Waiting time distributions for pattern occurrence in a constrained sequence
Valeri T. Stefanov, Wojciech Szpankowski

To cite this version:

HAL Id: hal-00966498
https://hal.inria.fr/hal-00966498
Submitted on 26 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Waiting Time Distributions for Pattern Occurrence in a Constrained Sequence

Valeri T. Stefanov\textsuperscript{1} and Wojciech Szpankowski\textsuperscript{2}†

\textsuperscript{1} School of Mathematics and Statistics, The University of Western Australia, Crawley (Perth) 6009, Australia.
\textsuperscript{2} Department of Computer Science, Purdue University, W. Lafayette, IN 47907, U.S.A.
\textsuperscript{1} stefanov@maths.uwa.edu.au
\textsuperscript{2} spa@cs.purdue.edu

received 15\textsuperscript{th} February 2007, revised 30\textsuperscript{nd} October 2007, accepted 5\textsuperscript{th} November 2007.

A binary sequence of zeros and ones is called a \((d, k)\)-sequence if it does not contain runs of zeros of length either less than \(d\) or greater than \(k\), where \(d\) and \(k\) are arbitrary, but fixed, non-negative integers and \(d < k\). Such sequences find an abundance of applications in communications, in particular for magnetic and optical recording. Occasionally, one requires that \((d, k)\)-sequences do not contain a specific pattern \(w\). Therefore, distribution results concerning pattern occurrence in \((d, k)\)-sequences are of interest. In this paper we study the distribution of the waiting time until the \(r\)-th occurrence of a pattern \(w\) in a random \((d, k)\)-sequence generated by a Markov source. Numerical examples are also provided.

\textbf{Keywords:} pattern; probability generating function; \((d, k)\)-sequence

1 Introduction

In many communication systems, including magnetic and optical recording ones, one must restrict the structure of a bit stream (binary sequence) to a class of sequences satisfying certain constraints. The simplest constrained binary sequences are those in which runs of zeros (between two consecutive 1’s) must have length at least \(d\) and at most \(k\), where \(d < k\). Such sequences are called \((d, k)\)-sequences (cf. [18, 19, 32]). For example, in \((1, 4)\)-sequence 11 and 00000 are forbidden runs. In some situations, as observed in [20], one needs to avoid certain patterns in \((d, k)\)-sequences. In this paper, for a given pattern (word) \(w\) \((w = w_1w_2\ldots w_m)\) we study the exact distribution of the waiting time until the \(r\)-th occurrence of the pattern \(w\) in a random \((d, k)\)-sequence generated by a Markov source.

Pattern matching is a well studied problem. It is motivated by applications in communication theory as well as computational biology where one looks for over-represented or under-represented patterns in order to find useful signals. In general, for a given set of patterns \(W = \{W_1,\ldots, W_K\}\), where the \(W_i\) are words of the same length, one searches for all \(W\) occurrences in a \textit{text} of length \(n\). (In this

\textsuperscript{†}The work of this author was supported in part by the NSF Grants CCR-0208709, CCF-0513636, DMS-0503742, and NIH Grant R01 GM068959-01.

1365–8050 © 2007 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France
paper we only consider a single pattern of length \( m \) that we denote by \( w \). In computer science literature several fast algorithms (e.g., Knuth-Morris-Pratt and Boyer-Moore algorithms) were designed to search for such patterns. Here, we are rather interested in the distribution theory associated with the number of \( w \) occurrences in a probabilistic framework where the (constrained) text is generated randomly (a Markov source in our case).

The pattern matching problem (in a probabilistic framework) goes back, at least, to Feller. The number of word occurrences in a random text has been intensively studied over the last two decades, with significant progress in this area being reported [3, 4, 5, 6, 7, 10, 11, 13, 14, 15, 18, 21, 24, 25, 26, 27, 28, 29, 31]. For instance, Guibas and Odlyzko [14] revealed the fundamental role played by autocorrelation sets and their associated polynomials. Li [15] and Gerber and Li [13] introduced martingale techniques to the area and combined the latter with a relevant Markov chain embedding. Markov chain embeddings have been widely used by a number of authors (see [6, 10, 11] and the references in [2, 12]). Blom and Thorburn [5] made connections with Markov renewal theory and Biggins and Cannings [4] elaborated on these. Stefanov and Pakes [29] introduced exponential family methodology to the area and Stefanov [27] extended it in combination with suitable Markov renewal embeddings. Régnier and Szpankowski [22, 23] established that the number of occurrences of a word is asymptotically normal under a diversity of models that include Markov chains. Nicodème, Salvy, and Flajolet [21] showed generally that the number of places in a random text at which a 'motif' (i.e., a general regular expression pattern) terminates is asymptotically normally distributed. Bender and Kochman [3] studied a generalized pattern \( W \) occurrences using (in a nutshell) the deBruijn graph representation that allowed the authors to establish the central limit theorem, but without explicit mean and variance. Recent surveys on pattern matching can be found in Lothaire [18] (Chaps. 6 and 7). To the best of our knowledge, the distribution theory associated with pattern occurrence in a constrained sequence, such as a \((d, k)\)-sequence, has not been treated in the literature.

A brief description of our problem and methodology follows. Let \( N_n \) be the number of \( w \) \((w = w_1 \ldots w_m)\) occurrences in a binary sequence, of length \( n \), generated by a two-state Markov chain \( X \). Throughout the paper such sequences will be called \textit{unconstrained sequences} whereas \((d, k)\)-sequences will be called \textit{constrained sequences}. By \( Y_r \) we define the waiting time until the \( r \)-th occurrence of the pattern \( w \) in an unconstrained sequence. Bearing in mind that the initial symbol at time zero counts to the sequence length we have

\[
P(N_n \geq r) = P(Y_r \leq n - 1) \quad (1)
\]

for all \( r, n \geq 1 \). This basic renewal equation is the starting point of two different approaches to the analysis of pattern occurrences, on finite alphabets, in unconstrained sequences as surveyed in Chaps. 6 and 7 of [18]. For example, [3, 14, 21, 22, 23] analyze \( N_n \), whereas the authors of [24, 28, 31] study the waiting time \( Y_r \), for unconstrained sequences. In the case of constrained sequences we may be interested in either the distribution of \( N_n \) given the sequence is constrained up to time \( n \), or the distribution of \( Y_r \) given the sequence is constrained up to time \( Y_r \). In other words, denoting by \( \tilde{N}^{(d,k)}_n \) the number of runs of zero of length either \textit{less} than \( d \) or \textit{greater} than \( k \) in an unconstrained sequence of length \( n \), the probabilities of interests are \( P(N_n \geq r|N^{(d,k)}_n = 0) = P(Y_r \leq n - 1|N^{(d,k)}_n = 0) \) and \( P(Y_r \leq n - 1|\tilde{N}^{(d,k)}_n = 0) \) \((= P(N_n \geq r|\tilde{N}^{(d,k)}_n = 0))\). Clearly \( P(N_n \geq r|\tilde{N}^{(d,k)}_n = 0) \) is not equal to \( P(Y_r \leq n - 1|\tilde{N}^{(d,k)}_n = 0) \). Also the evaluation of each of these two conditional probabilities lead to two different problems. For the
latter probability we also have

\[ P(Y_r \leq n - 1 | \bar{N}_r^{(d,k)} = 0) = \frac{\sum_{i=0}^{n-1} P(Y_r = i, \bar{N}_i^{(d,k)} = 0)}{\sum_{i=0}^{\infty} P(Y_r = i, \bar{N}_i^{(d,k)} = 0)}. \]

In the present paper we deal with \( P(Y_r \leq n - 1 | \bar{N}_r^{(d,k)} = 0) \), whereas in a forthcoming paper we will treat \( P(N_n = \bar{r} | \bar{N}_r^{(d,k)} = 0) \). Of course, the latter probability is of relevance in situations when the constrained sequence has been observed up to time \( n \) whereas the former is such when the constrained sequence is observed up to an \( r \)-th occurrence of the pattern of interest.

Stefanov [28] provides an original approach for a recursive evaluation of the generating functions of the waiting time conditioned on seeing a portion of the pattern in an unconstrained sequence. Also the approach provides the joint generating functions of the aforementioned waiting time \( Y_r \) together with the associated counts of relevant events. This paper extends the analysis of [28] to constrained sequences. The case of constrained sequences, when the probability of interest is \( P(Y_r \leq n - 1 | \bar{N}_r^{(d,k)} = 0) \), leads to more general type of events, associated with the above waiting time, than those considered in [28]. The key points of that extension are explained in Idea of the Proof inserted immediately after Theorem 1 in the next section.

The paper is organized as follows. In the next section we present our main theoretical results. These provide recursive formulae for computing the joint generating function of the waiting time until seeing the \( r \)-th occurrence of a pattern and the associated count of runs of zero of length either less than \( d \) or greater than \( k \) (the so called forbidden patterns). In the last section we provide numerical examples.

## 2 Main Results

We assume that the binary sequences are generated by a two-state Markov chain \( X = (X(n), \ n = 0, 1, \ldots) \). Its transition probabilities are denoted by

\[ P(X(k) = j | X(k - 1) = i) = p_{i,j}, \quad i, j = 0, 1. \]

Recall that \( \bar{N}_n^{(d,k)} \) counts the number of the so called forbidden patterns up to time \( n \). Denote by

\[ G_{Y_r^{(s)}}(z_1, z_2) = E[z_1^{Y_r^{(s)}} z_2^{Y_r^{(s)}}] \]

the joint generating function of the waiting time, \( Y_r \), until seeing the \( r \)-th occurrence of the pattern \( w = w_1 w_2 \ldots w_m \), given the initial symbol of the sequence is \( s \), and the associated count, \( \bar{N}_r^{(d,k)} \), of occurrences of runs of zeros of length either less than \( d \) or greater than \( k \) up to that waiting time.

Note that if \( G_{Z_1, Z_2}(z_1, z_2) \) is the joint generating function of two nonnegative integer random variables, then for the generating function \( G_{Z_1 | Z_2 = 0}(z) \) of the conditional distribution of \( Z_1 \), given \( Z_2 = 0 \), we have

\[ G_{Z_1 | Z_2 = 0}(z) = \frac{\sum_{n=0}^{\infty} z^n P(Z_1 = n, Z_2 = 0)}{P(Z_2 = 0)} = \frac{G_{Z_1, Z_2}(z, 0)}{G_{Z_1, Z_2}(1, 0)}. \]
Therefore, the joint generating function $G_{Y_1(s)}(z_1, z_2)$ renders the generating function of the conditional distribution of $Y_r$ given $\hat{N}_{Y_1(s)}^{(d,k)} = 0$. Further assume that $Y_1^{(s)}, Y_2, Y_3, \ldots$ are independent random variables such that $Y_2, Y_3, \ldots$ are identically distributed with the following distributions. The distribution of $Y_1^{(s)}$ is equal to the conditional distribution of the waiting time $Y_1$ to see for the first time the pattern $w$, if the starting symbol is $s$ ($s = 0$ or $1$) and given no forbidden pattern has occurred up to time $Y_1$, that is, given $\hat{N}_{Y_1(s)}^{(d,k)} = 0$. The distribution of $Y_2$ is equal to the conditional distribution of the intersite distance between two consecutive occurrences of the pattern $w$ given no forbidden pattern has occurred within that intersite distance. Then, in view of the strong Markov property, the conditional distribution of $Y_r$ given $\hat{N}_{Y_1(s)}^{(d,k)} = 0$, is equal to that of

$$Y_1^{(s)} + \sum_{i \geq 2} Y_i$$

and of course its generating function equals $G_{Y_1(s)}(z)(G_{Y_2}(z))^{-1}$, where

$$G_{Y_1(s)}(z) = \frac{G_m(z, 0)}{G_m(1, 0)} \quad G_{Y_2}(z) = \frac{G_{m_{\text{intersite}}}(z, 0)}{G_{m_{\text{intersite}}}(1, 0)}$$

and by $G_m(z_1, z_2)$ we denoted the joint generating function of the waiting time, $Y_1$, until the first occurrence of the pattern $w = w_1w_2 \ldots w_m$ with initial symbol $s$, and the associated count, $\hat{N}_{Y_1(s)}^{(d,k)}$, of occurrences of the so-called forbidden patterns, and with $G_{m_{\text{intersite}}}(z_1, z_2)$ we denoted the joint generating function of the intersite distance between two consecutive occurrences of the pattern $w$ and the associated count of occurrences of the forbidden patterns.

The remaining part of the paper is devoted to a method for an explicit derivation of the joint generating functions $G_m(z_1, z_2)$ and $G_{m_{\text{intersite}}}(z_1, z_2)$. Recall that these generating functions are associated with unconstrained sequences.

Let $\nu_{i,j}$ be the transition time from state $i$ to state $j$ in the two-state Markov chain $X$ introduced earlier, that is,

$$\nu_{i,j} = \inf\{n : X(n) = j | X(0) = i\}, \quad i, j = 0, 1,$$

and let $I_{\nu_{i,j}}$ be the associated indicator function of the event "a run of zeros of length either less than $d$ or greater than $k$ has occurred during the transition time $\nu_{i,j}$." It is assumed that $\nu_{i,i} = 0$.

Introduce the functions $g_{i,j}(z_1, z_2)$ for $i, j = 0, 1$, as follows: Let

$$g_{i,j}(z_1, z_2) = \mathbb{E}[z_1^{\nu_{i,j}} z_2^{\nu_{i,j}}], \quad (i, j) \neq (1, 0)$$

be the joint generating function of $(\nu_{i,j}, I_{\nu_{i,j}})$, if $i, j = 0, 1$ and $(i, j) \neq (1, 0)$. Of course

$$g_{0,0}(z_1, z_2) = g_{1,1}(z_1, z_2) = 1,$$

because $\nu_{0,0} = \nu_{1,1} = 0$, and subsequently $I_{\nu_{0,0}} = I_{\nu_{1,1}} = 0$. We define $g_{1,0}(z_1, z_2)$ to be the generating function $G_{\nu_{1,0}}(z_1)$ of $\nu_{1,0}$, that is,

$$g_{1,0}(z_1, z_2) = G_{\nu_{1,0}}(z_1) = \frac{z_1 p_{1,0}}{1 - p_{1,1} z_1}$$

(2)
The second identity comes from noticing that \( \nu_{1,0} \) is a geometrically distributed random variable with probability of success \( p_{1,0} \) and support \( \{1, 2, \ldots\} \). In other words, the meaning of \( g_{1,0}(z_1, z_2) \) is the same as that for the other \( g_{i,j}(z_1, z_2) \) with the convention that \( I_{\nu_{1,0}} = 0 \). The reasons for defining \( g_{1,0}(z_1, z_2) \) as if ignoring a possible occurrence of the event of interest within the passage time \( \nu_{1,0} \) will become clear in the proof of Theorem 1 below.

Another generating function of relevance is the joint generating function \( (0, I_{\nu_{0,1}}) \), given that exactly \( r \) zeros are preceding the starting state zero; it is assumed that these zeros are allowed to be counted towards the formation of the event marked by the indicator function \( I_{\nu_{0,1}} \). Denote this joint generating function by \( g_{r-0,1}(z_1, z_2) \). Clearly, \( g_{r-0,1}(z_1, z_2) \) equals the joint generating function of \( \nu_{0,1} \) and the indicator function of the event "a run of zeros of length either less than \( d - r \) or greater than \( k - r \) has occurred within that transition time".

**Lemma 1** The following explicit expressions hold for the joint generating functions \( g_{0,1}(z_1, z_2) \) and \( g_{r-0,1}(z_1, z_2) \):

\[
g_{0,1}(z_1, z_2) = p_{0,1}z_1 \left( \frac{z_2}{1 - p_{0,0}z_1} + (1 - z_2) \sum_{i=\max(1,d)}^{k} (p_{0,0}z_1)^{i-1} \right) \tag{3}
g_{r-0,1}(z_1, z_2) = p_{0,1}z_1 \left( \frac{z_2}{1 - p_{0,0}z_1} + (1 - z_2) \sum_{i=\max(0,k-r)}^{\max(1,d-r)} (p_{0,0}z_1)^{i-1} \right) \tag{4}
\]

where \( \max(i, j) \) is the maximal of the two integers \( i \) and \( j \), and the convention \( \sum_{i=1}^{\infty} = 0 \) applies.

**Proof:** Denote by \( p_i = P(\nu_{0,1} = i) \) and note that \( p_i = p_{0,1}p_{0,0}^{i-1} \) because \( \nu_{0,1} \) is geometrically distributed with probability of success \( p_{0,1} \) and support \( \{1, 2, \ldots\} \). Also note that \( I_{\nu_{0,1}} = 0 \) if and only if \( d \leq \nu_{0,1} \leq k \). Thus, for the joint generating function \( g_{0,1}(z_1, z_2) \) we get

\[
g_{0,1}(z_1, z_2) = \sum_{i=1}^{d-1} z_2 z_1^i p_i + \sum_{i=d}^{k} z_2^i z_1 p_i + \sum_{i=k+1}^{\infty} z_2 z_1^i p_i
\]

\[
= \sum_{i=1}^{d-1} z_2 z_1^i p_{0,1} p_{0,0}^{i-1} + \sum_{i=d}^{k} z_2^i z_1 p_{0,1} p_{0,0}^{i-1} + \sum_{i=k+1}^{\infty} z_2 z_1^i p_{0,1} p_{0,0}^{i-1}.
\]

Simplifying the above expression leads to (3) above. Similar arguments apply for the derivation of the expression for \( g_{r-0,1}(z_1, z_2) \) and the details are therefore omitted. The proof of Lemma 1 is complete. \( \square \)

We will derive simple recurrence relations leading to an exact evaluation of the joint generating function \( G_{m}^{(s)}(z_1, z_2) \) and \( G_{m}^{(intersite)}(z_1, z_2) \), which have been introduced above.

For the pattern of interest \( w = w_1 w_2 \ldots w_m \), denote

\[
I(s_1, s_2, \ldots, s_r) = \begin{cases} 1 & \text{if } (s_1, s_2, \ldots, s_r) = (w_1, w_2, \ldots, w_r) \\ 0 & \text{otherwise}, \end{cases} \\
T(s_1, s_2, \ldots, s_r) = 1 - I(s_1, s_2, \ldots, s_r), \quad r = 1, 2, \ldots, m, \tag{5}
\]

Waiting Time Distributions for Pattern Occurrence in a Constrained Sequence
that is, $I(s_1, s_2, \ldots, s_r) = I_{w_1}(s_1)I_{w_2}(s_2) \cdots I_{w_r}(s_r)$, where $I_{w}(\cdot)$ is an indicator function. For the sake of brevity we introduce the following notation. For each $j, j = 2, 3, \ldots, m - 1$, and each $a, a = 0, 1$, let

\[
\begin{align*}
L_1(j, a) &= 1 \\
L_2(j, a) &= I(w_2, w_3, \ldots, w_j, a) \\
L_3(j, a) &= I(w_2, w_3, \ldots, w_j, a)I(w_3, w_4, \ldots, w_j, a) \\
& \quad \vdots \\
L_r(j, a) &= I(w_2, w_3, \ldots, w_j, a) \cdots I(w_r, w_{r+1}, \ldots, w_j, a)I(w_{r+1}, w_{r+2}, \ldots, w_j, a) \\
& \quad \vdots \\
L_j(j, a) &= I(w_2, w_3, \ldots, w_j, a)I(w_3, w_4, \ldots, w_j, a) \cdots I(w_j, a),
\end{align*}
\]

where it is assumed that $I(w_i, w_{i+1}, \ldots, w_j, a) = 1$ if $i > j$. Note that $L_r(j, a) = 1$ for $r < j$ if and only if none of $w_i w_{i+1} \ldots w_j a$ for $i = 2, 3, \ldots, r$ is a prefix to $w_j w_{j+1} \ldots w_m$, whereas $w_{r+1} w_{r+2} \ldots w_j a$ is such. Also, $L_j(j, a)$ is equal to one if and only if none of $w_i w_{i+1} \ldots w_j a$ for $i = 2, 3, \ldots, j$ is a prefix to $w_j w_{j+1} \ldots w_m$. In other words, the $L_i(j, a)$ are relevant indicator functions related to the self-overlapping structure of the pattern $w = w_1 w_2 \ldots w_m$. In passing we observe that our definition of $L_r(j, a)$ is related to the autocorrelation set and polynomial of Guibas and Odlyzko [14] (cf. also [18, 23]).

Let now $Y_1^{(s)}(w_i^j)$ be the waiting time to see the pattern $w_i^j = w_i w_{i+1} \ldots w_j$, given the initial state is $s$. Then we define by

\[
G_j^{(s)}(z_1, z_2) = \mathbb{E}[z_1^{Y_1^{(s)}(w_i^j)} z_2^{N^{(d,k)}_{Y_1(w_i^j)}}]
\]

the joint generating function of $Y_1^{(s)}(w_i^j)$ and the associated count, $N^{(d,k)}_{Y_1(w_i^j)}$, of forbidden patterns (runs of zeros of length either less than $d$ or greater than $k$). Here we allow the first symbol, that is, $s$, to contribute to the pattern (of course this matters if $s = w_1$). Recall that for $j = m$ the joint generating function $G_m^{(s)}(z_1, z_2)$ has been introduced earlier. Let $G_j^{(r-0)}(z_1, z_2)$ be the joint generating function of the same quantities as above, given the initial state (assumed to be zero) is preceded by exactly $r$ zeros and the latter zeros are allowed to count towards the formation of the relevant event concerning the forbidden patterns (in other words the length of the first zero run within the waiting time $Y_1^{(0)}(w_i^j)$ is increased by $r$); also the initial state zero is allowed to contribute to the pattern. Further, let $G_j^{(w_i w_2 \ldots w_h)}(z_1, z_2)$ be the joint generating function of the same quantities as above, given the sub-pattern $w_1 w_2 \ldots w_h (h \leq j)$ has been reached.

Throughout the article it is assumed that the pattern of interest $w = w_1 w_2 \ldots w_m$ does not contain forbidden patterns. Each pattern of zeros and ones can be viewed as a sequence of alternating blocks of ones and zeros where the length of the $i$-th block is denoted by $k_i$ and $k_i > 0$ for $i = 2, 3, \ldots$ and $k_1 > 0$ if the initial symbol of the pattern is one whereas $k_1 = 0$ if the initial symbol is zero. For example, for the pattern 111000011000000 we have $k_1 = 3, k_2 = 4, k_3 = 2, k_4 = 5$, and for the pattern 001111100011 we have $k_1 = 0, k_2 = 2, k_3 = 5, k_4 = 3, k_5 = 2$.

Denote by $J_1$ and $J_2$ the following subsets of $\{1, 2, \ldots, m\}$, which are associated with the pattern
where $w = w_1 w_2 \ldots w_m$:

$$J_1 = \bigcup_{n=1}^{b} \{ j : \sum_{i=1}^{n} k_{2i-1} + 1 \leq j \leq \sum_{i=1}^{n} k_{2i-1} + d - 1 \},$$

$$J_2 = \bigcup_{n=1}^{b} \{ j : j = \sum_{i=1}^{n} k_{2i} \},$$

where $b$ is the number of zero blocks of the pattern $w$. For example, if the pattern of interest is 00111100011 and $k = 5$ we get $J_2 = \{2, 10\}$, and $J_1 = \{1, 8\}$.

Note that

$$G_1^{(1)}(z_1, z_2) = g_{*,w_1}(z_1, z_2),$$

where * stands for either 0 or 1, or $r - 0$. Actually, $G_1^{(1)}(z_1, z_2)$ has the same meaning as that of $g_{*,w_1}(z_1, z_2)$ unless $w_1 = 0$ and * = 1 (recall our definition of $g_{1,0}$ given in (8) above). For the latter case we formally assume that (8) holds and the reason for that assumption will become clear in the Idea of Proof of Theorem 1 below. Closed explicit expressions for $g_{i,j}(z_1, z_2)$ and $g_{r-1}(z_1, z_2)$ are found in Lemma 1 and prior to its statement. The formal proof is presented in the next section.

**Theorem 1** The following recurrence relations hold for the joint generating functions $G_j(z_1, z_2)$:

(i) For $j \notin (J_1 \cup J_2)$, and $h = 1, 2, \ldots, j$, and $r = 1, 2, \ldots$, we have

$$G_{j+1}^{(s)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(s)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j},$$

$$G_{j+1}^{(w_1 w_2 \ldots w_h)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(w_1 w_2 \ldots w_h)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j},$$

$$G_{j+1}^{(r-0)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(r-0)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j},$$

where

$$A_j = \sum_{i=1}^{j-1} L_i(j, 1 - w_{j+1})G_j^{(w_1 w_2 \ldots w_j \ldots i+1)}(z_1, z_2) + L_j(j, 1 - w_{j+1})G_j^{(1-w_{j+1})(z_1, z_2)},$$

with the convention $\sum_{i=1}^{0} = 0$, and the $L_i(j, a)$ have been introduced in (6).

(ii) For $j \in J_1$, and $h = 1, 2, \ldots, j$, and $r = 1, 2, \ldots$, we have

$$G_{j+1}^{(s)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(s)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j},$$

$$G_{j+1}^{(w_1 w_2 \ldots w_h)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(w_1 w_2 \ldots w_h)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j},$$

$$G_{j+1}^{(r-0)}(z_1, z_2) = \frac{p_{w_j, w_{j+1}} z_1 G_j^{(r-0)}(z_1, z_2)}{1 - p_{w_j, 1-w_{j+1}} z_1 A_j}.$$
where $A_j$ is as above.  
(iii) For $j \in J_2$, and $h = 1, 2, \ldots, j$, and $r = 1, 2, \ldots, \text{the same relations as those for } j \notin (J_1 \cup J_2)$ above hold after replacing $A_j$ by $B_j$, where

$$B_j = \sum_{i=1}^{j-1} L_i(j, 1 - w_{j+1}) G_j^{(w_1, w_2, \ldots, w_{j-1})}(z_1, z_2) + L_j(j, 1 - w_{j+1}) G_j^{(k_{2n} - 0)}(z_1, z_2), \quad (15)$$

and $n$ is associated with $j$ through $j = \sum_{i=1}^{n} k_{2i}$. Recall that $j \in J_2$ if and only if $j = k_1 + k_2 + \ldots + k_{2n}$ for some $n$, $n = 1, 2, \ldots.$

**Idea of the Proof:** The proof is based on a suitable extension of the methodology introduced in Stefanov [28]. The latter treated patterns formed on finite alphabets in strings generated by general discrete- and continuous-time models. In particular, Theorem 4.1 (cf. [28] p. 890) provides recurrence relations leading to exact evaluation of the joint generating function of the waiting time until reaching a pattern together with the associated counts of occurrences of the corresponding symbols of the alphabet. In this paper we deal with a simpler model (binary alphabet and discrete-time parameter) but the joint generating function of interest is that of the waiting time until reaching a pattern together with the associated count of occurrences of an event which is not as simple as the events considered in [28]. A careful scrutiny of [28] proofs reveals that the recurrence relations provided there are applicable to the joint generating function of the waiting time till reaching a pattern together with the associated count of occurrences of an ’event’ if the following two conditions are satisfied concerning that ’event’:

(i) the joint generating functions for the following quantities are available: the waiting time to reach a letter from another (or the same) letter of the alphabet together with the associated count of occurrences of the ’event’ of interest.

(ii) All occurrences of the ’event’ of interest are captured within the passage times between the states, that is, occurrence or non occurrence of the ’event’ does not depend on the history prior to a passage time or the future after that passage time.

Note that nominating the event of interest to be a run of zeros of length either less than $d$ or greater than $k$ we get that condition (ii) is not satisfied in general. For example, in a passage time from state zero to state one the occurrence or non occurrence of our event of interest (a constrained zero-run) depends on the number of zeros just preceding the starting state zero. As for a passage time from state one to state zero, note that, on one hand, the occurrence or non occurrence of the ’event’ of interest is not affected by the outcomes preceding the initial state one. On the other hand, within that passage time a run of zeros of length 1 occurs (the last observation within such passage time is zero which is preceded by one), that is, the event of interest occurs if $d > 1$ and given we stop observing the generated random sequence with such passage time. If we do not stop observing the generated sequence at such passage time then the occurrence or non occurrence of the event of interest depends on the future outcomes (that is, on how many zeros will follow after the first zero achieved from state one). Note that we stop observing the generated random sequences at occurrences of the pattern of interest, which we assume does not contain constrained zero-runs. That is, we do not stop observing the generated random sequence at a passage time from state one to state zero if $d > 1$. Therefore, within a passage time from state one to state zero we should not account for a possible occurrence of the event of interest, because such occurrence will be accounted for within the following passage time from state zero to state one. This is the reason for defining $g_{1,0}(z_1, z_2)$ to be equal to the generating function of $\nu_{1,0}$, as if assuming that within a passage
time from state one to state zero a constrained zero-run does not occur. By the same reason we assumed that $G_j^{(1)}(z_1, z_2) = g_1 w_j(z_1, z_2)$ if $w_0 = 0$ (cf. the comment prior to Theorem 1). In particular, we may assume that condition (i) above is satisfied for our problem because the relevant joint generating functions are provided in Lemma 1 and prior to it.

Further we show how the methodology in [28] can be extended to derive relevant recurrence relations for the case of the waiting time until reaching a pattern and the associated count of occurrences of constrained zero-runs.

Recall that we consider a pattern $w = w_1 w_2 \ldots w_m$ whose consecutive blocks of ones and zeros are of lengths $k_1, k_2, k_3, \ldots$, respectively.

Assume first the pattern of interest consists of the first $k_1 + 1$ symbols of $w$. Note that in this case condition (ii) is satisfied because there are no zeros preceding an initial zero state at a passage time from state zero to state one. Therefore, the relations given by (8), (9), (10) for $j$, such that $1 \leq j \leq k_1$ are a special case of the recurrence relations in Theorem 4.1 of [28]. Since the model treated in [28] is more general and the uninitiated reader may find it not quite transparent how to write the relations for our special case of the recurrence relations in Theorem 4.1 of [28] to get the relations (8), (9), (10).

Assume now that the pattern of interest consists of the first $k_1 + 2$ ($< k_1 + k_2$) symbols. Then note that from the time epoch at which we have reached the subpattern $w_1 w_2 \ldots w_{k_1 + 1}$ till reaching the pattern $w_1 w_2 \ldots w_{k_1 + 2}$ one may miss counts of the event of interest (a run of zeros of length less than $d$) due to the following observation. Upon reaching $w_1 w_2 \ldots w_{k_1 + 1}$ assume that in the next step the pattern $w_1 w_2 \ldots w_{k_1 + 2}$ is not reached (that is, a mismatch occur at this stage). Thus, a run of zeros of length $1 (< d$, of course given $d > 1$) has occurred and it will not be accounted for by the recurrence relations provided in [28]. To account for such occurrences one should multiply by $z_2$ the relevant joint generating functions each time such mismatch occurs. It is achieved by replacing the denominator $1 - p_{w_j,1-w_{j+1}, z_1} A_j$ by $1 - p_{w_j,1-w_{j+1}, z_1} z_2 A_j$ in the recurrence relations. Therefore, relations (12), (13), (14) hold for $j = k_1 + 1$.

Similarly, if the pattern consists of the first $k_1 + 3$ ($< k_1 + k_2$) symbols then upon reaching the subpattern $w_1 w_2 \ldots w_{k_1 + 2}$ on the following step one may miss a count of the event of interest. More specifically, this is a run of zeros of length 2 ($< d$, of course given $d > 2$). To account for such occurrences of the event of interest again the denominator $1 - p_{w_j,1-w_{j+1}, z_1} A_j$ is to be replaced by $1 - p_{w_j,1-w_{j+1}, z_1} z_2 A_j$. Clearly, the same argument applies if the pattern consists of the first $k_1 + j - 1$ ($< k_1 + k_2$) symbols of $w$, where $j \leq d$. Therefore, the recurrence relations (12), (13), (14) hold for $j$ such that $k_1 + 1 \leq j \leq k_1 + d - 1$ (these $j$’s belong to $J_1$). Further, note that for larger $j$, such that $k_1 + d \leq j \leq k_1 + k_2 - 1$ (note that such $j$’s do not belong to $J_1 \cup J_2$) the recurrence relations given by (8), (9), (10) hold, because in mismatch situations constrained zero-runs do not occur.

Assume now that the pattern consists of the first $k_1 + k_2 + 1$ symbols of $w$. Then note that in a mismatch situation at the next step after reaching the sub-pattern consisting of the first $k_1 + k_2$ symbols, we are at state zero with exactly $k_2$ zeros preceding it. The method in [28] implies that $G_j^{(1-w_{j+1})}(z_1, z_2)$ (this is the generating function, in the expression for $A_j$, which accounts for the evolution of the sequence after such a mismatch situation, and given no overlap occurred after that mismatch) from the expression for $A_j$, given in (11), is to be substituted by $G_j^{(k_2-0)}(z_1, z_2)$ in order to account for all occurrences of the event of interest. That is, for $j = k_1 + k_2$, the relations (8), (9), (10) hold with $G_j^{(1-w_{j+1})}(z_1, z_2)$ replaced by $G_j^{(k_2-0)}(z_1, z_2)$ in the expressions for the $A_j$; that is, $A_j$ is replaced by $B_j$.

For larger $j$ ($j > k_1 + k_2$) similar arguments to those above apply.
3 Proof of Theorem 1

First, recall that the g.f., $G(z)$, of the geometric distribution on $\{0, 1, \ldots\}$, with probability of ‘success’ $p$, is given by $p/(1-qz)$, where $q = 1-p$. Also recall that for the g.f. of the random sum $Y = \sum_{i=1}^{\nu} Y_i$ we have

$$G_Y(z) = G_{\nu}(G_{Y_i}(z)),$$

where $z = (z_1, z_2, \ldots, z_n)$ and the $Y_i$ are independent and identically distributed (i.i.d.) random vectors with g.f. $G_{Y_i}(z)$ and $\nu$ is a non-negative r.v., independent of the $Y_i$, with g.f. $G_{\nu}(z)$. If the distribution of $\nu$ is geometric then the random sum $Y$ is called a geometric sum.

The following quantity is called briefly the first return time to the pattern $w_1w_2\ldots w_j$:

$$\inf\{n \geq 1 : X(n+1)\ldots X(n+1+j) = w_1\ldots w_j | X(1)\ldots X(j) = w_1\ldots w_j\}.$$  

Recall that the pattern of interest is denoted by $w = w_1w_2\ldots w_m$. Note that $j \notin J_1 \cup J_2$ if and only if either $w_j+1 = 1$, or $w_jw_{j+1} = 10$, or $w_{j+1} = 0$ and the number of zeros preceding $w_{j+1}$, in the block of zeros to which $w_{j+1}$ belongs, is less than $d$ (recall that $d$ pertains to the term $(d,k)$-sequence).

We will prove the validity of (8), (9) and (10) first for $j = 1, 2, \ldots, k_1$. Recall that $k_1$ is the length of the first block of ones of the pattern of interest $w$. Of course these $j$’s do not belong to the set $(J_1 \cup J_2)$. Now consider the subpattern $w_1w_2$ consisting of the first two symbols of the pattern $w$, of course assuming that $k_1 \geq 1$. Note that the joint generating function of the first return time to state $w_1$ and the associated count of the forbidden patterns within that return time, conditional on not entering state $w_2$ at the first step, is

$$H_1(z_1, z_2) = \frac{p_{w_1,1-w_2} z_1 g_{1-w_2,w_1}(z_1, z_2)}{1-p_{w_1, w_2}}.$$  

Actually, (17) is derived via conditioning on the first step. It is easy to see, using the strong Markov property and applied to the consecutive entry times to state $w_1$, that the joint distribution of the waiting time to reach the pattern $w_1w_2$ from state $s$ and the associated count of forbidden patterns up to that waiting time, is the same as the joint distribution of

$$K_1 + e_1 + \sum_{i=0}^{\nu_1} Y_{i,1},$$

where $e_1$ is the unit vector $(1, 0)$, the $Y_{i,1}$ are i.i.d. (two-dimensional) random vectors, also independent of $K_1$, and $\nu_1$ is a geometric random variable, independent of the $Y_{i,1}$ and $K_1$ with a probability of ‘success’ $p_{w_1, w_2}$. Further, the (two-dimensional) random vector $K_1$ has the same joint distribution as that of the waiting time to reach $w_1$ from state $s$, and the associated count of forbidden patterns, that is, its joint g.f. $G_{K_1}(z_1, z_2)$ is equal to $g_{s,w_1}(z_1, z_2)$. The random vector $Y_{i,1}$ has a joint generating function given by $H_1(z_1, z_2)$. Thus, in view of (16) and (17) and recalling that $G_{1(s)}^{(s)}(z_1, z_2) = g_{s,w_1}(z_1, z_2)$ we get that

$$G_2(z_1, z_2) = g_{s,w_1}(z_1, z_2) \left( \frac{p_{w_1, w_2} z_1}{1 - (1-p_{w_1, w_2})H_1(z_1, z_2)} \right) = \frac{p_{w_1, w_2} z_1 G_{1(s)}^{(s)}(z_1, z_2)}{1 - p_{w_1, 1-w_2} z_1 G_{1}^{(s)}(z_1, z_2)}.$$
Waiting Time Distributions for Pattern Occurrence in a Constrained Sequence

Using the same arguments as those above we get that for \( r = 1, 2, \ldots \)

\[
G_2^{(w_1)}(z_1, z_2) = \frac{p_{w_1,w_2} z_1 G_1^{(w_1)}(z_1, z_2)}{1 - p_{w_1,1-w_2} z_1 G_1^{(1-w_2)}(z_1, z_2)},
\]

\[
G_2^{r-0}(z_1, z_2) = \frac{p_{w_1,w_2} z_1 G_1^{(r-0)}(z_1, z_2)}{1 - p_{w_1,1-w_2} z_1 G_1^{(1-w_2)}(z_1, z_2)}.
\]

That is, noting that \( A_1 = G_1^{(1-w_2)}(z_1, z_2) \) (cf. (11)), we get that (8), (9) and (10) hold for \( j = 1 \).

Now consider the subpattern \( w_1 w_2 w_3 \), assuming that \( k_1 \geq 2 \) (that is \( j = 2 \)). Similarly to the preceding case (when \( j = 1 \)), conditioning on the first step, note that the joint generating function of the first return time to the subpattern \( w_1 w_2 \) and the associated count of the forbidden patterns within that return time, conditional on not entering state \( w_3 \) at the first step, is given by

\[
H_2(z_1, z_2) = \frac{p_{w_2,1-w_3} z_1 \left( L_1(2, 1 - w_3) + L_2(2, 1 - w_3) G_2^{(1-w_3)}(z_1, z_2) \right)}{1 - p_{w_2,w_3}},
\]

where the \( L_i(j, a) \) have been introduced in (6). Again using the strong Markov property and applied to the consecutive entry times to the subpattern \( w_1 w_2 \), we get that the joint distribution of the waiting time to reach the subpattern \( w_1 w_2 w_3 \) from state \( s \) and the associated count of forbidden patterns up to that waiting time, is the same as the joint distribution of

\[
K_2 + e_1 + \sum_{i=0}^{\nu_2} Y_{i,2}
\]

where \( e_1 \) is the unit vector \((1, 0)\), the \( Y_{i,2} \) are i.i.d. random vectors, also independent of \( K_2 \), and \( \nu_2 \) is a geometric random variable, independent of the \( Y_{i,2} \) and \( K_2 \) with a probability of 'success' \( p_{w_2,w_3} \). Further, the random vector \( K_2 \) has the same joint distribution as that of the waiting time to reach \( w_1 w_2 \) from state \( s \), and the associated count of forbidden patterns, that is, its joint g.f. \( G_{K_2}(z_1, z_2) \) is equal to \( G_{2}^{(s)}(z_1, z_2) \). The random vector \( Y_{i,2} \) has a joint generating function given by \( H_2(z_1, z_2) \). Therefore, similarly to the preceding case, and using (18) we get that

\[
G_3^{(s)}(z_1, z_2) = \frac{p_{w_2,w_3} z_1 G_2^{(s)}(z_1, z_2)}{1 - (1 - p_{w_2,w_3}) H_2(z_1, z_2)} = \frac{p_{w_2,w_3} z_1 G_2^{(s)}(z_1, z_2)}{1 - p_{w_2,1-w_3} z_1 \left( L_1(2, 1 - w_3) + L_2(2, 1 - w_3) G_2^{(1-w_3)}(z_1, z_2) \right)}.
\]

That is, (8) holds for \( j = 2 \). Likewise, (9) and (10) hold for \( j = 2 \). The same arguments, as those used in the cases for \( j = 1, 2 \) apply to any \( j \) such that \( 1 \leq j \leq k_1 \). Therefore (8), (9) and (10) hold for \( j = 1, 2, \ldots k_1 \).

Now consider the case when \( j \in J_1 \). First, we will consider the \( j \)'s belonging to \( \{ j : \sum_{i=1}^{j} k_{2i-1} + 1 \leq j \leq \sum_{i=1}^{k_1} k_{2i-1} + d - 1 \} \), that is, for \( j = k_1 + 1, k_1 + 2, \ldots, k_1 + d - 1 \). Let \( j = k_1 + 1 \), that is, we consider
the subpattern \(w_1 w_2 \ldots w_{k_1+2}\). Again, conditioning on the first step, note that the joint generating function of the first return time to the subpattern \(w_1 w_2 \ldots w_{k_1+1}\) and the associated count of the forbidden patterns within that return time, conditional on not entering state \(w_{k_1+2}\) at the first step, is given by

\[
H_{k_1+1}(z_1, z_2) = \frac{p_{w_{k_1+1}, 1-w_{k_1+2} z_1 z_2 A_{k_1+1}}}{1 - p_{w_{k_1+1}, w_{k_1+2}}},
\]

(19)

where the \(A_j\) and \(L_i(j, a)\) have been introduced in (11) and (6), respectively. Actually, \(H_{k_1+1}(z_1, z_2)\) differs from its counterparts \(H_j(z_1, z_2)\), \(j \leq k_1\), (cf. (17) and (18)) by the presence of \(z_2\) in front of \(A_{k_1+1}\). The presence of \(z_2\) accounts for unaccounted otherwise occurrence of a forbidden pattern (a zero run of length less than \(d\)) at the first step when one fails to reach in one step the state \(w_{k_1+2}\) from the already reached subpattern \(w_1 w_2 \ldots w_{k_1+1}\). Further, similarly to the preceding cases and applying the strong Markov property to the consecutive entry times to the subpattern \(w_1 w_2 \ldots w_{k_1+1}\), one gets that

\[
G_{k_1+2}^{(s)}(z_1, z_2) = \frac{p_{w_{k_1+1}, w_{k_1+2} z_1 G_{k_1+1}^{(s)}(z_1, z_2)}}{1 - (1 - p_{w_{k_1+1}, w_{k_1+2}}) H_{k_1+1}(z_1, z_2)} = \frac{p_{w_{k_1+1}, w_{k_1+2} z_1 G_{k_1+1}^{(s)}(z_1, z_2)}}{1 - p_{w_{k_1+1}, 1-w_{k_1+2} z_1 z_2 A_{k_1+1}}},
\]

Thus, \(12\) holds for \(j = k_1 + 1\). Likewise, one gets that \(13\) and \(14\) hold for \(j = k_1 + 1\). Exactly the same arguments, as those used in the case for \(j = k_1 + 1\) apply to any \(j\) such that \(k_1 + 1 \leq j \leq k_1 + d - 1\). Therefore, \(12\), \(13\) and \(14\) hold for \(j = k_1 + 1, k_1 + 2, \ldots, k_1 + d - 1\). Consider now \(j = k_1 + d, k_1 + d + 1, \ldots, k_1 + k_2 - 1\). These \(j\)’s do not belong to \(J_1 \cup J_2\). Note that the relevant \(H_{j}(z_1, z_2)\) is given by

\[
H_{j}(z_1, z_2) = \frac{p_{w_{j+1}, 1-w_{j+1} z_1 A_j}}{1 - p_{w_{j}, w_{j+1}}},
\]

(20)

that is, \(20\) has the same form as that of \(17\) and \(18\). This is due to the observation that at the first step when \(w_{j+1}\) is not reached from the already reached subpattern \(w_1 w_2 \ldots w_j\) a forbidden pattern does not occur (the reached zero run is of length at least \(d\) and of course less than \(k\)). Therefore, \(8\), \(9\) and \(10\) hold for \(j = k_1 + d, k_1 + d + 1, \ldots, k_1 + k_2 - 1\).

Consider now the case \(j = k_1 + k_2\). This \(j\) belongs to \(J_2\). Note that the relevant \(H_{j}(z_1, z_2)\) for the joint g.f. of the first return time to the subpattern \(w_1 w_2 \ldots w_{k_1+k_2}\) and the associated count of the forbidden patterns within that return time, conditional on not entering state \(w_{k_1+k_2+1}\) at the first step, is given by

\[
H_{k_1+k_2}(z_1, z_2) = \frac{p_{w_{k_1+k_2}, 1-w_{k_1+k_2+1} z_1 B_{k_1+k_2}}}{1 - p_{w_{k_1+k_2}, w_{k_1+k_2+1}}},
\]

(21)

where \(B_j\) is given in \(15\). More specifically, note first that \(B_j\) differs from \(A_j\) only through the generating function associated with the indicator function \(L_j(j, 1 - w_{j+1})\) (cf. \(6\)). This g.f. is \(G_j^{(1-w_j)}\) for \(A_j\) and \(G_j^{(k_2n-0)}\) \((n = 1\) for \(j = k_1 + k_2\)) for \(B_j\), and it accounts for what happens after the first step given \(L_j(j, 1 - w_{j+1}) = 1\). Further, note that conditioning on not entering state \(w_{k_1+k_2+1}\) at the first step means that a run of zeros of length exactly \(k_1 + k_2 + 1\) has been reached, that is after the first step the current state zero is preceded by exactly \(k_1 + k_2\) zeros. After that first step is made one waits until
subpattern $w_1 w_2 \ldots w_{k_1 + k_2}$ is reached. Therefore, noticing that after the first step no overlap occurs, that is

$$L_{k_1 + k_2}(k_1 + k_2, 1 - w_{k_1 + k_2 + 1}) = 1,$$

one can see that the relevant generating function which will capture all occurrences of forbidden patterns up to this waiting time is given by $G(k_2 - 0)(z_1, z_2)$. Therefore, it is clear now that (8), (9) and (10) with $A_j$ replaced by $B_j$ hold for $j = k_1 + k_2$. Note that for the other $j$’s from $J_2$, such as say $j = k_1 + k_2 + k_3 + k_4$, an overlap may occur (for example if $k_4 < k_2$) after the first step but then the relevant g.f. $(G^{(w_1 w_2 \ldots w_{k_1 + k_2})})$ in the expression for $B_j$ will capture all occurrences of forbidden patterns. Recall that the indicator functions appearing in the expressions for $A_j$ and $B_j$ sum to 1.

It is easy to see that the same arguments as those used for $j = 1, 2, \ldots, k_1 + k_2$ apply for $j > k_1 + k_2$.

The proof of Theorem 1 is complete. $\square$

**Remark 3.1** Note that the generating functions $G_j^{(r - 0)}(z_1, z_2)$ play a pivotal role, through the relevant entry in the expression for $B_j$, in the process of evaluating the relevant generating functions $G_j^{(s)}(z_1, z_2)$ and $G_j^{(w_1 w_2 \ldots w_h)}(z_1, z_2)$. Therefore, only $G_j^{(k_2 - 0)}(z_1, z_2)$ for $i = 1, 2, \ldots, r$, are to be evaluated for each $j$.

Theorem provides a route for exact evaluation of the $G_j^{(s)}(z_1, z_2)$ and $G_j^{(w_1 w_2 \ldots w_h)}(z_1, z_2)$. In particular, these contain the generating functions of interest, that is $G_j^{(s)}(z_1, z_2)$ and $G_j^{(intersite)}(z_1, z_2)$. More specifically, note that $G_j^{(intersite)}(z_1, z_2)$ is equal to $G_j^{(w_1 w_2 \ldots w_h)}(z_1, z_2)$ for $h$ such that $w_1 w_2 \ldots w_h$ is the longest prefix, which is also a suffix, to the pattern $w$.

### 4 Numerical Analysis

In this section, we present some numerical results. Using Maple 9, we derived explicit expressions for the joint generating functions given in Theorem 1 for the patterns in our numerical examples. These provided us with explicit expressions for the probability generating function $G^{(n)}(z)$ of $Y_n^{(s)}$ and subsequently explicit expressions for the cumulative generating function of $Y_n^{(s)}$, since the latter g.f. is equal to $G^{(n)}(z)/(1 - z)$. Recall that the distribution of $Y_n^{(s)}$, is equal to the conditional distribution of the waiting time until the $r$-th occurrence of the pattern of interest given there were no forbidden patterns up to time $Y_n^{(s)}$. The relevant probabilities, that is $P(Y_n \leq n - 1|\bar{N}^{(d,k)}_n = 0)$, $P(N_n \geq r|\bar{N}^{(d,k)}_n = 0)$, are computed via a numerical inversion of the cumulative generating function $G^{(n)}(z)/(1 - z)$. We used the numerical procedure introduced by Abate and Whitt. This procedure is very fast and computes the exact probabilities with any given, in advance, accuracy. The computation was implemented on Powerbook G4 using Maple 9. The derivation of the expression for $G^{(n)}(z)/(1 - z)$ is almost instantaneous.

In the example presented in Table 1 we compute the probabilities $P(N_n \geq r|\bar{N}^{(d,k)}_n = 0)$, $r = 1, 2, \ldots$, with initial symbol $X_1 = 0$ or 1, for $w = 100100100$, for $d = 1$ and $k = 4$ and transition probabilities $p_{0,0} = 0.4, p_{0,1} = 0.6, p_{1,0} = 1, p_{1,1} = 0$. Note that if $d > 0$ then $p_{1,0} = 1$ because runs of 1’s are not allowed in such sequences. Runs of 1’s are allowed only in $(0,k)$-sequences.
Tab. 1: Probabilities for the number of occurrences, $N_n$, of $w = 100100100$ in a random $(1, 4)$-sequence of length $n = 500$ with $p_{0,0} = 0.4, p_{0,1} = 0.6, p_{1,0} = 1, p_{1,1} = 0$.

| $r$ | $P(N_n \geq r | \bar{N}_n^{(1,4)} = 0, X_1 = 0)$ | $r$ | $P(N_n \geq r | \bar{N}_n^{(1,4)} = 0, X_1 = 1)$ |
|-----|---------------------------------|-----|---------------------------------|
| 1   | 0.9998229893                    | 1   | 0.9998277951                    |
| 2   | 0.998829172                     | 2   | 0.9988712924                    |
| 3   | 0.995704487                     | 3   | 0.995854122                     |
| 4   | 0.988586721                     | 4   | 0.988809846                     |
| 5   | 0.9748567886                    | 5   | 0.9753124330                    |
| 6   | 0.9521461776                    | 6   | 0.9529272243                    |
| 7   | 0.9185406129                    | 7   | 0.9197338459                    |
| 8   | 0.8731122847                    | 8   | 0.8747731853                    |
| 9   | 0.8161570287                    | 9   | 0.8182964131                    |
| 10  | 0.7491883799                    | 10  | 0.7517679679                    |
| 11  | 0.6747008052                    | 11  | 0.677640655                     |
| 12  | 0.5958347059                    | 12  | 0.5990149589                    |
| 13  | 0.515805941                     | 13  | 0.5191630134                    |
| 14  | 0.4379186414                    | 14  | 0.441925930                     |
| 15  | 0.365829805                     | 15  | 0.3677218233                    |
| 16  | 0.2977955830                    | 16  | 0.3007062001                    |
| 17  | 0.2387613333                    | 17  | 0.2413791679                    |
| 18  | 0.1880023708                    | 18  | 0.1902906906                    |
| 19  | 0.1454637033                    | 19  | 0.1474121477                    |
| 20  | 0.110650150                     | 20  | 0.1122769106                    |
| 21  | 0.0828116490                    | 21  | 0.0841261739                    |
| 22  | 0.0609984853                    | 22  | 0.0620430520                    |
| 23  | 0.0442483799                    | 23  | 0.0450616861                    |
| 24  | 0.0316263751                    | 24  | 0.0322475305                    |
| 25  | 0.022837706                     | 25  | 0.022745761                     |
| 26  | 0.0154852671                    | 26  | 0.0158285526                    |
| 27  | 0.010617613                     | 27  | 0.0108665930                    |
| 28  | 0.0071864591                    | 28  | 0.0073639909                    |
| 29  | 0.0048032867                    | 29  | 0.0049280424                    |
| 30  | 0.0031715243                    | 30  | 0.0032579386                    |
| 31  | 0.0020694696                    | 31  | 0.0021284793                    |
| 32  | 0.0013349299                    | 32  | 0.0013746927                    |
| 33  | 0.0008515403                    | 33  | 0.0008779885                    |
References


