1}^{(E)}(f_1(f_0^k(x)),s) & \text{for } \ r>1. \\ \qquad + P_x(A_{k+1})s^{k+1}\phi_r^{(E)}(f_0^{k+1}(x),s) & \end{cases}$$

As before, this gives us the following set of equations,

$$\begin{aligned} \Psi^{(E)}(x,z) &= \sum_{i=0}^{k-1} P_x(A_i)s^{i+1}\Psi^{(E)}(f_1(f_0^i(x)),z) + P_x(A_k)s^{k+1}z \\ &\quad + P_x(A_k)s^{k+1}z\Psi^{(E)}(f_1(f_0^k(x)),z) + P_x(A_{k+1})s^{k+1}\Phi^{(E)}(f_0^{k+1}(x),z). \end{aligned}$$

The probabilities $P_x(A_i)$ are specified in equation (3.8). Define, for $x, y \in N'_{m-1}$,

$$a_{22}^{(E)}(x,y) = \begin{cases} 1 & \text{if } \ y = x \\ -P_x(A_i)s^{i+1} & \text{if } \ y = f_1(f_0^i(x)) \ \text{ for some } i = 0,1,\ldots,k-1 \\ -P_x(A_k)s^{k+1}z & \text{if } \ y = f_1(f_0^k(x)) \\ 0 & \text{otherwise} \end{cases}$$

and for $x \in N'_{m-1}$ and $y \in N_m$,

$$a_{21}^{(E)}(x,y) = \begin{cases} -P_x(A_{k+1})s^{k+1} & \text{if } \ y = f_0^{k+1}(x) \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$\boldsymbol{A}_{21}^{(E)} = (a_{21}^{(E)}(x,y))_{x\in N'_{m-1},y\in N_m} \qquad \text{and} \qquad \boldsymbol{A}_{22}^{(E)} = (a_{22}^{(E)}(x,y))_{x,y\in N'_{m-1}}$$

and $\boldsymbol{b}^{(E)} = P_x(A_k)s^{k+1}z$. Hence, we can write the set of above equations as

$$\boldsymbol{A}_{21}^{(E)}\boldsymbol{\Phi}^{(E)} + \boldsymbol{A}_{22}^{(E)}\boldsymbol{\Psi}^{(E)} = \boldsymbol{b}^{(E)} \tag{3.11}$$

where $\boldsymbol{\Phi}^{(E)}$ and $\boldsymbol{\Psi}^{(E)}$ are as defined earlier. Therefore, as before, we have

THEOREM 3.2. *The generating function of* $\{\Phi_r^{(E)}(x,s) : r \geq 1\}$ *is given by*

$$\boldsymbol{\Phi}^{(E)} = (\boldsymbol{A}_{12}^{(E)}(\boldsymbol{A}_{22}^{(E)})^{-1}\boldsymbol{A}_{21}^{(E)} - \boldsymbol{A}_{11}^{(E)})^{-1}\boldsymbol{A}_{12}^{(E)}(\boldsymbol{A}_{22}^{(E)})^{-1}\boldsymbol{b}^{(E)}. \tag{3.12}$$

4. Limit distributions

In this section, we obtain several limit laws of the distribution of the waiting times. We will take a direct evaluation route for the results. For this, we require the following lemmas about weak convergence of discrete random variables. We could not find the exact result that we require, in any of the standard references (Feller (1968), Billingsley (1986), Chung (1974)) and hence we incorporate it for the completeness of the paper.

LEMMA 4.1. *Let $\{\xi_r : r \geq 1\}$ be a sequence of random variables taking values on $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ such that*

$$\liminf_{r\to\infty} P(\xi_r = t) \geq p_t \qquad \text{for all} \quad t \in \mathbb{Z}$$

where $\sum_{t=-\infty}^{\infty} p_t = 1$, then $\xi_r \Rightarrow \xi$ where $P(\xi = t) = p_t$ for all $t \in \mathbb{Z}$.

PROOF. First we claim that the sequence of random variables $\{\xi_r : r \geq 1\}$ is tight. Indeed, for any $\epsilon > 0$, we first choose K so large that $\sum_{t=-K}^{K} p_t > 1 - \epsilon/2$. Now, for every $t, |t| \leq K$, we choose N_t so large that whenever $r \geq N_t$, $P(\xi_r = t) > p_t - \epsilon/(8K)$. Now, setting $N = \max\{N_t : |t| \leq K\}$, we have for $t \geq N$,

$$P(|\xi_r| \leq K) = \sum_{t=-K}^{K} P(\xi_r = t) > \sum_{t=-K}^{K} p_t - (2K+1)\epsilon/(8K) > 1 - \epsilon.$$

Now, by Corollary of Theorem 25.10 of Billingsley (1986), it is enough to show that if for any sub-sequence $\{\xi_{r_i} : i \geq 1\}$ such that $\xi_{r_i} \Rightarrow \xi'$, then $\xi' \stackrel{d}{=} \xi$. We have, $P(\xi' = t) = \lim_{i\to\infty} P(\xi_{r_i} = t) \geq \liminf_{r\to\infty} P(\xi_r = t) = P(\xi = t)$. If for some $t_0 \in \mathbb{Z}$, $P(\xi' = t_0) > P(\xi = t_0)$, then we have

$$1 = \sum_{t=-\infty}^{\infty} P(\xi' = t) = P(\xi' = t_0) + \sum_{t \neq t_0} P(\xi' = t) > P(\xi = t_0) + \sum_{t \neq t_0} P(\xi = t) = 1.$$

This is a contradiction, proving that $\xi' \stackrel{d}{=} \xi$. □

We also need the following lemma for proving several of our results.

LEMMA 4.2. *Let $\{\xi_r : r \geq 1\}$ be a sequence of random variables taking values on $\mathbb{Z}$ such that*

$$\liminf_{r\to\infty} P(\xi_r = t, A_j^{(r)}) \geq p_j^{(1)} p_{t-j}^{(2)} \qquad \text{for all} \qquad t, j \in \mathbb{Z}$$

where $\sum_{t=-\infty}^{\infty} p_t^{(1)} = 1 = \sum_{t=-\infty}^{\infty} p_t^{(2)}$ and $\{A_j^{(r)} : j \in \mathbb{Z}\}$ are disjoint events for each r. Then

$$\xi_r \Rightarrow \xi^{(1)} + \xi^{(2)}$$

where $P(\xi^{(1)} = t) = p_t^{(1)}$ and $P(\xi^{(2)} = t) = p_t^{(2)}$ for all $t \in \mathbb{Z}$ and $\xi^{(1)}$ & $\xi^{(2)}$ are independent.

PROOF. Clearly, for any $t \in \mathbb{Z}$, $P(\xi^{(1)} + \xi^{(2)} = t) = \sum_{j=-\infty}^{\infty} p_j^{(1)} p_{t-j}^{(2)}$. Thus, it is enough, by Lemma 4.1, to show that

$$\liminf_{r\to\infty} P(\xi_r = t) \geq \sum_{j=-\infty}^{\infty} p_j^{(1)} p_{t-j}^{(2)}.$$

To prove this, we consider any $\epsilon > 0$ and choose J so large that $\sum_{j=-J}^{J} p_j^{(1)} p_{t-j}^{(2)} > \sum_{j=-\infty}^{\infty} p_j^{(1)} p_{t-j}^{(2)} - \epsilon$. Now, we have,

$$\begin{aligned}
\liminf_{r\to\infty} P(\xi_r = t) &\geq \liminf_{r\to\infty} \sum_{j=-\infty}^{\infty} P(\xi_r = t, A_j^{(r)}) \\
&\geq \liminf_{r\to\infty} \sum_{j=-J}^{J} P(\xi_r = t, A_j^{(r)}) \\
&\geq \sum_{j=-J}^{J} \liminf_{r\to\infty} P(\xi_r = t, A_j^{(r)}) \\
&\geq \sum_{j=-J}^{J} p_j^{(1)} p_{t-j}^{(2)} \\
&\geq \sum_{j=-\infty}^{\infty} p_j^{(1)} p_{t-j}^{(2)} - \epsilon.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we get the desired result. □

We say that a random variable Z_k follows a Poisson distribution of multiplicity k with parameter λ ($k \in \mathbb{Z}$ and $\lambda > 0$) if

$$P(Z_k = kt) = \frac{\exp(-\lambda)\lambda^t}{t!} \quad \text{for} \quad t = 0, 1, \ldots.$$

In the sequel, we will denote this by $Z_k \sim \mathrm{Poi}(k, \lambda)$. Note that, when $k = 1$ it is the usual Poisson distribution.

4.1 *Strong tendency towards success*

We first consider the case of overlapping runs of type at most k. The assumption that we make on the probabilities is that system has a strong tendency towards success. We formalize this by stating that for $u \in N_m$, p_u (as a function of r) converges to 1 in such a way that

$$r(1 - p_u) \to \mu_u \quad \text{where} \quad \mu_u > 0 \quad \text{is a positive constant.} \tag{4.1}$$

We require the probability of the following event to establish our results. Fix $w \geq m$ where m is the order of the Markov chain. For $t \geq 0$ and $l \geq w(t+1) + t$, define

$B_t^{(w)}(l)$ = {all strings of length l consisting of **0**'s and **1**'s with exactly t **0**'s such that the number of **1**'s, before the first occurrence of **0** or between i-th and $(i+1)$-th occurrence of **0** for $i = 1, 2, \ldots, t-1$ or after the last (t-th) occurrence of **0**, is at least w}.

In other words, if r_0 is the number of $\mathbf{1}$'s before the first occurrence of $\mathbf{0}$ and r_i is the number of $\mathbf{1}$'s between i-th and $(i+1)$-th occurrences of $\mathbf{0}$ for $i = 1, 2, \dots, t-1$ and r_t is the number of $\mathbf{1}$'s after t-th occurrence of $\mathbf{0}$, then $r_i \geq w$ for all $i = 0, 1, \dots, t$ and $\sum_{i=0}^{t} r_i = l - t$. We obtain the probability of the event $B_t^{(w)}(l)$ in the next lemma.

LEMMA 4.3. *For any initial condition $x \in N_m$,*

$$P_x(B_t^{(w)}(l)) = \binom{l - w(t+1)}{t} \left(p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} \right) (p_{2^m-1})^{l-t-m(t+1)} \times (1 - p_{2^m-1})^t \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t . \tag{4.2}$$

PROOF. It is easy to note that, under the initial condition $x \in N_m$, the probability of any string in $B_t^{(w)}(l)$ is given by

$$\begin{aligned} & p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} (p_{2^m-1})^{r_0 - m} (1 - p_{2^m-1}) \\ & \times p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} (p_{2^m-1})^{r_1 - m} (1 - p_{2^m-1}) \\ & \cdots p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} (p_{2^m-1})^{r_{t-1} - m} (1 - p_{2^m-1}) \\ & \times p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} (p_{2^m-1})^{r_t - m} \\ = & \left(p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} \right) (p_{2^m-1})^{l-t-m(t+1)} (1 - p_{2^m-1})^t \\ & \times \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t \end{aligned} \tag{4.3}$$

since $\sum_{j=0}^{t} r_j = l - t$ and f_0 and f_1 are as defined earlier. Clearly, the probability (4.3) for any string in $B_t^{(w)}(l)$ is independent of the choice of r_i's. Thus, the probability of $B_t^{(w)}(l)$ is obtained by multiplying (4.3) with number of all such possible strings.

Now, using combinatorial arguments we calculate the number of such possible strings. Indeed, it is equivalent to distributing $l - t$ similar objects to $(t+1)$ groups so that each group has at least w objects. This is given by $\binom{l-w(t+1)}{t}$. This completes the proof of lemma. □

THEOREM 4.1. *Under any initial condition, if the condition* (4.1) *holds, then as* $r \to \infty$,

$$\text{(a)} \quad T_r^{(M)} - (r+1) \Rightarrow Z_1 \qquad \textit{when} \quad k \geq 1$$

$$\text{(b)} \quad T_r^{(M)} - (r+1) \Rightarrow Z_2 \qquad \textit{when} \quad k = 0$$

where $Z_i \sim \mathrm{Poi}(i, \mu_{2^m-1})$ *for* $i = 1, 2$.

PROOF OF THEOREM 4.1. We use Lemma 4.1 for this proof. To show part (a), we fix any $t \in \{0, 1, 2, \ldots\}$ and obtain a lower bound for the probability $P_x(T_r^{(M)} - (r+1) = t)$. In order to do so, we choose $w = \max(m, 1)$ and consider the event $B_t^{(w)}(r+1+t)$ for $r \geq w(t+1)$. Clearly,

$$B_t^{(w)}(r+1+t) \subseteq \{T_r^{(M)} - (r+1) = t\}$$

since for each string in $B_t^{(w)}(r+1+t)$, there are exactly r overlapping runs of type at most k $(k \geq 1)$. Thus we have,

$$\begin{aligned}
&P_x(T_r^{(M)} - (r+1) = t)\\
&\quad \geq P_x(B_t^{(w)}(r+1+t))\\
&\quad = \binom{r+1-w(t+1)+t}{t} p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} (p_{2^m-1})^{r+1-(t+1)m}\\
&\qquad \times (1 - p_{2^m-1})^t \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t\\
&\quad = \frac{r^t}{t!}(1+o(1))(1-p_{2^m-1})^t (p_{2^m-1})^r\\
&\qquad \times \left(p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} \right) (p_{2^m-1})^{(1-(t+1)m)} \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t\\
&\quad \longrightarrow \frac{\exp(-\mu_{2^m-1})(\mu_{2^m-1})^t}{t!}
\end{aligned}$$

since $r(1-p_u) \longrightarrow \mu_u$ as $r \longrightarrow \infty$ for $u \in N_m$. This, by Lemma 4.1, completes the proof of part (a).

For part (b), we consider the probability $P_x(T_r^{(M)} - (r+1) = 2t)$ and obtain a lower bound in a similar manner. Again, for $w = \max(m, 1)$ and $r \geq w(t+1)$, note that

$$B_t^{(w)}(r+1+2t) \subseteq \{T_r^{(M)} - (r+1) = 2t\}$$

since for any string in $B_t^{(w)}(r+1+2t)$, we have exactly r overlapping runs of type at most k $(k = 0)$. Therefore, we have

$$\begin{aligned}
&P_x(T_r^{(M)} - (r+1) = 2t)\\
&\quad \geq P_x(B_t^{(w)}(r+1+2t))\\
&\quad = \binom{r+1-w(t+1)+2t}{t} \left(p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} \right) (p_{2^m-1})^{r+1+t-(t+1)m}
\end{aligned}$$

$$\times (1-p_{2^m-1})^t \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t$$
$$= \frac{r^t}{t!}(1+o(1))(1-p_{2^m-1})^t (p_{2^m-1})^r$$
$$\times \left(p_x \prod_{j=1}^{m-1} p_{f_1^j(x)} \right) (p_{2^m-1})^{-(t+1)(m-1)} \left(p_{2^m-2} \prod_{j=1}^{m-1} p_{f_1^j(2^m-2)} \right)^t$$
$$\to \frac{\exp(-\mu_{2^m-1})(\mu_{2^m-1})^t}{t!},$$

as $r \to \infty$ by using the condition of the theorem. □

Remark. Since the limiting distribution is independent of the initial condition, we can put any initial distribution on the initial conditions. Suppose that θ is the probability distribution on $\{0,1\}^m$. As we have already discussed, θ can be identified as a probability measure on N_m by the mapping $(x_0, x_1, \ldots, x_{m-1}) \to x = \sum_{j=0}^{m-1} 2^j x_j$ where each $x_i \in \{0,1\}$. Let $T_r^{(M)}(\theta)$ be the waiting time for the r-th occurrence of the run of type at most k where the initial condition is governed by the distribution θ. From the theorem 4.1, we can easily conclude that

$$\text{(a)} \quad T_r^{(M)}(\theta) - (r+1) \Rightarrow Z_1 \quad \text{when} \quad k \geq 1$$
$$\text{(b)} \quad T_r^{(M)}(\theta) - (r+1) \Rightarrow Z_2 \quad \text{when} \quad k = 0$$

by first conditioning on $x \in N_m$ and then summing over all possible values of $x \in N_m$. The random variables, Z_1 and Z_2, are as in the Theorem 4.1.

Now, we derive the convergence results for scan waiting time variables $T_{r,k}^{(k_1,j)}$ for $j = I, II$ and III. It is clear that when $k = k_1$, the scan is equivalent to a run of successes of length k. In such a case, the waiting time variables $T_{r,k}^{(k_1,I)}$ and $T_{r,k}^{(k_1,II)}$ are same and they represent the waiting time for the r-th occurrence of the non-overlapping run of successes of length k, while $T_{r,k}^{(k_1,III)}$ represents the waiting time of the r-th occurrence of the overlapping run of successes of length k. Sarkar and Anuradha (2002) obtained the Poisson convergence of the waiting time distribution for the r-th occurrence for a more generalized run for the m-th order Markov chain under the same condition (4.1). This generalized run includes both overlapping runs as well as non-overlapping runs of successes of length k. We quote those results here without giving proofs.

THEOREM 4.2. *Under any initial condition, if* (4.1) *holds, then as* $r \to \infty$,

$$\text{(a)} \quad T_{r,k}^{(k_1,I)} - rk \Rightarrow \sum_{i=1}^{k} Z_i \quad \textit{when} \quad k_1 = k \geq 1$$
$$\text{(b)} \quad T_{r,k}^{(k_1,I)} - rk_1 \Rightarrow Z_1^\star \quad \textit{when} \quad k_1 < k$$
$$\text{(c)} \quad T_{r,k}^{(k_1,II)} - rk \Rightarrow \sum_{i=1}^{k} Z_i \quad \textit{when} \quad k_1 = k \geq 1$$

(d) $T_{r,k}^{(k_1,II)} - rk \Rightarrow 0$ *when* $k_1 < k$

(e) $T_{r,k}^{(k_1,III)} - (r+k-1) \Rightarrow Z_k$ *when* $k_1 = k \geq 1$

(f) $T_{r,k}^{(k_1,III)} - (r+k-1) \Rightarrow 0$ *when* $k_1 < k$

where for $i \geq 1$, $Z_i \sim \mathrm{Poi}(i, \mu_{2^m-1})$ *and are independent and* $Z_1^* \sim \mathrm{Poi}(1, k_1\mu_{2^m-1})$.

PROOF. We prove the cases (b) and (d) only. For proof of part (b), fix $t \geq 0$ and set $w = \max(k_1+1, m)$. We note that, $\{T_{r,k}^{(k_1,I)} - rk_1 = t\} \supseteq B_t^{(w)}(rk_1 + t)$ for all r such that $rk_1 \geq (t+1)w$. Now, using Lemma 4.3, we have

$$\liminf_{r\to\infty} P_x(T_{r,k}^{(k_1,I)} - rk_1 = t) \geq \liminf_{r\to\infty} P_x(B_t^{(w)}(rk_1+t)) = \exp(-k_1\mu_{2^m-1})\frac{(k_1\mu_{2^m-1})^t}{t!}.$$

This, by Lemma 4.1, is sufficient for our purpose.

For part (d), fix $\epsilon > 0$ and choose $J \geq 1$ such that $\sum_{j=0}^{J} \exp(-k\lambda_{2^m-1})(k\lambda_{2^m-1})^j/j! \geq 1-\epsilon$. Set $w = \max(k+1, m)$ and note that, for r so large that $rk \geq (J+1)w + J$, we have

$$\bigcup_{t=0}^{J} B_t^{(w)}(rk) \subseteq \{T_{r,k}^{(k_1,II)} = rk\}.$$

Thus, we have,

$$\begin{aligned}\liminf_{r\to\infty} P_x(T_{r,k}^{(k_1,II)} = rk) &\geq \liminf_{r\to\infty} \sum_{t=0}^{J} P_x(B_t^{(w)}(rk)) \\ &= \sum_{t=0}^{J} \exp(-k\lambda_{2^m-1})(k\lambda_{2^m-1})^t/t! \\ &\geq 1-\epsilon.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, this proves the result. The proof of part (f) is similar and we omit it. □

Remark. As remarked earlier, since the limiting distribution does not depend upon the $x \in N_m$, same results continue to hold, when we replace the initial condition by a distribution on θ on N_m.

4.2 *Strong tendency towards reversing state*

More interesting results are obtained when the system has a strong tendency of reversing its state, i.e., if it is in state **1** it would tend to switch to state **0** and vice versa. Formally speaking, we set,

$$\left.\begin{array}{ll} rp_u \to \mu_u & \text{if } u \text{ is odd (i.e., } X_0 = 1) \\ r(1-p_u) \to \mu_u & \text{if } u \text{ is even (i.e., } X_0 = 0) \end{array}\right\} \quad \text{as} \quad r \to \infty \tag{4.4}$$

where $\mu_u > 0$ for all $u \in N_m$. For next few theorems we require a lower bound for the probability of the following event which we obtain in the next lemma. Suppose that

$j_0, j_1 \geq 0$ are given and l is sufficiently large and $w \geq m$. Define the event

$$B_l^{(w)}(j_0, j_1) = \{\text{all strings of length } l \text{ consisting of } \mathbf{1}\text{'s and } \mathbf{0}\text{'s, starting with } \mathbf{1} \text{ and having exactly } j_0/j_1 \text{ occurrences of } \mathbf{00}/\mathbf{11} \text{ respectively and no runs of } \mathbf{1}\text{'s of length 3 or more and no runs of } \mathbf{0}\text{'s of length 3 or more are present and between any two occurrences of } (\mathbf{00} \text{ or } \mathbf{11}) \text{ there are at least } w \text{ symbols (including both types of } \mathbf{1}\text{'s and } \mathbf{0}\text{'s)}\}.$$

The event $B_l^{(w)}(j_0, j_1)$ comprises of all strings of length l, consisting of alternating $\mathbf{1}$'s and $\mathbf{0}$'s, starting with $\mathbf{1}$; however sometimes $\mathbf{1}$ may follow a $\mathbf{1}$ and $\mathbf{0}$ may follow a $\mathbf{0}$. There are exactly j_0 occurrences of $\mathbf{00}$'s and j_1 occurrences of $\mathbf{11}$'s. Further we insist that between any two such occurrences ($\mathbf{00}$ or $\mathbf{11}$) there are at least w symbols (including both types of $\mathbf{1}$'s and $\mathbf{0}$'s).

Define now

$$\alpha(m) = \begin{cases} 2(2^m - 1)/3 & \text{if } m \text{ is even} \\ (2^m - 2)/3 & \text{if } m \text{ is odd} \end{cases} \tag{4.5}$$

and

$$\beta(m) = (2\alpha(m) + 1)(\text{mod } 2^m) = \begin{cases} (2^m - 1)/3 & \text{if } m \text{ is even} \\ (2^{m+1} - 1)/3 & \text{if } m \text{ is odd.} \end{cases} \tag{4.6}$$

Let

$$\kappa_0^{(m)}(l; j_0, j_1) = [(l - (j_0 + j_1))/2] - (j_0 + 1)[m/2] - j_1[(m+1)/2] \tag{4.7}$$

and

$$\kappa_1^{(m)}(l; j_0, j_1) = [(l + 1 - (j_0 + j_1))/2] - (j_0 + 1)[(m+1)/2] - j_1[m/2] \tag{4.8}$$

where $[s]$ is the largest integer less than or equal to s. Let $y \in N_m$ be any initial condition. Define

$$\gamma_0(y, m) = \begin{cases} \prod_{j=0}^{s-1}(1 - p_{h_j^{(0)}(y)})p_{f_0(h_j^{(0)}(y))} & \text{if } m = 2s \text{ is even} \\ \prod_{j=0}^{s-1}(1 - p_{h_j^{(0)}(y)})p_{f_0(h_j^{(0)}(y))}(1 - p_{h_s^{(0)}(y)}) & \text{if } m = 2s+1 \text{ is odd} \end{cases} \tag{4.9}$$

and

$$\gamma_1(y, m) = \begin{cases} \prod_{j=0}^{s-1} p_{h_j^{(1)}(y)}(1 - p_{f_1(h_j^{(1)}(y))}) & \text{if } m = 2s \text{ is even} \\ \prod_{j=0}^{s-1} p_{h_j^{(1)}(y)}(1 - p_{f_1(h_j^{(1)}(y))})p_{h_s^{(1)}(y)} & \text{if } m = 2s+1 \text{ is odd} \end{cases} \tag{4.10}$$

where $h_0^{(0)}(y) = h_0^{(1)}(y) = y$ and $h_j^{(0)}(y) = f_1(f_0(h_{j-1}^{(0)}(y)))$ and $h_j^{(1)}(y) = f_1(f_0(h_{j-1}^{(1)}(y)))$ for all $j \geq 1$. Now we have,

LEMMA 4.4. *For any initial condition $x \in N_m$ which is even, we have*

$$P_x(B_l^{(w)}(j_0, j_1)) \tag{4.11}$$

$$
\begin{aligned}
&\geq p_{\alpha(m)}^{\kappa_1^{(m)}(l;j_0,j_1)}(1-p_{\beta(m)})^{\kappa_0^{(m)}(l;j_0,j_1)} \\
&p_{\beta(m)}^{j_1}(1-p_{\alpha(m)})^{j_0}\gamma_1(x,m)\gamma_1(\alpha(m)+1,m)^{j_1}\gamma_0(\beta(m)-1,m)^{j_0} \\
&\binom{[(l-2w-(j_0+j_1)(2w+1))/2]}{j_0} \\
&\binom{[(1+l-2w-(j_0+j_1)(2w+1))/2]}{j_1}.
\end{aligned}
$$

PROOF. To find the probability of a string in $B_l^{(w)}(j_0,j_1)$, we note that, since $w \geq m$, each occurrence of **11** is followed by at least m symbols of alternating **0**'s and **1**'s, starting with **0** ($\underbrace{\mathbf{0101}\ldots}_{\geq m}$) and each occurrences of **00** is followed by at least m symbols of alternating **1**'s and **0**'s, starting with **1** ($\underbrace{\mathbf{1010}\ldots}_{\geq m}$). Therefore, except the **1**'s which occur, at trials not greater than m from the start of the string, or within m trials from an occurrence of either **00** or **11** or the **1**'s which is preceded by another **1** (the second **1** in **11**), all **1**'s will be preceded by m symbols of alternating **0**'s and **1**'s ending with a **0** ($\underbrace{\ldots\mathbf{1010}}_{m}$). Note that we only have to look m-trials backwards to calculate the probability of any occurrence (**0** or **1**) since the model is m-dependent. Similarly, except the **0**'s which occur, at trials less than or equal to m from the start of the string, or within m trials from an occurrence of either **00** or **11** or the **0** which is preceded by another **0** (the second **0** in **00**), all **0**'s will be preceded by $\underbrace{\ldots\mathbf{0101}}_{m}$. We illustrate this by an example with $m=5$ and $j_0=1$ and $j_1=2$:

$$
\overbrace{\mathbf{10101}}^{m}\mathbf{0100}\overbrace{\mathbf{10101}}^{m}\mathbf{01011}\overbrace{\mathbf{01010}}^{m}\mathbf{1011}\overbrace{\mathbf{01010}}^{m}\mathbf{101}
$$

$$
\star\circ\star\circ\star \qquad \circ\star\circ\star\circ\star \qquad \star\circ\star\circ\star\circ \qquad \star\circ\star\circ\star\circ
$$

Here, **1**'s and the **0**'s which are marked by a $\star$ and a $\circ$ respectively, are the symbols which we leave out in the counting. Remaining **1**'s and **0**'s are preceded by **01010** and **10101** respectively.

Now, the probability of a **1** which is preceded by $\underbrace{\ldots\mathbf{1010}}_{m}$ is given by $p_{\alpha(m)}$ and the probability of a **0** which is preceded by $\underbrace{\ldots\mathbf{0101}}_{m}$ is given by $1-p_{\beta(m)}$ where $\alpha(m)$ and $\beta(m)$ are defined in (4.5) and (4.6). We can find the number of **1**'s, denoted by $\kappa_1(m)$, which are preceded by $\underbrace{\ldots\mathbf{1010}}_{m}$. When m is odd, $m=2s+1$, at the start of the string and after each occurrence of **00**, there are $(s+1)=[(m+1)/2]$ many **1**'s and $s=[m/2]$ many **0**'s which occur at trials no larger than m from the start of the string or the occurrence of **00** respectively. Similarly, for each occurrence of **11**, we have $s=[m/2]$ many **1**'s and $(s+1)=[(m+1)/2]$ many **0**'s which occur at trials no larger than m from the occurrence of **11**. Therefore, the number of **1**'s which are preceded by $\underbrace{\ldots\mathbf{1010}}_{m}$, is given by $[(l+1-(j_0+j_1))/2]-j_1s-(j_0+1)(s+1)=[(l+1-(j_0+j_1))/2]-(j_0+1)[(m+1)/2]-j_1[m/2]$. Arguing similarly, we have, if m is even, $m=2s$, then the number of successes preceded by $\underbrace{\ldots 1010}_{m}$ is given by, $[(l+1-(j_0+j_1))/2] - s(j_0+j_1+1) = [(l+1-(j_0+j_1))/2] - [m/2](j_0+j_1+1) = [(l+1-(j_0+j_1))/2] - (j_0+1)[(m+1)/2] - j_1[m/2]$. Similarly, the number of **0**'s which are preceded by $\underbrace{\ldots 0101}_{m}$ is denoted by $\kappa_0^{(m)}(l; j_0, j_1)$ and is given in (4.7).

For **1**'s which are preceded by $\underbrace{\ldots 0101}_{m}$, the probability is given by $p_{\beta(m)}$ and **0**'s which are preceded by $\underbrace{\ldots 1010}_{m}$, the probability is given by $1 - p_{\alpha(m)}$.

Finally, for any string of length m, starting with **1** and consisting of alternating **1**'s and **0**'s, preceded by any initial condition $y = \sum_{j=0}^{m-1} 2^j y_j \in N_m$, the probability can be computed in the following way. Suppose that $m = 2s$, then, the probability

$$\begin{aligned}
\gamma_1(y,m) &= P(X_1 = 1, X_2 = 0, \ldots, X_{2s-1} = 1, \\
&\qquad X_{2s} = 0 \mid X_0 = y_0, X_{-1} = y_1, \ldots, X_{-m+1} = y_{m-1}) \\
&= p_{h_0^{(1)}(y)}(1 - p_{f_1(h_0^{(1)}(y))}) p_{h_1^{(1)}(y)}(1 - p_{f_1(h_1^{(1)}(y))}) \cdots p_{h_{s-1}^{(1)}(y)}(1 - p_{f_1(h_{s-1}^{(1)}(y))}) \\
&= \prod_{j=0}^{s-1} p_{h_j^{(1)}(y)}(1 - p_{f_1(h_j^{(1)}(y))})
\end{aligned}$$

where $h_0^{(1)}(y) = y$ and $h_j^{(1)}(y) = f_0(f_1(h_{j-1}^{(1)}(y)))$ for $j = 1, 2, \ldots, s-1$. When $m = 2s+1$, the probability can be computed similarly and is given by

$$\begin{aligned}
\gamma_1(y,m) &= P(X_1 = 1, X_2 = 0, \ldots, X_{2s} = 0, \\
&\qquad X_{2s+1} = 1 \mid X_0 = y_0, X_{-1} = y_1, \ldots, X_{-m+1} = y_{m-1}) \\
&= \prod_{j=0}^{s-1} p_{h_j^{(1)}(y)}(1 - p_{f_1(h_j^{(1)}(y))}) p_{h_s^{(1)}(y)}.
\end{aligned}$$

Similar computations can be carried out for strings of length m starting with **a 0** and consisting of alternating **0**'s and **1**'s. Indeed, the probability is denoted by $\gamma_0(y, m)$ and is given in (4.9).

Combining all the results, the probability of any string in $B_l^{(w)}(j_0, j_1)$ is given by

$$\begin{aligned}
(4.12)\qquad & p_{\alpha(m)}^{\kappa_1^{(m)}(l;j_0,j_1)}(1 - p_{\beta(m)})^{\kappa_0^{(m)}(l;j_0,j_1)} p_{\beta(m)}^{j_1}(1 - p_{\alpha(m)})^{j_0} \\
& \times \gamma_1(x,m)\gamma_1(\alpha(m)+1,m)^{j_1}\gamma_0(\beta(m)-1,m)^{j_0}.
\end{aligned}$$

We observe that the probability of the string is actually independent of the positions of occurrences of **00** and **11**. Therefore, the probability of $P_x(B_l^{(w)}(j_0, j_1))$ can be obtained by multiplying the above probability with the number of possible permutations. However, it is rather difficult to calculate the exact number of permutations, hence we obtain a lower bound which is sufficient for our purpose.

Now, to obtain a lower bound of the number of permutations, we proceed as follows: we assume that at the start or after each occurrence of **00**, there are w pairs of **10**, i.e., $\underbrace{1010\ldots 1010}_{2w}$ and each occurrence of **11** is followed by w pairs of **01**, i.e.,

$\underbrace{0101\ldots0101}_{2w}$. We can employ combinatorial methods to compute the lower bound as follows: suppose that we have two classes, A and B. The class A, consists of two types of elements **1** and $\mathbf{11}\underbrace{0101\ldots0101}_{2w}$, with exactly j_1 elements of type $\mathbf{11}\underbrace{0101\ldots0101}_{2w}$ while the class B contains two types of elements **0** and $\mathbf{00}\underbrace{1010\ldots1010}_{2w}$, with j_0 elements of type $\mathbf{00}\underbrace{1010\ldots1010}_{2w}$. First, we fill up the first $2w$ places with $\underbrace{1010\ldots1010}_{2w}$ which can only be done in one way. Now, we are left with $l-2w-j_0(2w+2)-j_1(2w+2)+j_0+j_1 = l-2w-(j_0+j_1)(2w+1)$ many places, the odd places of which should be filled up with elements from class A and the even places by elements from class B. This can be done in

$$\binom{[(l-2w-(j_0+j_1)(2w+1))/2]}{j_0}\binom{[(1+l-2w-(j_0+j_1)(2w+1))/2]}{j_1} \tag{4.13}$$

ways. This certainly gives a lower bound for the number of permutations. Therefore, combining (4.12) and (4.13), we get the result. □

Now, using Lemma 4.4, we will analyze the waiting time distributions for various types of runs.

THEOREM 4.3. *If the initial condition is that x is even (i.e., $X_0 = 0$) and the condition* (4.4) *holds, we have*

$$\text{(a)}\quad T_r^{(M)} - (2r+1) \Rightarrow Z_3^{(0)} + Z_{-1}^{(1)} \quad \textit{when} \quad k = 1$$
$$\text{(b)}\quad T_r^{(M)} - (2r+1) \Rightarrow Z_1^{(0)} + Z_{-1}^{(1)} \quad \textit{when} \quad k \geq 2$$

where $Z_1^{(0)} \sim \mathrm{Poi}(1, \mu_{\alpha(m)})$, $Z_3^{(0)} \sim \mathrm{Poi}(3, \mu_{\alpha(m)})$ and $Z_{-1}^{(1)} \sim \mathrm{Poi}(-1, \mu_{\beta(m)})$ and $Z_{-1}^{(1)}$ is independent of both $Z_1^{(0)}$ and $Z_3^{(0)}$.

PROOF. For part (a), fix any $t \in \mathbb{Z}$ and consider the set $L_t = \{(j_0, j_1) : j_0, j_1 \geq 0,\ 3j_0 - j_1 = t\}$. Fix any $(j_0, j_1) \in L_t$ and define the event $A_{(j_0,j_1)}^{(2r+t+1)} = \{$the string of length $(2r+1)+t$ has exactly j_0 occurrences of **00** and exactly j_1 occurrences of **11**$\}$. Whenever it is not possible to find runs of the above kind, we set $A_{(j_0,j_1)}^{(2r+t+1)}$ as the empty set. Note that for any fixed t and $(j_1, j_0) \in L_t$, the set $A_{(j_0,j_1)}^{(2r+t+1)}$ will be non-empty for all sufficiently large r.

Set $w = \max(4, m)$ and note that each string in $B_{2r+1+t}^{(w)}(j_0, j_1)$ also belongs to $A_{(j_0,j_1)}^{(2r+t+1)}$ and further for each string in $B_{2r+1+t}^{(w)}(j_0, j_1)$, we have $T_r^{(M)} = (2r+1)+t$ for all sufficiently large r. Thus, we get,

$$\{T_r^{(M)} - (2r+1) = t, A_{(j_0,j_1)}^{(2r+t+1)}\} \supseteq B_{2r+1+t}^{(w)}(j_0, j_1)$$

for all r sufficiently large. Hence, from Lemma 4.4, we have,

$$P_x(T_r^{(M)} - (2r+1) = t, A_{(j_0,j_1)}^{(2r+t+1)})$$

$$\begin{aligned}
&\geq P_x(B^{(w)}_{2r+1+t}(j_0,j_1))\\
&\geq p_{\alpha(m)}^{\kappa_1^{(m)}(2r+t+1;j_0,j_1)}(1-p_{\beta(m)})^{\kappa_0^{(m)}(2r+t+1;j_0,j_1)}p_{\beta(m)}^{j_1}(1-p_{\alpha(m)})^{j_0}\\
&\quad\times \gamma_1(x,m)\gamma_1(\alpha(m)+1,m)^{j_1}\gamma_0(\beta(m)-1,m)^{j_0}\\
&\quad\times \binom{[(2r+t+1-2w-(2w+1)(j_0+j_1))/2])}{j_0}\\
&\quad\times \binom{[(2r+t+2-2w-(j_0+j_1)(2w+1))/2]}{j_1}\\
&= p^r_{\alpha(m)}(1-p_{\beta(m)})^r(1-p_{\alpha(m)})^{j_0}p^{j_1}_{\beta(m)}(1+o(1))\times\frac{r^{j_1+j_0}}{j_1!j_0!}(1+o(1))\\
&\to \frac{\exp(-\mu_{\alpha(m)})(\mu_{\alpha(m)})^{j_0}}{j_0!}\frac{\exp(-\mu_{\beta(m)})(\mu_{\beta(m)})^{j_1}}{j_1!}
\end{aligned}$$

and this proves the result.

For part (b), we follow a similar procedure. For any $t\in\mathbb{Z}$, let L_t be the set $\{(j_1,j_0):j_0,j_1\geq 0, j_0-j_1=t\}$. Set $w=\max(4,m)$. Since $k\geq 2$, for each string in $B^{(w)}_{2r+1+t}(j_0,j_1)$, the r-th occurrence of the run of type at most $k(\geq 2)$ is completed exactly on $(2r+1)+t$. Thus, we have, for all sufficiently large r,

$$\{T^{(M)}_r-(2r+1)=t, A^{(2r+t+1)}_{(j_0,j_1)}\}\supseteq B^{(w)}_{2r+1+t}(j_0,j_1)$$

where $A^{(2r+t+1)}_{(j_0,j_1)}$ is as defined above. So, using Lemma 4.4, we get

$$\begin{aligned}
&P_x(T^{(M)}_r-(2r+1)=t, A^{(2r+t+1)}_{(j_0,j_1)})\\
&\quad\geq P_x(B^{(w)}_{2r+1+t}(j_0,j_1))\\
&\quad\geq p^r_{\alpha(m)}(1-p_{\beta(m)})^r(1-p_{\alpha(m)})^{j_0}p^{j_1}_{\beta(m)}(1+o(1))\times\frac{r^{j_0+j_1}}{j_0!j_1!}(1+o(1))\\
&\quad\to \frac{\exp(-\mu_{\alpha(m)})(\mu_{\alpha(m)})^{j_0}}{j_0!}\frac{\exp(-\mu_{\beta(m)})(\mu_{\beta(m)})^{j_1}}{j_1!}
\end{aligned}$$

as $r\to\infty$. This completes the proof of the theorem. □

Remark. If the initial condition x is odd, i.e., $X_0=1$, then the string will start with a **0** at the first trial and then the rest of the argument can be easily carried over to prove that,

$$\begin{aligned}
&\text{(a)}\quad T^{(M)}_r-(2r+2)\Rightarrow Z^{(0)}_3+Z^{(1)}_{-1}\quad\text{when}\quad k=1\\
&\text{(b)}\quad T^{(M)}_r-(2r+2)\Rightarrow Z^{(0)}_1+Z^{(1)}_{-1}\quad\text{when}\quad k\geq 2
\end{aligned}$$

where $Z^{(0)}_1$, $Z^{(0)}_3$ and $Z^{(1)}_{-1}$ are defined in Theorem 4.3. Suppose that θ is any probability distribution on N_m, the set of all initial conditions. By first conditioning on $x\in N_m, x$ even and then summing over all x even, we can establish that with probability $\theta(\text{even})=\sum_{x\text{even}}\theta(x)$, $T^{(M)}_r(\theta)-(2r+1)$ converges to $(Z^{(0)}_3+Z^{(1)}_{-1})$ if $k=1$ and to $(Z^{(0)}_1+Z^{(1)}_{-1})$ if $k\geq 2$. Similarly, with probability $\theta(\text{odd})$, $T^{(M)}_r(\theta)-(2r+1)$ converges to $1+(Z^{(0)}_3+Z^{(1)}_{-1})$

if $k = 1$ and to $1 + (Z_1^{(0)} + Z_{-1}^{(1)})$ if $k \geq 2$. These can be combined together as follows: let X be Bernoulli random variable with $P(X = 1) = \theta(\text{odd})$ and be independent of random variables $Z_1^{(0)}$, $Z_3^{(0)}$ and $Z_{-1}^{(1)}$. Then,

$$\begin{aligned}
&\text{(a)} \quad T_r^{(M)}(\theta) - (2r+1) \Rightarrow X + Z_3^{(0)} + Z_{-1}^{(1)} \quad \text{when} \quad k = 1\\
&\text{(b)} \quad T_r^{(M)}(\theta) - (2r+1) \Rightarrow X + Z_1^{(0)} + Z_{-1}^{(1)} \quad \text{when} \quad k \geq 2.
\end{aligned}$$

Next, we study the case of runs of type exactly k.

THEOREM 4.4. *If the initial condition is that x is even (i.e., $X_0 = 0$) and the condition* (4.4) *holds, we have*

$$T_r^{(E)} - (2r+1) \Rightarrow Z_3^{(0)} + Z_1^{(1)} \quad \textit{when} \quad k = 1$$

where $Z_1^{(1)} \sim \text{Poi}(1, \mu_{\beta(m)})$ and $Z_3^{(0)} \sim \text{Poi}(3, \mu_{\alpha(m)})$ and $Z_1^{(1)}$ and $Z_3^{(0)}$ are independent.

PROOF. The proof of this is similar. For $t \in \{0, 1, \ldots\}$, set $L_t = \{(j_1, j_0) : j_0, j_1 \geq 0, 3j_0 + j_1 = t\}$ and consider the probability $P_x(T_r^{(E)} - (2r+1) = t, A_{(j_0,j_1)}^{(2r+t+1)})$ where $A_{(j_0,j_1)}^{(2r+t+1)}$ is as defined earlier. Again, for each string in $B_{2r+1+t}^{(w)}(j_0, j_1)$, we have $T_r^{(E)} = (2r+1) + t$ where $w = \max(4, m)$. Hence, we have, for all sufficiently large r,

$$\{T_r^{(E)} - (2r+1) = t, A_{(j_0,j_1)}^{(2r+t+1)}\} \supseteq B_{2r+1+t}^{(w)}(j_0, j_1).$$

So, using Lemma 4.4, we have

$$\begin{aligned}
&P_x(T_r^{(E)} - (2r+1) = t, A_{(j_0,j_1)}^{(2r+t+1)})\\
&\quad \geq P_x(B_{2r+1+t}^{(w)}(j_0, j_1))\\
&\quad \geq p_{\alpha(m)}^r (1 - p_{\beta(m)})^r (1 - p_{\alpha(m)})^{j_0} p_{\beta(m)}^{j_1} (1 + o(1)) \times \frac{r^{j_0+j_1}}{j_0! j_1!} (1 + o(1))\\
&\quad \to \frac{\exp(-\mu_{\alpha(m)})(\mu_{\alpha(m)})^{j_0}}{j_0!} \frac{\exp(-\mu_{\beta(m)})(\mu_{\beta(m)})^{j_1}}{j_1!}
\end{aligned}$$

as $r \to \infty$. This completes the proof of the theorem. □

Remark. If the initial condition x is odd, as before, we can establish

$$T_r^{(E)} - (2r+2) \Rightarrow Z_3^{(0)} + Z_1^{(1)} \quad \text{when} \quad k = 1$$

where $Z_3^{(0)}$ and $Z_1^{(1)}$ are as specified in Theorem 4.4. Hence, given any probability distribution θ on N_m,

$$T_r^{(E)}(\theta) - (2r+1) \Rightarrow X + Z_3^{(0)} + Z_1^{(1)} \quad \text{when} \quad k = 1$$

where X is an independent Bernoulli random variable with $P(X = 1) = \theta(\text{odd})$.

Now, we consider the case of runs of type at least k. The proof of the next theorem is similar to the proof of the Theorem 4.3 and we omit it.

THEOREM 4.5. *If the initial condition is that x is even (i.e., $X_0 = 0$) and the condition* (4.4) *holds, we have*

$$\begin{aligned} &\text{(a)}\quad T_r^{(L)} - (2r+1) \Rightarrow Z_1^{(0)} + Z_{-1}^{(1)} \qquad \textit{when} \quad k=0\\ &\text{(b)}\quad T_r^{(L)} - (2r+1) \Rightarrow Z_1^{(0)} + Z_1^{(1)} \qquad \textit{when} \quad k=1 \end{aligned}$$

where $Z_1^{(0)} \sim \mathrm{Poi}(1, \mu_{\alpha(m)})$, $Z_1^{(1)} \sim \mathrm{Poi}(1, \mu_{\beta(m)})$ and $Z_{-1}^{(1)} \sim \mathrm{Poi}(-1, \mu_{\beta(m)})$ and $Z_1^{(0)}$ is independent of both $Z_{-1}^{(1)}$ and $Z_1^{(1)}$.

Remark. If θ is any distribution on the initial conditions, we have

$$\begin{aligned} &\text{(a)}\quad T_r^{(L)}(\theta) - (2r+1) \Rightarrow X + Z_1^{(0)} + Z_{-1}^{(1)} \qquad \text{when} \quad k=0\\ &\text{(b)}\quad T_r^{(L)}(\theta) - (2r+1) \Rightarrow X + Z_1^{(0)} + Z_1^{(1)} \qquad \text{when} \quad k=1 \end{aligned}$$

where $Z_1^{(0)}$, $Z_{-1}^{(1)}$ and $Z_1^{(1)}$ are as in Theorem 4.5 and X is an independent Bernoulli random variable with $P(X=1) = \theta(\text{odd})$.

Finally we investigate the limiting distributions of scan waiting times under this set up. Unfortunately the methods employed here are not sophisticated enough to obtain the limiting distributions of $T_{r,k}^{(k_1,I)}$ and $T_{r,k}^{(k_1,II)}$. However, these are good enough for the waiting time distribution of $T_{r,k}^{(k_1,III)}$.

Two different scenarios emerge for the scan waiting time $T_{r,k}^{(k_1,III)}$, namely: (a) k is even and (b) k is odd. We study them separately.

THEOREM 4.6. *Under any initial condition, if k is even and the condition* (4.4) *holds, we have*

$$\begin{aligned} &\text{(a)}\quad T_{r,k}^{(k_1,III)} - (r+k-1) \Rightarrow Z_{k_1}^{(0)} \qquad \textit{when} \quad k = 2k_1\\ &\text{(b)}\quad T_{r,k}^{(k_1,III)} - (r+k-1) \Rightarrow 0 \qquad \textit{when} \quad k > 2k_1 \end{aligned}$$

where $Z_{k_1}^{(0)} \sim \mathrm{Poi}(k_1, \mu_{\alpha(m)}/2)$.

PROOF. For part (a), fix $\epsilon > 0$. Choose J large so that $\sum_{j=0}^{J} \exp(-\frac{\mu_{\beta(m)}}{2})(\frac{\mu_{\beta(m)}}{2})^j / j! \geq 1-\epsilon$. Fix any $t \geq 0$. For $w = \max(4, m, 2k)$ and for all sufficiently large r, we have

$$\begin{aligned} \{T_{r,k}^{(k_1,III)} - (r+k-1) = k_1 t\} &\supseteq \bigcup_{j=0}^{J} \{T_{r,k}^{(k_1,III)} - (r+k-1) = k_1 t, A_{(t,j)}^{(r+k-1+k_1 t)}\}\\ &\supseteq \bigcup_{j=0}^{J} B_{r+k-1+k_1 t}^{(w)}(t,j). \end{aligned}$$

Therefore, we obtain,

$$\begin{aligned}
&P_x\{T_{r,k}^{(k_1,III)} - (r+k-1) = k_1 t\} \\
&\geq \sum_{j=0}^{J} P_x(B_{r+k-1+k_1 t}^{(w)}(t,j)) \\
&\geq \sum_{j=0}^{J} p_{\alpha(m)}^{r/2}(1-p_{\beta(m)})^{r/2}(1-p_{\alpha(m)})^t p_{\beta(m)}^j (1+o(1)) \times \frac{(r/2)^{t+j}}{t!j!}(1+o(1)) \\
&\to \sum_{j=0}^{J} \frac{\exp(-\mu_{\alpha(m)}/2)(\mu_{\alpha(m)}/2)^t}{t!} \frac{\exp(-\mu_{\beta(m)}/2)(\mu_{\beta(m)}/2)^j}{j!} \\
&\geq (1-\epsilon)\frac{\exp(-\mu_{\alpha(m)}/2)(\mu_{\alpha(m)}/2)^t}{t!}.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, this proves that $\liminf_{r\to\infty} P\{T_{r,k}^{(k_1,III)} - (r+k-1) = k_1 t\} \geq \exp(-\mu_{\alpha(m)}/2)(\mu_{\alpha(m)}/2)^t/t!$. This proves part (a).

For part (b), we choose $\epsilon > 0$ and J so large that $\sum_{j=0}^{J} \exp(-\mu_{\alpha(m)}/2)(\mu_{\alpha(m)}/2)^j / j! \geq 1-\epsilon$ and $\sum_{j=0}^{J} \exp(-\mu_{\beta(m)}/2)(\mu_{\beta(m)}/2)^j/j! \geq 1-\epsilon$. Now, for all sufficiently large r and w as defined above,

$$\{T_{r,k}^{(k_1,III)} - (r+k-1) = 0\} \supseteq \bigcup_{j_0,j_1=1}^{J} B_{r+k-1}^{(w)}(j_0,j_1).$$

Therefore, using Lemma 4.4, we have that $\liminf_{r\to\infty} P\{T_{r,k}^{(k_1,III)} - (r+k-1) = 0\} \geq (1-\epsilon)^2$. Since $\epsilon > 0$ is arbitrary, this completes the proof. □

Remark. When $k < 2k_1$, it is easy to conclude that the scan waiting times converge to infinity. As earlier, we can start with a initial distribution θ on N_m. Since the limiting distribution is independent of the initial condition, we have,

(a) $T_{r,k}^{(k_1,III)}(\theta) - (r+k-1) \Rightarrow Z_{k_1}^{(0)}$ when $k = 2k_1$

(b) $T_{r,k}^{(k_1,III)}(\theta) - (r+k-1) \Rightarrow 0$ when $k > 2k_1$

where $Z_{k_1}^{(0)}$ is same as in Theorem 4.6.

Next we explore the case when k is odd.

THEOREM 4.7. *If the initial condition is that x is even (i.e., $X_0 = 0$), k is odd and the condition* (4.4) *holds, we have*

(a) $T_{r,k}^{(k_1,III)} - k - 2(r-1) \Rightarrow Z_k^{(0)} + Z_{-k}^{(1)}$ *when* $k = 2k_1 - 1$

(b) $T_{r,k}^{(k_1,III)} - k - (r-1) \Rightarrow 0$ *when* $k > 2k_1 - 1$

where $Z_k^{(0)} \sim \mathrm{Poi}(k, \mu_{\alpha(m)})$ *and* $Z_{-k}^{(1)} \sim \mathrm{Poi}(-k, \mu_{\beta(m)})$ *and* $Z_k^{(0)}$ *and* $Z_{-k}^{(1)}$ *are independent.*

PROOF. We only prove part (a). For any $t \in \mathbb{Z}$, we define $L_t = \{(j_0, j_1) : j_0, j_1 \geq 0, j_0 - j_1 = t\}$. For any $(j_0, j_1) \in L_t$, consider the event $A_{(j_0,j_1)}^{(k+2(r-1)+kt)}$. For all sufficiently large r, this is a well-defined event and also note,

$$\{T_{r,k}^{(k_1,III)} - k - 2(r-1) = kt, A_{(j_0,j_1)}^{(k+2(r-1)+kt)}\} \supseteq B_{k+2(r-1)+kt}^{(w)}(j_0, j_1)$$

where $w = \max(4, m, 2k)$. Therefore, using Lemma 4.4, we conclude that,

$$\begin{aligned}\liminf_{r\to\infty} P\{T_{r,k}^{(k_1,III)} - k - 2(r-1) = kt, A_{(j_0,j_1)}^{(k+2(r-1)+kt)}\} \\ \geq \frac{\exp(-\mu_{\alpha(m)})(\mu_{\alpha(m)})^{j_0}}{j_0!} \frac{\exp(-\mu_{\beta(m)})(\mu_{\beta(m)})^{j_1}}{j_1!}.\end{aligned}$$

This completes the proof. □

Remark. When $k < 2k_1 - 1$, it is easy to observe that the waiting times converge to infinity. If θ is any initial distribution on N_m, we have,

(a) $T_{r,k}^{(k_1,III)}(\theta) - k - 2(r-1) \Rightarrow Z_k^{(0)} + Z_{-k}^{(1)} + X$ when $k = 2k_1 - 1$

(b) $T_{r,k}^{(k_1,III)}(\theta) - k - 2(r-1) \Rightarrow 0$ when $k > 2k_1 - 1$

where $Z_k^{(0)}$ and $Z_{-k}^{(1)}$ are as in Theorem 4.7 and X is an independent Bernoulli random variable with $P(X = 1) = \theta(\text{odd})$.

5. Central limit theorem

In this section, we derive central limit theorem for the number of runs up to time n, of different types, viz., $M_n^{(M)}$, $M_n^{(E)}$ and $M_n^{(L)}$, as well as the number of scans up to time n, denoted by $S_{k_1,k}^{(j)}(n)$ for $j = I, II$ and III, for the stationary m-th order Markov chain set up. Further, for the same set up, we establish law of the iterated logarithm for $M_n^{(M)}$, $M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(III)}(n)$.

We define a sequence of new random variables which form a stationary Markov chain with finite state space and translate the description of runs of all types as well as the scans from the set of original random variables to the set of newly defined random variables. We assume that $0 < p_x < 1$ for all $x \in N_m$; therefore this new Markov chain is irreducible. Hence the new Markov chain, being stationary with finite state space and irreducible, forms an α-mixing sequence. Using results from the central limit theorem for α-mixing sequences, we derive a very general central limit theorem on arbitrary functions of the newly defined random variables. Further, we establish moment bounds on the new set of random variables and applying the bounds in the α-mixing setup, we obtain law of the iterated logarithm for any function of the newly defined random variables. Now, by appropriately choosing the function in the above result, we obtain the central limit theorem and law of the iterated logarithm for $M_n^{(M)}$, $M_n^{(E)}$ and $S_{(k_1,k)}^{(III)}(n)$. In case of runs of type at least k, the above method cannot be directly applied since there is no upper bound on the length of the run. However, we can construct a pattern, hence an associated function, so that $M_n^{(L)}$ can be approximated by the number of above patterns till time n. This approximation enables us to use the above general theorem to derive the asymptotic results for $M_n^{(L)}$.

In order to obtain the results for the scan enumerating variables, $S^{(I)}_{k_1,k}(n)$ and $S^{(II)}_{k_1,k}(n)$, associated with the non-overlapping schemes for counting of scans, we have to employ a new technique. We define a sequence of stopping times in terms of the newly defined random variables in such a way that the number of scans between two successive stopping times become i.i.d. This allows us to approximate both $S^{(I)}_{k_1,k}(n)$ and $S^{(II)}_{k_1,k}(n)$, by a random sum of i.i.d. random variables. Using asymptotic results on random sums of i.i.d. random variables, we derive the appropriate central limit theorem for $S^{(I)}_{k_1,k}(n)$ and $S^{(II)}_{k_1,k}(n)$. This method can be generalized to a wider class of statistics, which depend on a finite number of observations and has a finite order auto-correlation structure.

Let $l = \max(m, 2k, k+2)$. If $l > m$, set $X_{-m} = X_{-m-1} = \cdots = X_{-l+1} = 0$. Define a sequence of random variables $\{Y_n : n \geq 0\}$ as follows:

$$Y_n = \sum_{j=0}^{l-1} 2^j X_{n-j}.$$

Since $X_i \in \{0,1\}$ for all i, Y_n assumes values in the set N_l. Since the random variables X_n is stationary and m-dependent ($m \leq l$), we have

$$P(Y_{n+1} = y \mid Y_n = x) = \begin{cases} p_{\pi_m(x)} & \text{if } \ y = f_1^{(l)}(x) \\ q_{\pi_m(x)} = 1 - p_{\pi_m(x)} & \text{if } \ y = f_0^{(l)}(x) \\ 0 & \text{otherwise} \end{cases}$$

where $\pi_m : N_l \to N_m$ defined by $\pi_m(x) = x \operatorname{mod}(2^m)$ and $f_i^{(l)}(x) = (2x+i)\operatorname{mod}(2^l)$, $i = 0, 1$. Thus, $\{Y_n : n \geq 0\}$ is a stationary Markov chain with state space N_l with $Y_0 = x$. Therefore, $\{Y_n : n \geq 0\}$ is an α-mixing sequence with $\alpha_n = K\rho^n$ where $K > 0$ and $0 < \rho < 1$ are constants (see Billingsley (1986)). More formally, let $\mathcal{F}_0^n$ and $\mathcal{F}_n^\infty$ be the σ-algebras generated by the random variables $(Y_0, Y_1, \ldots, Y_n)$ and $(Y_n, Y_{n+1}, \ldots)$ respectively. Then, we have for any $n \geq 1$ and $t \geq 1$

$$\sup_{A \in \mathcal{F}_0^n, B \in \mathcal{F}_{n+t}^\infty} |P(A)P(B) - P(A \cap B)| \leq K\rho^t. \tag{5.1}$$

From the definition of Y_n, it is clear that $X_n = 1$ if and only if Y_n is odd. Further, Y_n contains all the information of the window of length l, starting at $n-l+1$ and ending at n. We now prove the central limit theorem and law of the iterated logarithm for arbitrary functions of the sequence $\{Y_n : n \geq 1\}$.

THEOREM 5.1. *Let $v : N_l \to \mathbb{R}$ be any function. Then, we have,*

$$\frac{\sum_{i=1}^n v(Y_i) - E(\sum_{i=1}^n v(Y_i))}{\sqrt{n}\sigma} \Rightarrow Z$$

where $\sigma > 0$ and Z follows a standard normal distribution. Further, law of the iterated logarithm holds for $\{v(Y_i) : i \geq 1\}$, i.e.,

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^n v(Y_i) - E(\sum_{i=1}^n v(Y_i))}{\sigma\sqrt{2n \log\log n}} = 1 \ \textit{almost surely.}$$

PROOF. Using example 27.6 of Billingsley ((1986), p. 363), we obtain that $\{v(Y_i) : i \geq 0\}$ is a stationary sequence of random variables. Further, we also obtain that $\{v(Y_i) : i \geq 0\}$ is a mixing sequence with mixing constants α_n, given by $\alpha_n = K\rho^n$ where K and ρ are defined in (5.1). Therefore, we are able to apply the Theorem 27.4 of Billingsley to obtain the central limit theorem.

For the proof of the law of the iterated logarithm, we use Theorem 1.2.1 of Philipp (1971). Let $M_1 = \max\{|v(j)| : j \in N_l\}$. Since, N_l is a finite set, $M_1 < \infty$. Thus, we have, for any $n \geq 0$,

$$|v(Y_n) - E(v(Y_n))| \leq |v(Y_n)| + E(|v(Y_n)|) \leq 2M_1.$$

Thus, the condition of Philipp (1971) is satisfied by the family $\{v(Y_n) : n \geq 0\}$ and hence law of the iterated logarithm holds. □

Now, we use this meta theorem, with special choices of functions $v : N_l \to \mathbb{R}$, to derive the central limit theorems as well as law of the iterated logarithm for the enumerating variables $M_n^{(M)}$, $M_n^{(E)}$, $M_n^{(L)}$ and $S_{k_1,k}^{(III)}(n)$.

THEOREM 5.2. *Let $M_n^{(M)}$ be the number of runs of type at most k up to trial n. Then*

$$\frac{M_n^{(M)} - E(M_n^{(M)})}{\sqrt{n}\sigma_M} \Rightarrow Z$$

where $\sigma_M > 0$ and Z follows a standard normal distribution. Further, law of the iterated logarithm holds for $M_n^{(M)}$, i.e.,

$$\limsup_{n\to\infty} \frac{M_n^{(M)} - E(M_n^{(M)})}{\sigma_M\sqrt{2n\log\log n}} = 1 \text{ almost surely.}$$

PROOF. Consider the function $v_M : N_l \to \{0,1\}$ defined as

$$v_M(x) = \begin{cases} 1 & \text{if} \quad x \bmod(2^j) = 2^{j-1}+1 \quad \text{for some} \quad j = 2,3,\dots,k+2 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to note that $Y_n \bmod(2^j) = 2^{j-1}+1$ if and only if $X_n = X_{n-j+1} = 1$ and $X_i = 0$ for $n-j+2 \leq i < n$, $j = 2,3,\dots,k+2$. Thus, $v_M(Y_n) = 1$ if and only if a run of type at most k ends at time n for $n \geq k+2$. Therefore, we have $k+1+\sum_{j=k+2}^n v_M(Y_j) \geq M_n^{(M)} \geq \sum_{j=k+2}^n v_M(Y_j)$. Since $|\sum_{j=1}^n v_M(Y_j) - \sum_{j=k+2}^n v_M(Y_j)| \leq k+1$, we have the central limit theorem and law of the iterated logarithm from Theorem 5.1. □

Considering the function $v_E : N_l \to \{0,1\}$ defined by

$$v_E(x) = \begin{cases} 1 & \text{if} \quad x \bmod(2^{k+2}) = 2^{k+1}+1 \\ 0 & \text{otherwise} \end{cases}$$

we obtain the following limit theorem for the runs of type exactly k, proof of which we omit.

THEOREM 5.3. *Let $M_n^{(E)}$ be the number of runs of type exactly k up to time n. Then*

$$\frac{M_n^{(E)} - E(M_n^{(E)})}{\sqrt{n}\sigma_E} \Rightarrow Z$$

where $\sigma_E > 0$ and Z follows a standard normal distribution. Further, law of the iterated logarithm holds for $M_n^{(E)}$, i.e.,

$$\limsup_{n\to\infty} \frac{M_n^{(E)} - E(M_n^{(E)})}{\sigma_E\sqrt{2n\log\log n}} = 1 \text{ almost surely.}$$

Now, we concentrate on the scan statistics $S_{k_1,k}^{(III)}(n)$ where $1 \le k_1 \le k$. Define a family of functions $\{\pi_t : t = 0, 1, \dots, k\}$ on $N_l \to \mathbb{R}$ in the following way: $\pi_0(x) = 0$ and $\pi_t(x) = x \bmod(2^t)$ for $t \ge 1$. Now, define the function $v_S : N_l \to \{0, 1\}$ defined by

$$v_S(x) = \begin{cases} 1 & \text{if } \displaystyle\sum_{t=1}^{k}(\pi_t(x) - \pi_{t-1}(x))/2^{t-1} \ge k_1 \\ 0 & \text{otherwise.} \end{cases}$$

It should be noted that $v_S(Y_n) = 1$ if and only if the window of length k starting at $n - k + 1$ and ending at n, contains at least k_1 many successes, i.e., a scan is observed in the window of length k ending at n. Therefore, using this function, we derive the following limit theorem for $S_{k_1,k}^{(III)}(n)$ from Theorem 5.1.

THEOREM 5.4. *Let $S_{k_1,k}^{(III)}(n)$ be the number of scans up to time n, obtained by using the overlapping scheme of counting scans. Then*

$$\frac{S_{k_1,k}^{(III)}(n) - E(S_{k_1,k}^{(III)}(n))}{\sqrt{n}\sigma_S^{(III)}} \Rightarrow Z$$

where $\sigma_S^{(III)} > 0$ and Z follows a standard normal distribution. Further, law of the iterated logarithm holds for $S_{k_1,k}^{(III)}(n)$, i.e.,

$$\limsup_{n\to\infty} \frac{S_{k_1,k}^{(III)}(n) - E(S_{k_1,k}^{(III)}(n))}{\sigma_S^{(III)}\sqrt{2n\log\log n}} = 1 \text{ almost surely.}$$

Remark. The arguments above use actually the mixing nature of the underlying random variables and will continue to hold for general stationary mixing sequences with appropriate conditions on the mixing constants. It is also evident that any pattern or a family of patterns which are determined by the values of finitely many X_i's, can be similarly represented by a suitably constructed function and therefore, the asymptotic results in such cases can be similarly derived from Theorem 5.1.

Next, we concentrate on the runs of type at least k. This situation is different from the previous cases, since, in this case, there is no upper bound on how far the run, starting at time n, can go. So, to apply the above technique, we need to approximate this run by some pattern which is of finite length.

THEOREM 5.5. *Let $M_n^{(L)}$ be the number of runs of type at least k up to time n. Then*

$$\frac{M_n^{(L)} - E(M_n^{(L)})}{\sqrt{n}\sigma_L} \Rightarrow Z$$

where $\sigma_L > 0$ and Z follows a standard normal distribution. Further, law of the iterated logarithm holds for $M_n^{(L)}$, i.e.,

$$\limsup_{n\to\infty} \frac{M_n^{(L)} - E(M_n^{(L)})}{\sigma_L\sqrt{2n\log\log n}} = 1 \text{ almost surely.}$$

PROOF. In order to approximate runs of type at least k, we define the function $v_L : N_l \to \{0,1\}$ as follows:

$$v_L(x) = \begin{cases} 1 & \text{if} \quad x \bmod(2^{k+1}) = 2^k \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $v_L(Y_n) = 1$ if $X_{n-k} = 1$ and $X_{n-k+1} = \cdots = X_n = 0$. Thus, this function will count the number of occurrences of the event that a success is followed by at least k failures.

Suppose that $i_1 < i_2 < \cdots < i_s$ are the starting points of runs of type at least k, up to time n. Hence, we must have $X_{i_t} = 1$ and $X_{i_t+j} = 0$ for $j = 1, 2, \ldots, k$. Hence, $v_L(Y_{i_t+k}) = 1$ for each $t = 1, 2, \ldots, s$. Therefore, we must have,

$$M_n^{(L)} \le \sum_{j=k+1}^{n} v_L(Y_i).$$

Conversely, if $k + 1 \le j_1 < j_2 < \cdots < j_s \le n$ be the trials for which $v_L(Y_{j_t}) = 1$ for $t = 1, 2, \ldots, s$, then $j_t - k$ must also be starting point of a run of type at least k, for $t = 1, 2, \ldots, s - 1$. Therefore, we must have,

$$M_n^{(L)} \ge \sum_{j=k+1}^{n} v_L(Y_i) - 1.$$

Combining this with above, we have,

$$\left| M_n^{(L)} - \sum_{j=1}^{n} v_L(Y_i) \right| \le k + 2.$$

From Theorem 5.1 and the above estimate, the result follows. □

Now, we consider the random variables $S_{k_1,k}^{(I)}(n)$ and $S_{k_1,k}^{(II)}(n)$. It is obvious that $X_n = 1$ if and only if Y_n is odd. So, the random variables $\{R_{k_1,k}^{(I)}(n) : n \ge 1\}$ and

$\{R^{(II)}_{k_1,k}(n) : n \geq 1\}$ can be re-defined using Y_n's in the following way:

$$R^{(I)}_{k_1,k}(n) = \begin{cases} 1 & \text{if } \pi_1(Y_n) + \sum_{j=n-k+1}^{n-1} \pi_1(Y_j) \prod_{t=j}^{n-1} (1 - R^{(I)}_{k_1,k}(t)) \geq k_1 \\ 0 & \text{otherwise} \end{cases}$$

$$R^{(II)}_{k_1,k}(n) = \prod_{j=n-k+1}^{n-1} (1 - R^{(II)}_{k_1,k}(j)) 1_{\{\sum_{j=n-k+1}^{n} \pi_1(Y_j) \geq k_1\}}$$

where $\pi_1(x) = x \bmod(2)$.

Now, we consider a sequence of increasing stopping times $\{\tau_t : t \geq 0\}$ defined as $\tau_0 = 0$ and for $t \geq 1$, $\tau_t = \inf\{n > \tau_{t-1} : Y_n = 0\}$. Clearly, as $t \to \infty$, $\tau_t \uparrow \infty$. The Markov chain $\{Y_n : n \geq 0\}$ has finite state space N_l. Further it is irreducible, since $0 < p_x < 1$ for all $x \in N_m$. Hence $\tau_t < \infty$ almost surely. It is evident that the stopping times occur if and only if we observe a sequence of **0**'s of length l, in the original sequence of random variables. This breaks the auto-correlation structure of the occurrences of scan and as a result, the number of scans before the stopping time and after the stopping time behave independently. We make this formal in the next lemma using the strong Markov property.

Define,

$$U^{(I)}_t = \sum_{i=\tau_t+1}^{\tau_{t+1}} R^{(I)}_{k_1,k}(i)$$

$$U^{(II)}_t = \sum_{i=\tau_t+1}^{\tau_{t+1}} R^{(II)}_{k_1,k}(i)$$

for $t \geq 0$. We have

LEMMA 5.1. *For any initial condition x, we have,*

- *the random variables $\{U^{(I)}_t : t \geq 0\}$ are independent. Further, $\{U^{(I)}_t : t \geq 1\}$ are identically distributed,*
- *the random variables $\{U^{(II)}_t : t \geq 0\}$ are independent. Further, $\{U^{(II)}_t : t \geq 1\}$ are identically distributed.*

First, we prove the central limit theorems assuming Lemma 5.1.

THEOREM 5.6. *Let $S^{(I)}_{k_1,k}(n)$ and $S^{(II)}_{k_1,k}(n)$ be the number of scans up to time n using the non-overlapping schemes of counting scans. Then we have*

$$\text{(a)} \quad \frac{S^{(I)}_{k_1,k}(n) - E(S^{(I)}_{k_1,k}(n))}{\sqrt{n}\sigma^{(I)}_S} \Rightarrow Z$$

$$\text{(b)} \quad \frac{S^{(II)}_{k_1,k}(n) - E(S^{(II)}_{k_1,k}(n))}{\sqrt{n}\sigma^{(II)}_S} \Rightarrow Z$$

where $\sigma^{(I)}_S$, $\sigma^{(II)}_S > 0$ and Z follows a standard normal distribution.

PROOF. We prove part (a) of Theorem 5.6 only. The other can be proved similarly. Since $\{Y_n : n \geq 0\}$ is an irreducible Markov chain with finite state space, it has an unique stationary distribution. Let $(\theta_0, \theta_1, \ldots, \theta_{2^l-1})$ be the stationary distribution. Further, we have that $E(\tau_1) < \infty$. Define $W(n) = \inf\{t \geq 1 : \tau_t > n\}$. Since the Markov chain is positive recurrent,

$$\frac{W(n)}{n} \longrightarrow \frac{1}{\theta_0} > 0 \quad \text{almost surely as} \quad n \to \infty.$$

Let $(\sigma_0^{(I)})^2 = \mathrm{Var}(U_1^{(I)})$ and set $(\sigma_S^{(I)})^2 = (\sigma_0^{(I)})^2/\theta_0 > 0$.

Now, we have,

$$\frac{S_{k_1,k}^{(I)}(n) - E(S_{k_1,k}^{(I)}(n))}{\sqrt{n}} = \frac{U_0^{(I)} - E(U_0^{(I)})}{\sqrt{n}} + \frac{\sum_{i=1}^{W(n)}(U_i^{(I)} - E(U_i^{(I)}))}{\sqrt{n}} - \frac{\sum_{i=n+1}^{\tau_{W(n)}} R_{k_1,k}^{(I)}(i) - E(R_{k_1,k}^{(I)}(i))}{\sqrt{n}}.$$

Since $E(|U_0^{(I)} - E(U_0^{(I)})|) \leq 2E(U_0^{(I)}) \leq 2E(\tau_1) < \infty$,

$$\frac{U_0^{(I)} - E(U_0^{(I)})}{\sqrt{n}} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Also, $E(|\sum_{i=n+1}^{\tau_{W(n)}} R_{k_1,k}^{(I)}(i) - E(R_{k_1,k}^{(I)}(i))|) \leq 2E(\tau_2 - \tau_1) < \infty$; as a result,

$$\frac{\sum_{i=n+1}^{\tau_{W(n)}}(R_{k_1,k}^{(I)}(i) - E(R_{k_1,k}^{(I)}(i)))}{\sqrt{n}} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Since we have $W(n)/n \to 1/\theta_0$ almost surely as $n \to \infty$, using proposition 10.1 of Bhattacharya and Waymire (1990), for the i.i.d. sequence of random variables $\{U_t^{(I)} : t \geq 1\}$, we have

$$\frac{\sum_{i=1}^{W(n)}(U_i^{(I)} - E(U_i^{(I)}))}{\sqrt{W(n)}} \Rightarrow N(0, (\sigma_0^{(I)})^2).$$

Combining all these, we conclude that $(S_{k_1,k}^{(I)}(n) - E(S_{k_1,k}^{(I)}(n)))/\sqrt{n} \Rightarrow N(0, (\sigma_S^{(I)})^2)$. □

Finally, we prove the Lemma 5.1.

PROOF OF LEMMA 5.1. Since $\tau_t \uparrow \infty$ as $t \to \infty$, for any $i \geq 1$, we can find $t \geq 0$ such that $\tau_t < i \leq \tau_{t+1}$. Now, we define new sequences of random variables as follows: set $R^{(I)}(i) = 0$ for $1 \leq i \leq k_1 - 1$ and $R^{(II)}(i) = 0$ for $1 \leq i \leq k - 1$ and

$$R^{(I)}(i) = \begin{cases} 1 & \text{if} \quad \pi_1(Y_i) + \displaystyle\sum_{j=\max(i-k+1,\tau_t+1)}^{i-1} \pi_1(Y_j) \prod_{t=j}^{i-1}(1 - R^{(I)}(t)) \geq k_1 \\ 0 & \text{otherwise} \end{cases}$$

$$R^{(II)}(i) = \prod_{j=\max(i-k+1,\tau_t+1)}^{i-1} (1 - R^{(II)}(j)) 1_{\{\sum_{j=\max(i-k+1,\tau_t+1)}^{i} \pi_1(Y_j) \geq k_1\}}.$$

Also, define, for each $t \geq 0$,

$$V_t^{(I)} = \sum_{j=\tau_t+1}^{\tau_{t+1}} R^{(I)}(i) \quad \text{and} \quad V_t^{(II)} = \sum_{j=\tau_t+1}^{\tau_{t+1}} R^{(II)}(i).$$

From the definition of $V_j^{(I)}$ and $V_j^{(II)}$, for $j \leq t$, it is clear that both of them are determined by the process $\{Y_n : n \leq \tau_{t+1}\}$, i.e., $V_j^{(I)}$ and $V_j^{(II)}$ are both $\mathcal{F}_{\tau_{t+1}} = \sigma(Y_n : n \leq \tau_{t+1})$ measurable random variables.

Now, for $\tau_t < i$, the random variables $R^{(I)}(i)$ and $R^{(II)}(i)$ are measurable with respect to $\mathcal{F}_{\tau_t+} = \sigma(Y_n : n \geq \tau_t + 1)$. Indeed, $R^{(I)}(\tau_t + 1) = 1$ if and only if $\pi_1(Y_{\tau_t+1}) \geq k_1$; hence it is $\mathcal{F}_{\tau_t+}$ measurable. Assume that it is true for i. Since $R^{(I)}(i+1)$ is a function of $\{Y_j : i \geq j \geq \tau_t + 1\}$ and $\{R^{(I)}(j) : i \geq j \geq \tau_t + 1\}$ and by induction hypothesis, $R^{(I)}(j)$ is measurable w.r.t. $\mathcal{F}_{\tau_t+}$ for $i \geq j \geq \tau_t + 1$, we have that $R^{(I)}(i+1)$ is $\mathcal{F}_{\tau_t+}$ measurable. A similar argument holds for $R^{(II)}(i)$. Therefore, both $V_t^{(I)}$ and $V_t^{(II)}$, are $\mathcal{F}_{\tau_t+}$ measurable.

Since $\{Y_n : n \geq 0\}$ is a Markov chain with finite state space, it obeys the strong Markov property. Thus, the conditional distribution of the process $\{Y_j : j \geq \tau_{t+1}\}$, given the process up to time $\tau_{t+1}(\mathcal{F}_{\tau_{t+1}})$, using the strong Markov property, is same as that of $\{Y_n : n \geq 0\}$ with the initial condition $Y_0 = Y_{\tau_{t+1}} = 0$. Since $V_{t+1}^{(I)}$ is measurable with respect to $\mathcal{F}_{\tau_{t+1}+}$, for any $\Gamma \subseteq \mathbb{R}$ we must have Γ_1 such that

$$P_x(V_{t+1}^{(I)} \in \Gamma) = P_x((Y_{\tau_{t+1}+1}, Y_{\tau_{t+1}+2}, \ldots,) \in \Gamma_1).$$

Using the strong Markov property, we have

$$\begin{aligned} P_x(V_{t+1}^{(I)} \in \Gamma \mid \mathcal{F}_{\tau_{t+1}}) &= P_x((Y_{\tau_{t+1}+1}, Y_{\tau_{t+1}+2}, \ldots) \in \Gamma_1 \mid \mathcal{F}_{\tau_{t+1}}) \\ &= P_0((Y_1, Y_2, \ldots) \in \Gamma_1). \end{aligned} \tag{5.2}$$

Since the conditional distribution of $V_{t+1}^{(I)}$ given $\mathcal{F}_{\tau_{t+1}}$ does not depend on $\mathcal{F}_{\tau_{t+1}}$, $V_{t+1}^{(I)}$ is independent of $\mathcal{F}_{\tau_{t+1}}$. As a consequence, $V_{t+1}^{(I)}$ is independent of random variables which are measurable with respect to $\mathcal{F}_{\tau_{t+1}}$. Therefore, $V_{t+1}^{(I)}$ is independent of $\{V_j^{(I)} : 0 \leq j \leq t\}$. Further, from (5.2), we must have,

$$P_x(V_{t+1}^{(I)} \in \Gamma) = P_0((Y_1, Y_2, \ldots) \in \Gamma_1)$$

which is independent of t for $t \geq 0$. Hence $\{V_{t+1}^{(I)} : t \geq 0\}$ is a sequence of i.i.d. random variables. Same arguments can be carried out for $V_{t+1}^{(II)}$.

Only thing we need to show now is $R^{(I)}(i) = R_{k_1,k}^{(I)}(i)$ and $R^{(II)}(i) = R_{k_1,k}^{(II)}(i)$ for all $i \geq 1$. By definition, $R^{(I)}(i) = R_{k_1,k}^{(I)}(i) = 0$ for $i = 1, \ldots, k_1 - 1$ and $R^{(II)}(i) = R_{k_1,k}^{(II)}(i) = 0$ for $i = 1, \ldots, k-1$. Assume that, it is true for $i-1$. Fix $t \geq 0$ such that $\tau_t < i \leq \tau_{t+1}$. Now, if $\tau_t(\omega) \leq i - k + 1$, then $R^{(I)}(i) = 1$ if and only if $\sum_{j=i-k+1}^{i-1} \pi_1(Y_j) \prod_{s=j}^{i-1}(1 - R^{(I)}(s)) + \pi_1(Y_i) = \sum_{j=i-k+1}^{i-1} \pi_1(Y_j) \prod_{s=j}^{i-1}(1 - R_{k_1,k}^{(I)}(s)) + \pi_1(Y_i) \geq k_1$ (by induction hypothesis) and therefore, if and only if $R_{k_1,k}^{(I)}(i) = 1$. Similarly, $R^{(II)}(i) = \prod_{j=i-k+1}^{i-1}(1 - R^{(II)}(j)) 1_{\{\sum_{j=i-k+1}^{i} \pi_1(Y_j) \geq k_1\}} = \prod_{j=i-k+1}^{i-1}(1 - R_{k_1,k}^{(II)}(j)) 1_{\{\sum_{j=i-k+1}^{i} \pi_1(Y_j) \geq k_1\}} =$

$R^{(II)}_{k_1,k}(i)$. If $\tau_t(\omega) > i - k + 1$, note that $Y_{\tau_t} = 0$ which implies that $X_j = 0$, hence $\pi_1(Y_j) = 0$, for all $j = \tau_t - l + 1, \dots, \tau_t$. Therefore, $\sum_{j=i-k+1}^{i-1} \pi_1(Y_j) \prod_{s=j}^{i-1}(1 - R^{(I)}_{k_1,k}(s)) = \sum_{j=\tau_t+1}^{i-1} \pi_1(Y_j) \prod_{s=j}^{i-1}(1 - R^{(I)}_{k_1,k}(s)) = \sum_{j=\tau_t+1}^{i-1} \pi_1(Y_j) \prod_{s=j}^{i-1}(1 - R^{(I)}(s))$ which implies that $R^{(I)}_{k_1,k}(i) = R^{(I)}(i)$. Similarly, we have, $\sum_{j=i-k+1}^{i} \pi_1(Y_j) = \sum_{j=\tau_t+1}^{i} \pi_1(Y_j)$. Further, since $l \geq 2k$, for $j = \tau_t - k + 1, \dots, \tau_t$, the window of length k, starting at $j - k + 1$ and ending at j, contains no successes. Since $k_1 \geq 1$, we have $R^{(II)}_{k_1,k}(j) = 0$. Thus, we have $\prod_{j=i-k+1}^{i-1}(1 - R^{(II)}_{k_1,k}(j)) = \prod_{j=\tau_t+1}^{i-1}(1 - R^{(II)}_{k_1,k}(j)) = \prod_{j=\tau_t+1}^{i-1}(1 - R^{(II)}(j))$. Therefore, we conclude that $R^{(II)}(i) = R^{(II)}_{k_1,k}(i)$. This completes the proof of lemma. □

Acknowledgements

The authors would like to thank the referees for their useful suggestions and constructive comments.

REFERENCES

Aki, S. (1985). Discrete distributions of order k on a binary sequence, *Annals of the Institute of Statistical Mathematics*, **37**, 205–224.

Aki, S. (1992). Waiting time problems for a sequence of discrete random variables, *Annals of the Institute of Statistical Mathematics*, **44**, 363–378.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first k consecutive successes, *Annals of the Institute of Statistical Mathematics*, **47**, 225–235.

Aki, S., Balakrishnan, N. and Mohanty, S. G. (1996). Sooner or later waiting time problems and failure runs in higher order Markov dependent trials, *Annals of the Institute of Statistical Mathematics*, **48**, 773–787.

Balakrishnan, N. and Koutras, M. V. (2002). *Runs and Scans with Applications*, Wiley, New York.

Balasubramanian, K., Viveros, R. and Balakrishnan, N. (1993). Sooner or later waiting time problem for Markovian Bernoulli trials, *Statistics & Probability Letters*, **18**, 153–161.

Bhattacharya, R. N. and Waymire, E. C. (1990). *Stochastic Processes with Applications*, Wiley, New York.

Billingsley, P. (1986). *Probability and Measure*, 2nd ed., Wiley, New York.

Boutsikas, M. V. and Koutras, M. V. (2001). Compound Poisson approximation for sums of dependent random variables, *Probability and Statistical Models with Applications* (eds. Ch. A. Charalambides, M. V. Koutras and N. Balakrishnan), 63–86, Chapman and Hall, Florida.

Chadjiconstantinidis, S. and Koutras, M. V. (2001). Distributions of the numbers of failures and successes in a waiting time problem, *Annals of the Institute of Statistical Mathematics*, **53**, 576–598.

Chadjiconstantinidis, S., Antzoulakos, D. L. and Koutras, M. V. (2000). Joint distributions of successes, failures and patterns in enumeration problems, *Advances in Applied Probability*, **32**, 866–884.

Chen, J. and Glaz, J. (1999). Approximation for the distribution and the moments of discrete scan statistics, *Scan Statistics and Applications* (eds. J. Glaz and N. Balakrishnan), 27–66, Birkhauser, Boston.

Chung, K. L. (1974). *A Course in Probability Theory*, 2nd ed., Academic Press, London.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner or later waiting time problems for Bernoulli trials: Frequency and run quotas, *Statistics & Probability Letters*, **9**, 5–11.

Feller, W. (1968). *An Introduction to Probability Theory & Its Applications*, 3rd ed., Vol. I, Wiley, New York.

Fu, J. C. and Koutras, M. V. (1994). Distribution theory of runs: a Markov chain approach, *Journal of American Statistical Association*, **89**, 1050–1058.

Glaz, J. and Balakrishnan, N. (eds.) (1999). *Scan Statistics and Applications*, Birkhauser, Boston.

Glaz, J., Naus, J. and Wallenstein, S. (2001). *Scan Statistics*, Springer, New York.

Hirano, K. (1986). Some properties of the distributions of order k, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou and A. F. Horadam), 43–53, Reidel, Dordrecht.

Hirano, K. and Aki, S. (1993). On number of occurrences of success runs of specified length in a two-state Markov chain, *Statistica Sinica*, **3**, 311–320.

Koutras, M. V. (1996). On a waiting time distribution in a sequence of Bernoulli trials, *Annals of the Institute of Statistical Mathematics*, **48**, 789–806.

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach, *Annals of the Institute of Statistical Mathematics*, **47**, 743–766.

Ling, K. D. (1988). On binomial distributions of order k, *Statistics & Probability Letters*, **6**, 247–250.

Philipp, W. (1971). *Mixing Sequence of Random Variables and Probabilistic Number Theory*, Memoirs of the American Mathematical Society, A.M.S., Providence.

Philippou, A. N. (1986). Distributions and Fibonacci polynomials of order k, longest runs, and reliability of consecutive-k-out-of-n:F system, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou and A. F. Horadam), 203–227, Reidel, Dordrecht.

Philippou, A. N. and Makri, F. S. (1986). Successes, runs and longest runs, *Statistics & Probability Letters*, **4**, 101–105.

Sarkar, A. and Anuradha (2002). Asymptotic results for generalized runs in higher order Markov Chains (preprint).

Uchida, M. (1998). On number of occurrences of success runs of specified length in a higher-order two-state Markov chain, *Annals of the Institute of Statistical Mathematics*, **50**, 587–601.

Uchida, M. and Aki, S. (1995). Sooner or later waiting time problems in a two-state Markov chain, *Annals of the Institute of Statistical Mathematics*, **47**, 363–378.