

OSCILLATION PROPERTIES OF EXPECTED STOPPING TIMES AND STOPPING PROBABILITIES FOR PATTERNS CONSISTING OF CONSECUTIVE STATES IN MARKOV CHAINS

AZER KERIMOV AND ABDULLAH ÖNER

We investigate a Markov chain with a state space $1, 2, \dots, r$ stopping at appearance of patterns consisting of two consecutive states. It is observed that the expected stopping times of the chain have surprising oscillating dependencies on starting positions. Analogously, the stopping probabilities also have oscillating dependencies on terminal states. In a nonstopping Markov chain the frequencies of appearances of two consecutive states are found explicitly.

1. Formulation of results

Let $(Z_n)_{n \geq 1}$ be a time homogeneous Markov chain with finite state space Ω . A pattern is a finite sequence of elements from Ω . There has been and still is much interest in stopping probabilities and mean waiting times of pattern collections. First results were obtained for the patterns generated by independent and identically distributed discrete random variables. Stopping probabilities and mean waiting times in this case were obtained by Li [9], Gerber and Li [5] by use of a martingale approach and by Guibas and Odlyzko [7], Blom and Thorburn [1] by use of a combinatorial method. Later on, stopping probabilities and mean waiting times for the patterns generated by dependent random variables were investigated. Han and Aki [8], Fu and Chang [3] used the Markov chain embedding method, Han and Aki [8] used a combinatorial method, Glaz et al. [6], Pozdnyakov [11], Fisher and Cui [2] used a martingale approach, Glaz et al. [6], Gava and Salotti [4] used the method of gambling teams for developing a procedure for calculating stopping probabilities and mean waiting times for Markov chains.

We consider a time-homogeneous Markov chain $(Z_n)_{n \geq 1}$ on a finite set $\Omega(r) = \{1, 2, \dots, r\}$ with transition probabilities $p_{ii} = 0$ and $p_{ij} = 1/(r-1)$ for all $i, j \in \Omega(r)$. A pattern is a finite sequence of elements from $\Omega(r)$. A pattern of length n will be denoted by a_1, a_2, \dots, a_n , where $a_i \in \Omega(r)$. In this paper we will deal with patterns of length two. We say that $A_{u,v}^{\leq s} = u, v$ is an s pattern if $|u - v| \leq s$. Let $\tau_{A_{u,v}^{\leq s}}$ be the waiting time until $A_{u,v}^{\leq s}$ occurs. That is,

$$\tau_{A_{u,v}^{\leq s}} = \inf\{m \geq 1 : Z_m, Z_{m+1} = A_{u,v}^{\leq s} \text{ for some } u \text{ and } v\}.$$

Kerimov is the corresponding author.

2010 AMS *Mathematics subject classification*: 60J10, 05C81.

Keywords and phrases: Markov chain, pattern, expected stopping time, stopping probability.

Received by the editors on December 30, 2019, and in revised form on February 24, 2020.

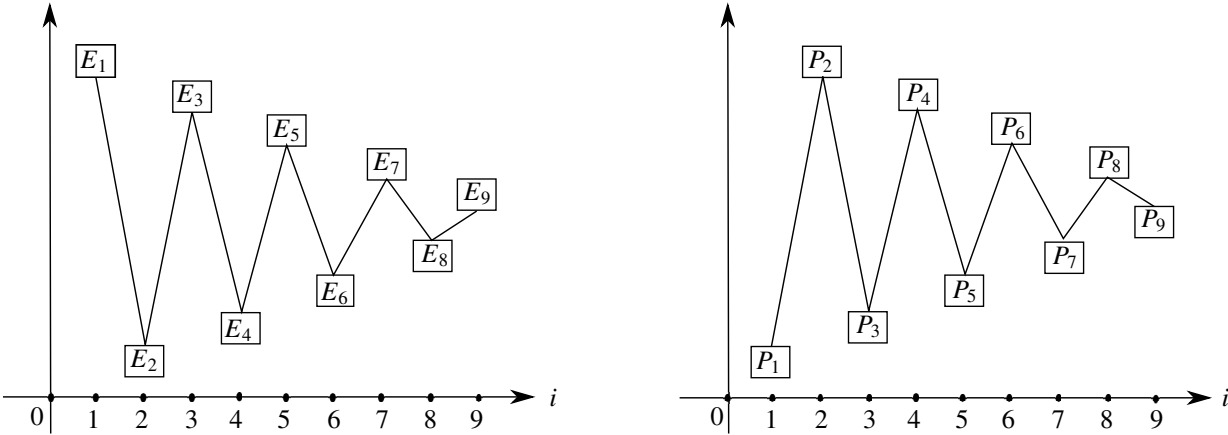


Figure 1. Oscillating hierarchy between expected stopping times and stopping probabilities.

The stopping time $\tau^{\leq s}$ of the Markov chain $(Z_n)_{n \geq 1}$ is the minimum number of steps required for getting one of the s patterns:

(1)
$$\tau^{\leq s} := \min_{(u,v)} \{ \tau_{A_{u,v}^{\leq s}} \}.$$

Let $E_i^{\leq s}$ be the expected stopping time of the chain represented by the conditional expectation

(2)
$$E_i^{\leq s} := E\{\tau^{\leq s} \mid Z_1 = i\}.$$

Below we focus on the case $s = 1$ and for simplicity use τ for $\tau^{\leq 1}$ and E_i for $E_i^{\leq 1}$. We will investigate the behavior of expected stopping times E_i as a function of i . We will only consider E_1, \dots, E_k for $r = 2k$ and E_1, \dots, E_{k+1} for $r = 2k + 1$, since by symmetry $E_i = E_{r-i+1}$ for $i = 1, \dots, \lfloor r/2 \rfloor$. It is easy to see that $E_1 > E_2$. One could expect monotonicity relationships $E_i > E_{i+1}$ for $i = 1, 2, \dots, \lfloor r/2 \rfloor$. On the contrary, it turns out that the smallest expectation is E_2 and there is the following oscillating hierarchy between expected stopping times:

Theorem 1. *The expected stopping times $E_i, i = 1, 2, \dots, r$ satisfy the following inequalities:*

$$E_1 > E_3 > \dots > E_{\lceil (r/2) - a - 2 \rceil} > E_{\lceil (r/2) - a \rceil} > E_{\lceil r/2 \rceil} > E_{\lceil (r/2) - b \rceil} > E_{\lceil (r/2) - b - 2 \rceil} > \dots > E_4 > E_2,$$

where $a = 1, b = 2$ if $r = 4l$ or $4l + 3$, and $a = 2, b = 1$ if $r = 4l + 1$ or $4l + 2$.

The dependence of expectations E_i on the starting states i in the case $r = 17$ is shown in Figure 1. Let $A_u := A_{u,v}^{\leq 1}$. It can be readily seen that there are $2r - 2$ stopping patterns of length two consisting of all consecutive elements of $\Omega(r)$: for each $u = 1, \dots, 2r - 2$ the pattern A_u is defined by

$$A_u = \begin{cases} u, u + 1 & \text{if } u = 1, \dots, r - 1, \\ u - r + 2, u - r + 1 & \text{if } u = r, \dots, 2r - 2. \end{cases}$$

If we admit the states 1 and r as neighbors and add the patterns 1, r and $r, 1$ to the collection of stopping patterns, then the stopping time of the obtained Markov chain $(Z_n^o)_{n \geq 1}$ is a geometric random variable and $E_1 = E_2 = \dots = E_r = (r - 1)/2$. Thus, $(Z_n)_{n \geq 1}$ is obtained from $(Z_n^o)_{n \geq 1}$ by symmetry breaking. Inherently the effect of this symmetry breaking diminishes as the starting point moves away from the

points 1 and r . [Theorem 1](#) shows that the natural decrease of dependence on the symmetry breaking points is not monotonic and moreover, surprisingly, oscillates.

Let $P\{\tau = \tau_{A_{j-1}} \mid Z_1 = i\}$ be the probability that the pattern $j - 1, j$ appears first among all stopping patterns in the realization of the chain starting at i . Then, under natural assumptions on uniform distribution of starting states, the probability of occurrence of $j - 1, j$ as the first pattern among all stopping patterns is

$$(3) \quad P\{\tau = \tau_{A_{j-1}}\} = \frac{1}{r} \sum_{i=1}^r P\{\tau = \tau_{A_{j-1}} \mid Z_1 = i\}.$$

Similarly, the probability of occurrence of $j + 1, j$ as the first pattern among all stopping patterns is

$$(4) \quad P\{\tau = \tau_{A_{j+r-1}}\} = \frac{1}{r} \sum_{i=1}^r P\{\tau = \tau_{A_{j+r-1}} \mid Z_1 = i\}.$$

For convenience, we set $P\{\tau = \tau_{A_0}\} = 0$. Hence by (3) and (4), the probability P_j that the chain terminates at state j satisfies

$$P_j = P\{\tau = \tau_{A_{j-1}}\} + P\{\tau = \tau_{A_{j+r-1}}\}$$

for each $1 \leq j \leq r$.

We will investigate the behavior of stopping probabilities P_j as a function of j . Since by symmetry $P_j = P_{r-j+1}$ for $j = 1, \dots, \lfloor r/2 \rfloor$, we will investigate P_1, \dots, P_k if $r = 2k$ and P_1, \dots, P_{k+1} if $r = 2k + 1$. One could expect natural monotonicity relationships $P_j < P_{j+1}$ for $j = 1, 2, \dots, \lfloor r/2 \rfloor$ between stopping probabilities. On the contrary, it turns out that P_2 is the greatest among all probabilities and there is the following oscillating hierarchy between stopping probabilities:

Theorem 2. *The stopping probabilities P_j , $j = 1, 2, \dots, r$ satisfy the following inequalities:*

$$P_1 < P_3 < \dots < P_{\lceil (r/2) - a - 2 \rceil} < P_{\lceil (r/2) - a \rceil} < P_{\lceil r/2 \rceil} < P_{\lceil (r/2) - b \rceil} < P_{\lceil (r/2) - b - 2 \rceil} < \dots < P_4 < P_2,$$

where $a = 1, b = 2$ if $r = 4l$ or $4l + 3$, and $a = 2, b = 1$ if $r = 4l + 1$ or $4l + 2$.

The dependence of probabilities P_j on the terminal states j in the case $r = 17$ is shown in [Figure 1](#). Thus, there is a duality relationship between expected stopping times and stopping probabilities: one can obtain orderings among stopping probabilities by simple reversing of orderings among expected stopping times.

The method of gambling teams and martingale approach provide a procedure for construction of equations and solving them to find the expectation of τ . By analogous approaches one can set up a technique for expressing each E_i , $i = 1, \dots, r$ as an entry of the solution and each P_j , $j = 1, \dots, r$ in terms of some entries of the solution for a certain linear system of equations; see [\[2\]](#), and also, for the independent case, see [\[9\]](#). Since this procedure does not provide explicit and simple expressions for the required expectations E_i and probabilities P_j , the comparison of the solutions of these equation systems becomes an extremely arduous task. In the proofs of [Theorems 1](#) and [2](#) we follow a totally different approach. We derive and explore a system of linear equations expressing dependencies of each expectation in terms of remaining expectations.

Consider a nonstopping Markov chain $(Z_n)_{n \geq 1}$ starting at state i . Let $\chi_m^i(j)$ be the total number of appearances of the patterns $j-1, j$ or $j+1, j$ among the first m trials. It turns out that as m increases, in spite of the ordering relationships between expected stopping times and stopping probabilities, the frequencies of appearances of these patterns have totally different behavior:

Theorem 3. *Let $i, j \in \Omega(r)$ be fixed and $\chi_m^i(j)$ be the total number of appearances of the patterns $j-1, j$ or $j+1, j$ in the first m trials of $(Z_n)_{n \geq 1}$. Then*

$$\lim_{m \rightarrow \infty} \frac{\chi_m^i(j)}{m} = \begin{cases} \frac{1}{2r-2} & \text{if } j \in \{1, r\}, \\ \frac{2}{2r-2} & \text{if } j \in \{2, \dots, r-1\}. \end{cases}$$

Let $G = (V, E)$ be an undirected connected graph with n vertices and m edges. It is well known that if a random walk jumps uniformly from any vertex to its adjacent vertices, its stationary distribution is

$$\pi(i) = \frac{d(i)}{2m},$$

where $d(i)$ is the degree of vertex i ; see [10]. We deal with a random walk jumping to all vertices of the complete graph with stopping rule (1) ($s = 1$) so we have a unique positive probability of stopping at each vertex i in the long run. Let us call the edges $A_j, j = 1, \dots, r-1$ the stopping edges. The stopping degree $d(i)$ of any vertex is the number of stopping edges incident to this vertex: $d(1) = 1$ and $d(i) = 2$ for $i = 2, \dots, \lceil r/2 \rceil$. Then, we may rewrite the stationary distribution in Theorem 3 as follows:

$$\pi(i) = \frac{d(i)}{2(r-1)}.$$

Thus, in this sense Theorem 3 is consistent with the formula $\pi(i) = d(i)/2m$ from [10].

2. Proofs

2.1. Proof of Theorem 1. In order to investigate the relationship between expected stopping times, we express each expectation E_i in terms of remaining expectations by using the properties of conditional expectation.

$$(5) \quad \begin{cases} E_1 = \frac{1}{r-1} + \sum_{j=3}^r \frac{1}{r-1} (1 + E_j) \\ E_2 = \frac{1}{r-1} + \frac{1}{r-1} + \sum_{j=4}^r \frac{1}{r-1} (1 + E_j) \\ E_i = \sum_{j=1}^{i-2} \frac{1}{r-1} (1 + E_j) + \frac{1}{r-1} + \frac{1}{r-1} + \sum_{j=i+2}^r \frac{1}{r-1} (1 + E_j) \quad \text{if } 3 \leq i \leq r-2, \\ E_{r-1} = \sum_{j=1}^{r-3} \frac{1}{r-1} (1 + E_j) + \frac{1}{r-1} + \frac{1}{r-1}, \\ E_r = \sum_{j=1}^{r-2} \frac{1}{r-1} (1 + E_j) + \frac{1}{r-1}. \end{cases}$$

Below we proceed with the odd case $r = 2k + 1$. Let $p := 1/(r - 1)$ and $\Delta_{i,j} := E_i - E_j$ for all i, j . As we will see in [Lemma 1](#), the mean stopping time E_i takes its maximum at $i = 1$:

Lemma 1. $\Delta_{1,j} > 0$ for $j = 2, \dots, k + 1$.

Proof. The proof of [Lemma 1](#) is based on the following trick. It directly follows from the system (5) that the expression $\Delta_{1,j} = E_1 - E_j$ depends on E_{j+1} and terms E_i , $2 \leq i \leq j - 1$. By iterative substitutions we show that the impact of all these terms results in a coefficient c_j of $\Delta_{1,j}$, which turns out to be positive.

Direct comparison of values of E_1 , E_2 and E_3 from (5) readily leads to

$$\frac{1}{p}\Delta_{1,2} = E_3 \quad \text{and} \quad \left(1 + \frac{1}{p}\right)\Delta_{1,3} = E_4.$$

Therefore, $\Delta_{1,2} > 0$ and $\Delta_{1,3} > 0$. Let $j \geq 4$. From the system of equations (5) we readily have

$$(6) \quad \left(1 + \frac{1}{p}\right)\Delta_{1,j} = E_{j+1} + \Delta_{j-1,2} \quad \text{for } j = 4, \dots, k + 1.$$

In (6), $\Delta_{1,j}$ is expressed in terms of E_{j+1} and $\Delta_{2,j-1}$. More generally, for all $i \geq 2$ by (5) one can express $\Delta_{i,j}$ hierarchically in terms of $\Delta_{i+1,j-1}$ and $\Delta_{i-1,j+1}$. Based on these formulas, below we will show that $\Delta_{i,j}$ is positive (negative) if and only if $\Delta_{i-1,j+1}$ is negative (positive). This observation is a main reason for the oscillation properties, as will be seen in (11) and (19).

Now in order to prove $\Delta_{1,j} > 0$ for the remaining cases $j \geq 4$, we consider the following two systems of equations depending on the parity of j :

For even $j \geq 4$ we have the following $j/2$ equations, where the first equation corresponds to $s = 1$ and coincides with (6) and the last equation corresponds to $s = j/2$:

$$(7) \quad \begin{cases} \left(1 + \frac{1}{p}\right)\Delta_{1,j} = E_{j+1} + \Delta_{j-1,2}, \\ \left(1 + \frac{1}{p}\right)\Delta_{s,t} = \Delta_{t+1,s-1} + \Delta_{t-1,s+1} \quad \text{for } s = 2, \dots, (j-2)/2 \text{ and } t = j+1-s, \\ \frac{1}{p}\Delta_{j/2,(j+2)/2} = \Delta_{(j+4)/2,(j-2)/2}. \end{cases}$$

For odd $j \geq 5$ we have the following $(j-1)/2$ equations, where the first equation corresponds to $s = 1$ and coincides with (6) and the last equation corresponds to $s = (j-1)/2$:

$$(8) \quad \begin{cases} \left(1 + \frac{1}{p}\right)\Delta_{1,j} = E_{j+1} + \Delta_{j-1,2}, \\ \left(1 + \frac{1}{p}\right)\Delta_{s,t} = \Delta_{t+1,s-1} + \Delta_{t-1,s+1} \quad \text{for } s = 2, \dots, (j-3)/2 \text{ and } t = j+1-s, \\ \left(1 + \frac{1}{p}\right)\Delta_{(j-1)/2,(j+3)/2} = \Delta_{(j+5)/2,(j-3)/2}. \end{cases}$$

Below, $[a_0; a_1, a_2, \dots, a_n]$ denotes a finite simple continued fraction

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}}$$

where a_i is an integer for $i \geq 0$. For even values of j consider the system (7). From the last equation corresponding to $s = j/2$ we express $\Delta_{j/2, (j/2)+1}$ in terms of $\Delta_{(j/2)+2, (j/2)-1}$ and insert it into the previous equation corresponding to $s = (i/2) - 1$. As a result, in the equation for $s = (i/2) - 1$ we get rid of $\Delta_{j/2, (j/2)+1}$. Similarly, in the equation corresponding to $s = (i/2) - 1$ we express $\Delta_{(j/2)-1, (j/2)+2}$ in terms of $\Delta_{(j/2)+3, (j/2)-2}$ and insert it into the previous equation corresponding to $s = (i/2) - 2$ and get rid of $\Delta_{(j/2)-1, (j/2)+2}$. We continue this process in (7) and finally insert the expression for $\Delta_{2, j-1}$ from the equation corresponding to $s = 2$ into the first equation corresponding to $s = 1$. We carry out the analogous process in (8) for odd values of j . As a result, we get the following identity:

$$(9) \quad c_j \Delta_{1, j} = E_{j+1},$$

where

$$c_j = \begin{cases} \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \frac{1}{p} \right] & \text{if } j \text{ is even,} \\ \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \left(1 + \frac{1}{p}\right) \right] & \text{if } j \text{ is odd.} \end{cases}$$

Here c_j has $n + 1 = j/2$ partial quotients for even j and has $n + 1 = (j - 1)/2$ partial quotients for odd j . It can be readily shown that the coefficient c_j is positive. Then, by (9) we get $\Delta_{1, j} > 0$ for $j = 4, \dots, k + 1$. Thus, Lemma 1 is proved. \square

In order to complete the proof of Theorem 2, we need the following:

Lemma 2. *Let $2 \leq i < j \leq k + 1$. Then $\Delta_{i, j} < 0$ if i is even, and $\Delta_{i, j} > 0$ if i is odd.*

Proof. The proof of Lemma 2 is also based on the following trick. It directly follows from the system (5) that the expression $\Delta_{s, t} = E_s - E_t$ depends on E_{s-1} , E_{t+1} and terms E_i , $s + 1 \leq i \leq t - 1$. By iterative substitutions we show that the impact of all these terms results in a coefficient $c_{s, t}$ or $d_{s, t}$ of $\Delta_{s, t}$, which turns out to be positive.

The proof consists of two parts: (a) $i + j \leq k + 2$, and (b) $k + 3 \leq i + j \leq 2k + 1$.

(a) We first consider the system (7). As in the proof of Lemma 1, starting from $s = \frac{1}{2}(j - 2)$ for $s = \frac{1}{2}(j - 2)$, $\frac{1}{2}(j - 2) - 1, \dots, 3$ we gradually express $\Delta_{s, t}$ in terms of $\Delta_{s-1, t+1}$ in the equation corresponding to s and insert it into the equation corresponding to $s - 1$. Similarly, in (8) starting from $s = \frac{1}{2}(j - 3)$ for $s = \frac{1}{2}(j - 3)$, $\frac{1}{2}(j - 3) - 1, \dots, 3$ we gradually express $\Delta_{s, t}$ in terms of $\Delta_{s-1, t+1}$ in the equation corresponding to s and insert it into the equation corresponding to $s - 1$. As a result, for $s = 2, \dots, \frac{1}{2}j$ (if j is even) and $s = 2, \dots, \frac{1}{2}(j - 1)$ (if j is odd) we get

$$(10) \quad c_{s, t} \Delta_{s, t} = \Delta_{t+1, s-1},$$

where

$$c_{s, t} = \begin{cases} \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \frac{1}{p} \right] & \text{if } s + t \text{ is odd,} \\ \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \left(1 + \frac{1}{p}\right) \right] & \text{if } s + t \text{ is even.} \end{cases}$$

Here $s + t = j + 1$ and $c_{s,t}$ has $\frac{1}{2}(t - s + 1)$ partial quotients for odd $s + t$ and $\frac{1}{2}(t - s)$ partial quotients for even $s + t$. As in the proof of [Lemma 1](#) the coefficient $c_{s,t}$ is positive. Therefore, by [\(10\)](#) we have

$$(11) \quad \Delta_{s,t} > 0 \ (\Delta_{s,t} < 0) \iff \Delta_{s-1,t+1} < 0 \ (\Delta_{s-1,t+1} > 0).$$

Using [Lemma 1](#) and [\(11\)](#), we readily get the following relationships for each $4 \leq j \leq k + 1$:

If j is even,

$$(12) \quad \Delta_{2,j-1} < 0 \iff \Delta_{3,j-2} > 0 \iff \Delta_{4,j-3} < 0 \iff \cdots \iff \Delta_{j/2,(j+2)/2} < 0 \ (\text{or } \Delta_{j/2,(j+2)/2} > 0)$$

for $j/2$ even or odd, respectively.

If j is odd,

$$(13) \quad \Delta_{2,j-1} < 0 \iff \Delta_{3,j-2} > 0 \iff \Delta_{4,j-3} < 0 \iff \cdots \iff \Delta_{(j-1)/2,(j+3)/2} < 0 \ (\text{or } \Delta_{(j-1)/2,(j+3)/2} > 0)$$

for $(j - 1)/2$ even or odd, respectively.

Thus, the signs of $\Delta_{i,j}$ are verified for $i + j \leq k + 2$.

(b) Now we will study the signs of $\Delta_{i,j}$ for $k + 3 \leq i + j \leq 2k + 1$. For each fixed $i \in \{2, \dots, k - 2\}$, depending on the parity of $i + k + 1$ we get the following two systems of equations:

For odd $i + k + 1$: we have

$$(14) \quad \begin{cases} \left(1 + \frac{1}{p}\right) \Delta_{i,k+1} = \Delta_{k+2,i-1} + \Delta_{k,i+1}, \\ \left(1 + \frac{1}{p}\right) \Delta_{s,t} = \Delta_{t+1,s-1} + \Delta_{t-1,s+1} \quad \text{for } s = i + 1, \dots, \frac{1}{2}(k + i - 2) \text{ and } t = i + k + 1 - s, \\ \frac{1}{p} \Delta_{\frac{1}{2}(k+i), \frac{1}{2}(k+i+2)} = \Delta_{\frac{1}{2}(k+i+4), \frac{1}{2}(k+i-2)}. \end{cases}$$

For even $i + k + 1$: we have

$$(15) \quad \begin{cases} \left(1 + \frac{1}{p}\right) \Delta_{i,k+1} = \Delta_{k+2,i-1} + \Delta_{k,i+1}, \\ \left(1 + \frac{1}{p}\right) \Delta_{s,t} = \Delta_{t+1,s-1} + \Delta_{t-1,s+1} \quad \text{for } s = i + 1, \dots, \frac{1}{2}(k + i - 3) \text{ and} \\ t = i + k + 1 - s \left(1 + \frac{1}{p}\right) \Delta_{\frac{1}{2}(k+i-1), \frac{1}{2}(k+i+3)} \\ = \Delta_{\frac{1}{2}(k+i+5), \frac{1}{2}(k+i-3)}. \end{cases}$$

Now, we take the difference of each E_i for $i \in \{k - 1, k\}$ with E_{k+1} from the system [\(5\)](#):

Note that for $i = k - 1$, we have

$$(16) \quad \left(1 + \frac{1}{p}\right) \Delta_{k-1,k+1} = \Delta_{k+2,k-2}$$

and for $i = k$,

$$(17) \quad \frac{1}{p} \Delta_{k,k+1} = \Delta_{k+2,k-1}.$$

We first consider the system (14). As in the proof of Lemma 1, starting from $s = \frac{1}{2}(k+i)$, for $s = \frac{1}{2}(k+i), \frac{1}{2}(k+i)-1, \dots, i+1$ we gradually express $\Delta_{s,t}$ in terms of $\Delta_{s-1,t+1}$ in the equation corresponding to s and insert it into the equation corresponding to $s-1$. Similarly, in (15) starting from $s = \frac{1}{2}(k+i-1)$, for $s = \frac{1}{2}(k+i-1), \frac{1}{2}(k+i-1)-1, \dots, i+1$ we gradually express $\Delta_{s,t}$ in terms of $\Delta_{s-1,t+1}$ in the equation corresponding to s and insert it into the equation corresponding to $s-1$. As a result, for $s = i, \dots, \frac{1}{2}(k+i)$ (if $i+k+1$ is odd) and for $s = i, \dots, \frac{1}{2}(k+i-1)$ (if $i+k+1$ is even), we get

$$(18) \quad d_{s,t} \Delta_{s,t} = \Delta_{t+1,s-1},$$

where

$$d_{s,t} = \begin{cases} \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \frac{1}{p} \right] & \text{if } s+t \text{ is odd,} \\ \left[1 + \frac{1}{p}; -\left(1 + \frac{1}{p}\right), \dots, (-1)^{n-1} \left(1 + \frac{1}{p}\right), (-1)^n \left(1 + \frac{1}{p}\right) \right] & \text{if } s+t \text{ is even.} \end{cases}$$

Here $s+t = i+k+1$ and $d_{s,t}$ has $\frac{1}{2}(t-s+1)$ partial quotients for odd $s+t$, and $\frac{1}{2}(t-s)$ partial quotients for even $s+t$. Since $d_{s,t} > 0$ by (16), (17) and (18), for $k+3 \leq s+t \leq 2k+1$ we have

$$(19) \quad \Delta_{s,t} > 0 \ (\Delta_{s,t} < 0) \iff \Delta_{s-1,t+1} < 0 \ (\Delta_{s-1,t+1} > 0).$$

Let us take $i = 2$ in (14) and (15) hence $s+t = k+3$. By symmetry, $\Delta_{1,k+2} = \Delta_{1,k}$ so $\Delta_{1,k+2} > 0$ by Lemma 1. Then, (19) gives

$$(20) \quad \begin{cases} \Delta_{1,k+2} > 0 \iff \Delta_{2,k+1} < 0 \iff \Delta_{3,k} > 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+4)} & \text{if } k+3 \text{ is odd,} \\ \Delta_{1,k+2} > 0 \iff \Delta_{2,k+1} < 0 \iff \Delta_{3,k} > 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+1), \frac{1}{2}(k+5)} & \text{if } k+3 \text{ is even,} \end{cases}$$

where the signs of $\Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+4)}$ and $\Delta_{\frac{1}{2}(k+1), \frac{1}{2}(k+5)}$ depend on the parity of their first indices:

$$\begin{cases} \text{the sign of } \Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+4)} \text{ is } (-1)^{\frac{1}{2}(k+2)-1}, \\ \text{the sign of } \Delta_{\frac{1}{2}(k+1), \frac{1}{2}(k+5)} \text{ is } (-1)^{\frac{1}{2}(k+1)-1}. \end{cases}$$

Similarly, we take $i = 3$ in (14) and (15) so $s+t = k+4$. By symmetry, $\Delta_{2,k+2} = \Delta_{2,k}$, hence $\Delta_{2,k+2} < 0$ by (12) and (13). Then, (19) gives

$$(21) \quad \begin{cases} \Delta_{2,k+2} < 0 \iff \Delta_{3,k+1} > 0 \iff \Delta_{4,k} < 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+5)} & \text{if } k+4 \text{ is odd,} \\ \Delta_{2,k+2} < 0 \iff \Delta_{3,k+1} > 0 \iff \Delta_{4,k} < 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+6)} & \text{if } k+4 \text{ is even,} \end{cases}$$

where the signs of $\Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+5)}$ and $\Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+6)}$ depend on the parity of their first indices:

$$\begin{cases} \text{the sign of } \Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+5)} \text{ is } (-1)^{\frac{1}{2}(k+3)-1}, \\ \text{the sign of } \Delta_{\frac{1}{2}(k+2), \frac{1}{2}(k+6)} \text{ is } (-1)^{\frac{1}{2}(k+2)-1}. \end{cases}$$

Let us now take $i = 4$ in (14) and (15) so $s+t = k+5$. By symmetry $\Delta_{3,k+2} = \Delta_{3,k}$, hence $\Delta_{3,k+2} > 0$ by (20). Then, (19) gives

$$(22) \quad \begin{cases} \Delta_{3,k+2} > 0 \iff \Delta_{4,k+1} < 0 \iff \Delta_{5,k} > 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+4), \frac{1}{2}(k+6)} & \text{if } k+5 \text{ is odd,} \\ \Delta_{3,k+2} > 0 \iff \Delta_{4,k+1} < 0 \iff \Delta_{5,k} > 0 \iff \dots \iff \Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+7)} & \text{if } k+5 \text{ is even,} \end{cases}$$

where the signs of $\Delta_{\frac{1}{2}(k+4), \frac{1}{2}(k+6)}$ and $\Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+7)}$ depend on the parity of their first indices:

$$\begin{cases} \text{the sign of } \Delta_{\frac{1}{2}(k+4), \frac{1}{2}(k+6)} \text{ is } (-1)^{\frac{1}{2}(k+4)-1}, \\ \text{the sign of } \Delta_{\frac{1}{2}(k+3), \frac{1}{2}(k+7)} \text{ is } (-1)^{\frac{1}{2}(k+3)-1}. \end{cases}$$

Thus, the signs of $\Delta_{s,t}$ are also verified for $k+3 \leq s+t \leq k+5$. The verification of the signs of $\Delta_{s,t}$ for $k+6 \leq s+t \leq 2k-1$ is absolutely analogous and will be omitted. The cases $s+t=2k$ and $s+t=2k+1$ are considerably different and need a different treatment:

For $i=k-1$ we get $s+t=2k$ in (16). By symmetry $\Delta_{k-2,k+2} = \Delta_{k-2,k}$, so $\Delta_{k-2,k+2} < 0$ if k is even and $\Delta_{k-2,k+2} > 0$ if k is odd by two steps before, which is the case $s+t=2k-2$.

Then, by (19) we have

$$(23) \quad \begin{cases} \Delta_{k-2,k+2} < 0 \iff \Delta_{k-1,k+1} > 0 & \text{if } k \text{ is even,} \\ \Delta_{k-2,k+2} > 0 \iff \Delta_{k-1,k+1} < 0 & \text{if } k \text{ is odd.} \end{cases}$$

Similarly, for $i=k$ we get $s+t=2k+1$ in (17). By symmetry $\Delta_{k-1,k+2} = \Delta_{k-1,k}$, hence $\Delta_{k-1,k+2} < 0$ if k is odd and $\Delta_{k-1,k+2} > 0$ if k is even by two steps before, which is the case $s+t=2k-1$.

Then, by (19) we have

$$(24) \quad \begin{cases} \Delta_{k-1,k+2} > 0 \iff \Delta_{k,k+1} < 0 & \text{if } k \text{ is even,} \\ \Delta_{k-1,k+2} < 0 \iff \Delta_{k,k+1} > 0 & \text{if } k \text{ is odd.} \end{cases}$$

Now the proof of Lemma 2 readily follows from (12), (13) and (20)–(24). \square

Since $\Delta_{i,j} = E_i - E_j$, for $r=2k+1$ Theorem 1 readily follows from Lemmas 1 and 2. The proof in the even case is absolutely analogous.

2.2. Proof of Theorem 2. Let $f_n(j)$ be the probability that the Markov chain stops at j in the n -th step, and let $g_n(j)$ be the probability that the Markov chain is at j in the n -th step but has not stopped yet. Then,

$$(25) \quad P_j = \sum_{n=1}^{\infty} f_n(j),$$

and the following recursion relations expresses the probabilities $\{f_n(j)\}$ in terms of $\{g_n(j)\}$:

$$(26) \quad \begin{cases} f_{n+1}(1) = \frac{1}{r-1} g_n(2), \\ f_{n+1}(j) = \frac{1}{r-1} g_n(j-1) + \frac{1}{r-1} g_n(j+1) & \text{if } 2 \leq j \leq r-1, \\ f_{n+1}(r) = \frac{1}{r-1} g_n(r-1). \end{cases}$$

Consider the following infinite sum:

$$(27) \quad G_j = \sum_{n=1}^{\infty} g_n(j),$$

which converges for any j due to (25) and (26). It turns out that G_j plays a crucial role in the relationship between oscillations of expected stopping times and stopping probabilities. By summing each equation of (26) over all positive values of n we get ($f_1(j) = 0$ for each j)

$$(28) \quad \begin{cases} P_1 = \frac{1}{r-1}G_2, \\ P_j = \frac{1}{r-1}G_{j-1} + \frac{1}{r-1}G_{j+1} & \text{if } 2 \leq j \leq r-1, \\ P_r = \frac{1}{r-1}G_{r-1}. \end{cases}$$

Let $f_{i,j}$ be the probability that a chain starting at i terminates at state j :

$$f_{i,j} := P\{Z_\tau = j \mid Z_1 = i\}.$$

The following lemma states that under the additional assumption that the chain does not stop in the first n steps, in spite of the natural inequality $f_{i,j} \neq f_{j,i}$ the transition probabilities from i to j and j to i coincide:

Lemma 3. *For each $i, j \in \Omega(r)$, we have*

$$P\{\tau > n, Z_n = j \mid Z_1 = i\} = P\{\tau > n, Z_n = i \mid Z_1 = j\}.$$

Proof. To each path $i \rightarrow k \rightarrow \dots \rightarrow l \rightarrow j$ of the Markov chain $\{Z_1 = i, Z_n = j, \tau > n\}$ corresponds the reversed path $j \rightarrow l \rightarrow \dots \rightarrow k \rightarrow i$ of the Markov chain $\{Z_1 = j, Z_n = i, \tau > n\}$. This correspondence is readily one-to-one and since the weight of each path is

$$\left(\frac{1}{r-1}\right)^{n-1} \frac{1}{r},$$

the proof is complete. □

We observe from (28) that in order to get a relation between stopping probabilities, we need to investigate the functions G_j for each j . By (27) and Lemma 3 we have

$$\begin{aligned} G_j &= \sum_{n=1}^{\infty} g_n(j) = \sum_{n=1}^{\infty} P\{\tau > n, Z_n = j\} = \sum_{n=1}^{\infty} \sum_{i=1}^r P\{\tau > n, Z_n = j \mid Z_1 = i\} P\{Z_1 = i\} \\ &= \frac{1}{r} \sum_{n=1}^{\infty} \sum_{i=1}^r P\{\tau > n, Z_n = j \mid Z_1 = i\} \\ &= \frac{1}{r} \sum_{n=1}^{\infty} \sum_{i=1}^r P\{\tau > n, Z_n = i \mid Z_1 = j\} \\ &= \frac{1}{r} \sum_{n=1}^{\infty} P\{\tau > n \mid Z_1 = j\} \\ &= \frac{1}{r} E\{\tau \mid Z_1 = j\}. \end{aligned}$$

The last equality is due to a well-known theorem on the expectation of nonnegative random variables. Therefore, we get

$$(29) \quad G_j = \frac{E_j}{r}, \quad j \in \Omega(r).$$

By plugging (29) into (28) we readily get

$$(30) \quad \begin{cases} P_1 = \frac{1}{r(r-1)} E_2, \\ P_j = \frac{1}{r(r-1)} E_{j-1} + \frac{1}{r(r-1)} E_{j+1} & \text{if } 2 \leq j \leq r-1, \\ P_r = \frac{1}{r(r-1)} E_{r-1}. \end{cases}$$

Let us set $E_0 = E_{r+1} = 0$. By (30) the value of the stopping probability P_j is exactly determined by the expected stopping times E_{j-1} and E_{j+1} . Now it can be readily seen that by (30), oscillation properties of stopping probabilities follow from oscillation properties of expected stopping times. Thus, the proof of Theorem 2 is complete.

2.3. Proof of Theorem 3. By the strong law of large numbers for Markov chains,

$$\Pr \left\{ \lim_{m \rightarrow \infty} \frac{\chi_m^i(j)}{m} = \pi_j \right\} = 1,$$

where π_j is the stationary probability of state j .

Let $f_{i,j}$ be the probability that the patterns $j-1, j$ or $j+1, j$ precede the remaining stopping patterns in the realization of the chain if it starts at i . We fix j and write the system of equations $f_{i,j}$ by conditioning on the first jump. Then, for $j = 1$, we get the following system of equations:

$$(31) \quad \begin{cases} f_{1,1} = \sum_{n=3}^r \frac{1}{r-1} f_{n,1}, \\ f_{2,1} = \frac{1}{r-1} + \sum_{n=4}^r \frac{1}{r-1} f_{n,1}, \\ f_{i,1} = \sum_{n=1}^{i-2} \frac{1}{r-1} f_{n,1} + \sum_{n=i+2}^r \frac{1}{r-1} f_{n,1} & \text{if } 3 \leq i \leq r-2, \\ f_{i,1} = \sum_{n=1}^{i-2} \frac{1}{r-1} f_{n,1} & \text{if } r-1 \leq i \leq r. \end{cases}$$

Side-by-side summation of (31) yields

$$(32) \quad f_{1,1} + f_{r,1} + 2 \sum_{i=2}^{r-1} f_{i,1} = 1.$$

It can be shown that for $j \geq 2$ the probabilities $f_{i,j}$ satisfy an analogous equation

$$(33) \quad f_{1,j} + f_{r,j} + 2 \sum_{i=2}^{r-1} f_{i,j} = 2, \quad j \geq 2.$$

The $r \times r$ matrix $P = \|f_{i,j}\|$ with entries $f_{i,j}$ is a stochastic matrix for the nonstopping Markov chain $(Z_n)_{n \geq 1}$ with the phase space $\Omega(r)$. Let us check that

$$\pi = (\pi_1, \pi_2, \dots, \pi_{r-1}, \pi_r) = \left(\frac{1}{2r-2}, \frac{2}{2r-2}, \dots, \frac{2}{2r-2}, \frac{1}{2r-2} \right)$$

satisfies $\pi = \pi P$. Indeed, by using (32) it can be shown that

$$\pi_1 = \frac{1}{2r-2}(f_{1,1} + f_{r,1}) + \frac{2}{2r-2} \sum_{j=2}^{r-1} (f_{2,1} + \dots + f_{r-1,1}) = \frac{1}{2r-2}.$$

Similarly, for $j \in \{2, \dots, \lceil r/2 \rceil\}$ by using (33) it can be shown that

$$\pi_j = \frac{1}{2r-2}(f_{1,j} + f_{r,j}) + \frac{2}{2r-2} \sum_{j=2}^{r-1} (f_{2,j} + \dots + f_{r-1,j}) = \frac{2}{2r-2}.$$

The proof of Theorem 3 is complete.

3. Concluding remarks

Let $\tau^{\leq s}$ be the stopping time of the Markov chain $(Z_n)_{n \geq 1}$ and let $E_i^{\leq s}$ be the expectation of $\tau^{\leq s}$ for a given starting state i as defined in (1) and (2), respectively. Numerical results show that for these stopping rules the effect of symmetry-breaking on the mean stopping times and stopping probabilities also decreases according to similar but more sophisticated oscillation rules.

For certainty we consider the odd case $r = 2k + 1$. Let $k + 1 = sp + q$. Let us distribute $k + 1$ expectations into $p + 1$ groups:

$$\begin{aligned} H_1 &= \{E_1^{\leq s}, \dots, E_s^{\leq s}\}, \\ H_2 &= \{E_{s+1}^{\leq s}, \dots, E_{2s}^{\leq s}\}, \\ &\vdots \\ H_p &= \{E_{(p-1)s+1}^{\leq s}, \dots, E_{ps}^{\leq s}\}, \\ H_{p+1} &= \{E_{ps+1}^{\leq s}, \dots, E_{k+1}^{\leq s}\}. \end{aligned}$$

The inequality $H_i \gg H_j$ means that $E' > E''$ for any two expectations $E' \in H_i$ and $E'' \in H_j$. Numerical studies suggest that $H_i, i = 1, \dots, p + 1$ satisfy Theorem 1 and expectations inside each H_k are monotonically increasing and decreasing for even and odd values of k , respectively. Similarly, if we divide stopping probabilities into groups of size s , these groups satisfy Theorem 2 and there is a monotonicity inside each group.

Let $(Z_n)_{n \geq 1}$ be a homogeneous Markov chain where the state space is the set of vertices of some regular graph G . Let the chain stop at step l if Z_l and Z_{l-1} are the vertices of the same edge. Most likely when one or several edges of G will be removed then the effect of this symmetry-breaking on the mean stopping times also will decrease according to some sophisticated oscillation rules.

Acknowledgements

The authors are deeply grateful to the referee for valuable suggestions.

References

- [1] G. Blom and D. Thorburn, “How many random digits are required until given sequences are obtained?”, *J. Appl. Probab.* **19**:3 (1982), 518–531.
- [2] E. Fisher and S. Cui, “Patterns generated by m th-order Markov chains”, *Statist. Probab. Lett.* **80**:15-16 (2010), 1157–1166.
- [3] J. C. Fu and Y. M. Chang, “On probability generating functions for waiting time distributions of compound patterns in a sequence of multistate trials”, *J. Appl. Probab.* **39**:1 (2002), 70–80.
- [4] R. J. Gava and D. Salotti, “Stopping probabilities for patterns in Markov chains”, *J. Appl. Probab.* **51**:1 (2014), 287–292.
- [5] H. U. Gerber and S.-Y. R. Li, “The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain”, *Stochastic Process. Appl.* **11**:1 (1981), 101–108.
- [6] J. Glaz, M. Kulldorff, V. Pozdnyakov, and J. M. Steele, “Gambling teams and waiting times for patterns in two-state Markov chains”, *J. Appl. Probab.* **43**:1 (2006), 127–140.
- [7] L. J. Guibas and A. M. Odlyzko, “String overlaps, pattern matching, and nontransitive games”, *J. Combin. Theory Ser. A* **30**:2 (1981), 183–208.
- [8] Q. Han and S. Aki, “Waiting time problems in a two-state Markov chain”, *Ann. Inst. Statist. Math.* **52**:4 (2000), 778–789.
- [9] S.-Y. R. Li, “A martingale approach to the study of occurrence of sequence patterns in repeated experiments”, *Ann. Probab.* **8**:6 (1980), 1171–1176.
- [10] L. Lovász, “Random walks on graphs: a survey”, pp. 353–397 in *Combinatorics, Paul Erdős is eighty* (Keszthely, 1993), Bolyai Soc. Math. Stud. **2**, János Bolyai Math. Soc., Budapest, 1996.
- [11] V. Pozdnyakov, “On occurrence of patterns in Markov chains: method of gambling teams”, *Statist. Probab. Lett.* **78**:16 (2008), 2762–2767.

AZER KERIMOV: kerimov@fen.bilkent.edu.tr

Department of Mathematics, Bilkent University, Bilkent, Ankara, Turkey

ABDULLAH ÖNER: abdullah.oner@fen.bilkent.edu.tr

Department of Mathematics, Bilkent University, Bilkent, Ankara, Turkey