Random 0-1 rectangular matrices: a probabilistic analysis

Guy Louchard* Helmut Prodinger†

May 29, 2002

Abstract

Carlitz considered in [5] $r \times n$ matrices with entries being zero or one and the number of changes, i. e., the number of (horizontally or vertically) adjacent entries which are different. We extend these results in many ways. For instance, we exhibit that the limiting distribution is Gaussian and get explicit formulæ for some moments even in the general instance of d dimensions (instead of just two).

1 Introduction

Carlitz considered in [5] $r \times n$ matrices with entries being zero or one. He was interested in the *number of changes*, i. e., the number of (horizontally or vertically) adjacent entries which are different. He was able to analyze the instances $r = 1, 2, 3$ and general n.

Here we extend these results in many ways and discuss various methods that are appropriate for these kind of problems. Thus, this paper can be partially seen as a tutorial effort.

To be more specific, we start with a matrix approch, by considering r to be fixed and adding "a new slice;" generating functions of interest can then be deduced by essentially inverting a matrix. Using techniques of Bender [2] we can get the Gaussian limit distribution from this, as well as moments of any order.

Then we consider the problem in a more probabilistic manner; the underlying Markov process allows us to deduce the essential results without inverting a (huge) matrix. In particular, we are able to translate the problem into the language of the *Brownian motion*. We also discuss (briefly) how the *Drazin inverse* can be used in this context.

We extend Carlitz' problem from two dimensions to d dimensions: Assume that d integers $n_1, \ldots, n_d$ are given. We consider functions (as generalizations of matrices)

$$f : \{1, \ldots, n_1\} \times \ldots \times \{1, \ldots, n_d\} \longrightarrow \{0, 1\},$$

and count the number of changes in adjacent "cells." Two cells, indexed by $(i_1, \ldots, i_d)$ resp. $(j_1, \ldots, j_d)$ are adjacent iff $\sum_{k=1}^{d} |i_k - j_k| = 1$.

Furthermore, we are able to get the moments explicitly.

Finally, we offer a section with simulation results which show an excellent fit with the predicted results.

Although our primary motivation was from analytic combinatorics, let us mention that our model is actually related to the Ising Partition function at infinite temperature (see, for instance, Baxter [1] for a good introduction). It could also be used in the problem of testing for nonhomogeneous white noise: see, for instance, Dufour, Hallin and Mizera [6]. This will be the object of future work.

*Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium, email: louchard@ulb.ac.be

†University of the Witwatersrand, The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, P.O. Wits, 2050 Johannesburg, South Africa, e-mail: helmut@staff.ms.wits.ac.za

The paper is organized as follows: Section 2 analyses the number of changes in simple cases. Section 3 starts the probabilistic approach. Section 4 gives the asymptotic distribution of the number of changes. Section 5 extends our results to multi-dimensional matrices. Section 6 provides an automatic procedure to generate some arrays we need in previous sections. Section 7 deals with the general case and Section 8 offers some simulations.

2 Number of changes

Let us consider a rectangular array $R[1..r, 1..n]$ with entries in $\{0,1\}$. We will analyze the case $r = 3$, which captures all useful ingredients. Let $K(n)$ be the total number of changes, with $f_{v,w,x}(n,k)$ denoting the number of arrays with last column (v,w,x) and k changes. We remark, as Carlitz did in [5], that a symmetrization can be applied: interchanging 0 and 1 or reversing the last column order: $(v,w,x) \leftrightarrow (x,w,v)$ lead to the same $f_\bullet$ values. Hence we will group some final columns values together, leading to three states:

$$[1] = \{000, 111\},\ [2] = \{001, 110, 011, 100\},\ [3] = \{010, 101\}.$$

So we replace the $f_\bullet$ by a (row) vector

$$\phi(n,y) := [\phi_1(n,y), \phi_2(n,y), \phi_3(n,y)],$$

where $\phi_i(n,y)$ is the generating function (GF) of the total number of changes with final column in state $[i]$, i.e.,

$$\phi_i(n,y) := \sum_{k\geq 0} \sum_{(v,w,x)\in[i]} f_{v,w,x}(n,k) y^k.$$

It is thus convenient to look for the **total** number of changes from **all** columns in $[i]$ to **any** column of $[j]$ (by definition, their values are identical). This leads to a matrix $H[1..3, 1..3]$, given by

$$H = \begin{pmatrix} 1+y^3 & y+y^2 & y+y^2 \\ 2(y+y^2) & 1+y+y^2+y^3 & 2(y+y^2) \\ y+y^2 & y+y^2 & 1+y^3 \end{pmatrix}.$$

Moreover, each state $[i]$ has its own changes GF given by the number of changes of **any** of his columns. This leads to a vector

$$v = (1, y, y^2).$$

Let

$$V = \begin{pmatrix} v[1] & 0 & 0 \\ 0 & v[2] & 0 \\ 0 & 0 & v[3] \end{pmatrix}.$$

The final changes GF matrix is thus given by

$$A = H \cdot V = \begin{pmatrix} 1+y^3 & y^2+y^3 & y^3+y^4 \\ 2(y+y^2) & y+y^2+y^3+y^4 & 2(y^3+y^4) \\ y+y^2 & y^2+y^3 & y^2+y^5 \end{pmatrix}.$$

Note that, for $y = 1$, we obtain

$$A_1 := A|_{y=1} = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 4 & 4 \\ 2 & 2 & 2 \end{pmatrix}.$$

Let us remark that A is the transpose of the matrix M mentioned in [5]. However, Carlitz derives this matrix at the end of his analysis, as a by-product of previous computations. Here, we start directly from the matrix A.

We obtain immediately $\phi(n) = \phi(1)A^{n-1}$, where $\phi(1) \equiv v = (1, y, y^2)$. Let

$$\begin{aligned} v_1 &:= (2,4,2), & (1)\\ \delta(n) &:= \phi(n) \cdot v_1^T. & (2) \end{aligned}$$

Note that v_1 is the transpose of the first column of A_1. The quantity $\delta(n)$ gives the GF of the total number of changes in **all** matrices. Of course, substituting $y = 1$ leads to $2^{rn} = 8^n$. Let

$$F(z,y) := \sum_{n\geq 1, k\geq 0} [\text{number of } 3\times n \text{ rectangles with } k \text{ changes}]\, z^n y^k,$$

then we obtain a rational function:

$$\begin{aligned} F(z,y) &= \sum_{n\geq 1} \phi(n) z^n \cdot v_1^T \\ &= \phi(1) z \sum_{n\geq 0} z^n A^n \cdot v_1^T \\ &= \phi(1) z (I - zA)^{-1} \cdot v_1^T \\ &= \frac{g}{h}, \end{aligned}$$

where

$$\begin{aligned} g &:= 2(1+y)^2 z - 2(1+y)y(3+2y+3y^2)(1-y)^2 z^2 + 8y^3(1+y)^2(1-y)^4 z^3, \\ h &:= 1 - (1+y)(1+y^2)^2 z + y(1+y^2)(1-y)^2(1+y)^4 z^2 - y^3(1-y)^4(1+y)^5 z^3. \end{aligned}$$

When $y = 1$, the dominant singularity of $F(z,1)$ is given by $z^* = 1/8$. Consider all matrices as equiprobable. By Bender's techniques, as given by Theorems 1, 3 and (3.2) in [2] we can derive the asymptotic distribution of the total number of changes $K(n)$. Set

$$\begin{aligned} r_1 &:= -h_y/h_z, \\ r_2 &:= -(r_1^2 h_{zz} + 2r_1 h_{zy} + h_y + h_{yy})/h_z. \end{aligned}$$

Let $y = 1$, $z = z^*$; this gives

$$h_y^* = -5/2,\ h_z^* = -8,\ h_{yy}^* = -5,\ h_{zz}^* = 0,\ h_{zy}^* = -20.$$

Bender's results lead to

$$\begin{aligned} M^* &:= -r_1/z^* = 5/2 \\ \sigma^{*2} &:= M^{*^2} - r_2/z^* = 5/4. \end{aligned}$$

Also,

$$K(n) \sim \mathcal{N}(nM^*, n\sigma^{*2}), \quad n \to \infty, \tag{3}$$

where $\mathcal{N}$:= a Gaussian (normal) random variable (RV).

We can even get more precise expressions for the mean $M(n)$ and the variance $\sigma^2(n)$ of $K(n)$. Indeed the generating functions of the expectations is

$$\frac{\partial F}{\partial y}\left(\frac{z}{8}, 1\right) = \frac{z(2+3z)}{2(1-z)^2},$$

from which we get the expectation

$$M(n) = 5/2n - 3/2. \tag{4}$$

Similarly,

$$\frac{\partial^2 F}{\partial y^2}\left(\frac{z}{8}, 1\right) = \frac{z(1+18z+6z^2)}{2(1-z)^3},$$

which gives the second moment m_2 from $F'' + F'$:

$$m_2 = (5n-2)(5n-3)/4$$

from which we get the variance

$$\sigma^2(n) = 5/4n - 3/4. \tag{5}$$

Also,

$$\frac{\partial^3 F}{\partial y^3}\left(\frac{z}{8}, 1\right) = \frac{15z^2(7+16z+2z^2)}{4(1-z)^4},$$

which gives the third moment m_3 from $F''' + 3F'' + F'$:

$$m_3(n) = 5n(5n-3)^2/8$$

and

$$\frac{\partial^4 F}{\partial y^4}\left(\frac{z}{8}, 1\right) = \frac{3z^2(13z^3+261z^2+314z+37)}{2(1-z)^5},$$

which gives m_4 from $F'''' + 6F''' + 7F'' + F'$:

$$m_4(n) = 625n^4/16 - 375n^3/8 + 75n^2/16 + 109n/8 - 6$$

. (The perspicacious reader has recognized the Stirling numbers of the second kind.) The centered moments are given by

$$\begin{aligned} \mu_3 &= m_3 - 3m_1m_2 + 2m_1^3 = 0, \\ \mu_4 &= m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 = 75n^2/16 - 13n/4 - 15/16. \end{aligned}$$

An efficient way of computing the moments will be given in Section 7. Note that $\mu_3 = 0$. This will be shown to be true for any matrix and is explained in Section 7.3.1.

3 Probabilistic approach

Now we consider a random filling with $\Pr(0) = \Pr(1) = 1/2$. To derive the corresponding Markov transition matrix, P, from A, we must apply two operators Ω_1 and Ω_2 defined as follows. First, we must divide each **line** of A by the **size** of the corresponding state. So Ω_1 amounts to computing

$$A^{(1)} = \begin{pmatrix} A(1,.)/2 \\ A(2,.)/4 \\ A(3,.)/2 \end{pmatrix}.$$

Next, we must multiply each term $(.,j)$ by the probability of stepping into state j, which is proportional to the **size** of the state, i. e., to $v_1 = (2,4,2)$, which leads to multiplying the **columns** by the probability distribution

$$v_2 = (1/4, 1/2, 1/4) = v_1/2^r. \tag{6}$$

So Ω_2 computes

$$P := (A^{(1)}(.,1)/4, A^{(1)}(.,2)/2, A^{(1)}(.,3)/4).$$

This gives

$$P = \begin{pmatrix} (1+y^3)/8 & (y^2+y^3)/4 & (y^3+y^4)/8 \\ (y+y^2)/8 & (y+y^2+y^3+y^4)/8 & (y^3+y^4)/8 \\ (y+y^2)/8 & (y^2+y^3)/4 & (y^2+y^5)/8 \end{pmatrix}.$$

The meaning of this probability generating function (PGF) matrix is that, for $k \in [0..5]$, $[y^k]P(i,j)$ is the probability of jumping from a global state i to a global state j, bringing k changes. Of course, and this decomposition is true for all r,

$$P^* = P|_{y=1} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} = 1 \times v_2, \tag{7}$$

and $P^* \cdot 1 = 1$ (1 denoting a column vector of ones). Also

$$P'|_{y=1} = \begin{pmatrix} 3/8 & 5/4 & 7/8 \\ 3/8 & 5/4 & 7/8 \\ 3/8 & 5/4 & 7/8 \end{pmatrix} = 1 \times v_3 \tag{8}$$

where

$$v_3 = (3/8, 5/4, 7/8).$$

This decomposition is also true for all r. Indeed, take for instance, for $r = 3$, one column $(a_1, a_2, a_3)^T$. The number of changes to an adjacent column $(b_1, b_2, b_3)^T$ is given by $(d_1, d_2, d_3)^T$, where $d_i := |a_i - b_i|$. Let $\Delta_1 := d_1 + d_2 + d_3$. However, in each state, we find also a column $(1 - a_1, 1 - a_2, 1 - a_3)^T$, leading to $\Delta_2 = 3 - \Delta_1$ changes, so the mean number of changes relative to these two columns is given by $(\Delta_1 + \Delta_2)/2 = 3/2$. Therefore the mean number of changes from state i to state j is given by

$$v_2[j][3/2 + v'[j]|_{y=1}],$$

independently of i. This gives exactly $v_3[j]$.

We have here a lumpability of Markov chain states, where the classical condition is that, for each elementary state l of a global state i, the probability of jumping from l to a global state j, given by $\sum_{s \in j} PE(l,s)$ (PE is the basic states transition matrix), is independent of l. This is the case in our problem, by symmetrization and the fact that the elementary transition probabilities from $l = (v_1, w_1, x_1)$ to $s = (v_2, w_2, x_2)$ are all given by $1/8$. However, note that here, we use instead the fact that the total probability mass brought from a global state i to some elementary state s of a global state j, given by $\sum_{l \in i} PE(l,s)$, is independent of s and that the starting probabilities are the same for all elementary states.

Starting from the state

$$\psi(1) = (1/4, y/2, y^2/4), \tag{9}$$

the number of changes PGF vector $\psi(n)$ is given by $\psi(1)P^{n-1}$ and the total number of changes PGF is given by

$$\kappa(y,n) = \psi(n) \cdot 1.$$

We should have $\kappa(y,n) \equiv \delta(n)/8^n$, as given by (2). This is easily checked, indeed operator Ω_1 is compensated by the weights of $\psi(1)$ in (9) and operator Ω_2 has the same effect as the components of v_1 in (2). Moreover, Ω_1 and Ω_2 cancel each other between successive applications of P.

4 Asymptotic distribution of K

4.1 The limiting Markov process

Our Markov chain is finite, φ-mixing and so we can apply the Functional Central Limit Theorem, Billingsley ([3, p. 174, Thm. 20.1]). We obtain the following result, where $B(t)$ is the standard Brownian Motion (BM).

Theorem 4.1

$$\frac{K(\lfloor nt \rfloor) - M^* nt}{\sigma^* \sqrt{n}} \Rightarrow B(t), \quad n \to \infty, \ t \in [0,1], \tag{10}$$

where nM^* and $n\sigma^{*2}$ are the dominant terms of the mean $M(n)$ and variance $\sigma^2(n)$ of $K(n)$. Of course, from (10),

$$K(n) \sim \mathcal{N}(nM^*, n\sigma^{*2})$$

which conforms to (3). Let us compute $M(n)$ and $\sigma^2(n)$. Note that,

$$\begin{aligned}
P^{*n} &\equiv P^* = 1 \times v_2, \\
v_4 P^* &= v_2 \quad \text{for } \textbf{any} \text{ probability vector } v_4, \\
\psi(1)|_{y=1} &= v_2.
\end{aligned}$$

We obtain (in the sequel, differentiation is always w.r.t. y and we set $y = 1$ after all operations)

$$\begin{aligned}
M(n) &= \mathbb{E}[K(n)] = \kappa'(n) = \psi'(1)P^* \cdot 1 + \psi(1)(n-1)P^{*(n-2)} \cdot P' \cdot 1 \\
&= \psi'(1) \cdot 1 + (n-1)v_2 P' \cdot 1, \\
\mathbb{E}[K(n)(K(n)-1)] &= \kappa''(n) = \psi''(1)P^* \cdot 1 + 2\psi'(1)(n-1)P^{*(n-2)} \cdot P' \cdot 1 \\
&+ \psi(1)(n-1)(n-2)P^{*(n-3)} \cdot (P')^2 \cdot 1 + \psi(1)(n-1)P^{*(n-2)} \cdot P'' \cdot 1 \\
&= \psi''(1) \cdot 1 + 2(n-1)\psi'(1)P^* P' \cdot 1 \\
&+ v_2(n-1)(n-2)(P')^2 \cdot 1 + v_2(n-1)P'' \cdot 1,
\end{aligned}$$

however $P^* = 1 \times v_2$ and $P'|_{y=1} = 1 \times v_3$.
Set $\alpha_1 := \psi'(1) \cdot 1$, $\alpha_2 := \psi''(1) \cdot 1$, $\beta_1 := v_3 \cdot 1$, $\beta_2 := v_2 P'' \cdot 1$. We obtain

$$\begin{aligned}
M(n) &= \alpha_1 + (n-1)\beta_1 = 5/2n - 3/2, \\
\sigma^2(n) &= (\alpha_2 - \beta_2 + \beta_1^2 - \alpha_1^2 + \alpha_1 - \beta_1) + (\beta_2 + \beta_1 - \beta_1^2)n = 5/4n - 3/4.
\end{aligned}$$

Hence

$$M^* = \beta_1 = 5/2, \tag{11}$$
$$\sigma^{*2} = \beta_2 + \beta_1 - \beta_1^2 = 5/4. \tag{12}$$

We recover, of course, (4) and (5). However, here, we don't have to invert matrices, which could be cumbersome for large r.

4.2 Another approach for mean and variance computation. The stationary number of changes distribution

Keeping the lumpability, let us consider a complete Markov chain, $\Pi[1..18, 1..18]$, with states given by the doublets (i, k), i denoting the global state, k denoting the one-step *number* of changes. We have

$$\Pi[(i, k_1), (j, k_2)] = [y^{k_2}]P(i, j),$$

i. e., $\Pi[(i,k_1),(j,k_2)]$ is the probability of jumping from a global state i to a global state j, bringing k_2 changes. We obtain

$$\Pi = \left(\begin{array}{cccccc|cccccc|cccccc}
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
1/8 & 0 & 0 & 1/8 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
\hline
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 1/8 & 1/8 & 0 \\
\hline
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8 \\
0 & 1/8 & 1/8 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 1/8 & 0 & 0 & 1/8
\end{array}\right).$$

Let

$$p := v_2 \cdot P = [(1+3y+3y^2+y^3)/32,\ (y+3y^2+3y^3+y^4)/16,\ (y^2+3y^3+3y^4+y^5)/32].$$

Obviously, $p \cdot 1$ gives the PGF of the stationary number of changes distribution, v_5 say, and π, the stationary vector of Π is given by

$$\pi[(i,k)] = [y^k]p(i).$$

This leads to

$$\begin{aligned}
v_5 &= (1/32, 5/32, 5/16, 5/16, 5/32, 1/32), \\
\pi &= (1/32, 3/32, 3/32, 1/32, 0, 0, 0, 1/16, 3/16, 3/16, 1/16, 0, 0, 0, 1/32, 3/32, 3/32, 1/32).
\end{aligned} \tag{13}$$

Now $K(n) = \sum_1^n X(i)$ where $X(i)$ is the number of changes random variable corresponding to column i. By the ergodic theorem, we know that

$$\begin{aligned}
M(n) &\sim n \sum \pi(m) \cdot m \equiv nM^*, \\
\sigma^2(n) &\sim n\Big[\sum \pi(m) \cdot m^2 - M^{*2}\Big] + 2n \sum_{u=1}^{\infty} \sum_m \sum_h \pi(m)(m - M^*)\Pi(m,h)^u(h - M^*)
\end{aligned} \tag{14}$$

(see for instance Billingsley [3, p. 172]), where, by a slight abuse of notation, we use the shorthand $\sum \pi(m) \cdot m^\alpha$ for $\sum_i \sum_k \pi(i,k) \cdot k^\alpha$. The second sum in (14) is given by

$$\sum_{u=1}^{\infty} \sum_m \sum_h \pi(m) m [\Pi(m,h)^u - \pi(h)] \cdot h. \tag{15}$$

We can now enter the realm of Markov chain potential theory, and $M_1 := \sum_{u \geq 0}(\Pi^u - 1 \times \pi)$ is the Drazin inverse of $I - \Pi$. We refer to Campbell and Meyer [4] for a detailed definition and analysis of the Drazin inverse. We have $M_1 = M_2 - 1 \times \pi$ where $M_2 := [I - \Pi + 1 \times \pi]^{-1} = \sum_{u \geq 0}[\Pi - 1 \times \pi]^u$ is the potential used in Kemeny, Snell and Knapp [7]. Equation (15) becomes

$$\sum_m \sum_h \pi(m) m [M_2 - I](m,h) \cdot h.$$

However, here, a lot of simplifications occur. Indeed, let

$$g := [g_1(y), g_2(y), g_3(y)],$$

where $g_\bullet(y)$ are polynomials in y. Let γ denote the [1..18] vector derived from g in the same way as π was derived from p. We readily check that $\gamma\Pi$ can again be derived, in the same way, from

$$[g_1(1), g_2(1), g_3(1)]P(y).$$

More generally, from a matrix $G(y)[1..3, 1..3]$ we can derive a matrix $\Gamma[1..18, 1..18]$ and $\Gamma\Pi$ can be derived from $G(1)P(y)$. This implies that, for $u \geq 2$, Π^u can be derived from $P^{*(u-1)}P(y)$, i.e., from $1 \times p(y)$. This leads in (15) to a drastic simplification: only the term $u = 1$ remains. We obtain

$$M_2 = I + \Pi - 1 \times \pi,$$

and (15) becomes

$$\sum_m \sum_h \pi(m) m[\Pi(m,h) - \pi(h)] \cdot h.$$

But now, the summation on h corresponds to (we set $y = 1$)

$$P' \cdot 1 - 1 \times p' \cdot 1, \tag{16}$$

which, in turn, is readily checked to be 0. Indeed, we observe by (8), that

$$P'(1,j) = P'(2,j) = P'(3,j), \quad \forall j,$$

and, from $p = v_2 \cdot P$, we derive

$$p' = v_2 \cdot P' = [P'(.,1), P'(.,2), P'(.,3)],$$

which confirms that (16) leads to 0.

To summarize, the dominant term σ^{*2} is simply given by $[\sum \pi(m) \cdot m^2 - M^{*2}]$, which is equivalent to (11) and (12) .

5 Extensions

5.1 $r \times n$ matrices

We briefly sketch the instance of $4 \times n$ matrices. An automatic procedure to generate all useful vectors and matrices will be given in Section 6. We have to consider the following 6 sets (states) of column vectors (written in linear form for convenience):

$$[1] = \{0000, 1111\},\ [2] = \{1001, 0110\},\ [3] = \{0010, 0100, 1101, 1011\},$$
$$[4] = \{0011, 1100\},\ [5] = \{1010, 0101\},\ [6] = \{0001, 1000, 1110, 0111\}.$$

The matrix A is

$$A = \begin{pmatrix} 1+y^4 & 2y^4 & y^3+y^5 & 2y^3 & 2y^5 & y^2+y^4 \\ 2y^2 & y^2+y^6 & y^3+y^5 & 2y^3 & 2y^5 & y^2+y^4 \\ 2y+2y^3 & 2y^3+2y^5 & y^2+2y^4+y^6 & 2y^2+2y^4 & 2y^4+2y^6 & 4y^3 \\ 2y^2 & 2y^4 & y^3+y^5 & y+y^5 & 2y^5 & y^2+y^4 \\ 2y^2 & 2y^4 & y^3+y^5 & 2y^3 & y^3+y^7 & y^2+y^4 \\ 2y+2y^3 & 2y^3+2y^5 & 4y^4 & 2y^2+2y^4 & 2y^4+2y^6 & y+2y^3+y^5 \end{pmatrix},$$

the vector $v = (1, y^2, y^2, y, y^3, y)$, $v_1 = (2, 2, 4, 2, 2, 4)$, $v_2 = (1/8, 1/8, 2/8, 1/8, 1/8, 2/8)$. The Markov chain transition matrix P is given by our usual operators Ω_1 and Ω_2 . The generating function $F(z, y) = zv(1 - zA)^{-1}v_1^T = g/h$, with

$g = 2(y+1)^3 z - 2y(5y^4 + 4y^3 + 14y^2 + 4y + 5)(y-1)^2(y+1)^2 z^2 + 8y^2(y^2 + y + 1)(y^2 + 1)(y-1)^4(y+1)^5 z^3 - 2y^4(11y^6 + 10y^5 + 29y^4 + 28y^3 + 29y^2 + 10y + 11)(y+1)^4(y-1)^6 z^4 + 2y^6(7y^4 - 4y^3 + 10y^2 - 4y + 7)(y+1)^7(y-1)^8 z^5 - 16y^9(y+1)^8(y-1)^{12} z^6$

and

$h = 1 - (y^2+1)(y^2-y+1)(y+1)^3 z + y(2y^8 + 3y^7 + 9y^6 + 9y^5 + 18y^4 + 9y^3 + 9y^2 + 3y + 2)(y-1)^2(y+1)^2 z^2 - y^2(y^2+1)(y^6 + 3y^5 + 3y^4 + 10y^3 + 3y^2 + 3y + 1)(y-1)^4(y+1)^5 z^3 + y^4(2y^8 + 3y^7 + 9y^6 + 9y^5 + 18y^4 + 9y^3 + 9y^2 + 3y + 2)(y-1)^6(y+1)^6 z^4 - y^6(y^2+1)(y^2-y+1)(y-1)^8(y+1)^{11} z^5 + y^9(y-1)^{12}(y+1)^{12} z^6$.

The generating function of the expectations is

$$\frac{\partial F}{\partial y}\left(\frac{z}{16}, 1\right) = \frac{z(3+4z)}{2(1-z)^2},$$

from which we get the expectation as $M(n) = 7n/2 - 2$. Similarly,

$$\frac{\partial^2 F}{\partial y^2}\left(\frac{z}{16}, 1\right) = \frac{z(3+36z+10z^2)}{2(1-z)^3},$$

from which we get

$$\begin{aligned} m_2 &= (7n-3)(7n-4)/4, \\ \sigma^2(n) &= 7n/4 - 1. \end{aligned}$$

Also

$$\begin{aligned} \frac{\partial^3 F}{\partial y^3} &= \frac{3z(1+116z+206z^2+20z^3)}{4(1-z)^4}, \\ \frac{\partial^4 F}{\partial y^4} &= \frac{3z^2(213+1321z+835z^2+32z^3)}{2(1-z)^5}, \end{aligned}$$

and this leads to

$$\begin{aligned} m_3 &= (7n-1)(7n-4)^2/8, \\ m_4 &= 2401n^4/16 - 1715n^3/8 + 1323n^2/16 + 57n/8 - 9, \\ \mu_3 &= 0, \\ \mu_4 &= 147n^2/16 - 55n/8 - 1. \end{aligned}$$

For further use, we have also computed the moments of the $2 \times n$ matrix: this leads to

$$\begin{aligned} m_1 &= 3n/2 - 1, \\ m_2 &= (3n-1)(3n-2)/4, \\ m_3 &= (3n+1)(3n-2)^2/8, \\ m_4 &= (9n^2+3n-10)(9n^2-9n+4)/16. \end{aligned}$$

5.2 Multi-dimensional matrices

Another extension is to go to three-dimensional space. As an example, we consider all $2 \times 2 \times n$ "matrices." The following four states must be considered:

$[1] = \{ \begin{smallmatrix} 00 \\ 00 \end{smallmatrix}, \begin{smallmatrix} 11 \\ 11 \end{smallmatrix} \}$, $[2] = \{ \begin{smallmatrix} 10 \\ 00 \end{smallmatrix}, \begin{smallmatrix} 01 \\ 00 \end{smallmatrix}, \begin{smallmatrix} 00 \\ 01 \end{smallmatrix}, \begin{smallmatrix} 00 \\ 10 \end{smallmatrix}, \begin{smallmatrix} 01 \\ 11 \end{smallmatrix}, \begin{smallmatrix} 10 \\ 11 \end{smallmatrix}, \begin{smallmatrix} 11 \\ 10 \end{smallmatrix}, \begin{smallmatrix} 11 \\ 01 \end{smallmatrix} \}$, $[3] = \{ \begin{smallmatrix} 00 \\ 11 \end{smallmatrix}, \begin{smallmatrix} 10 \\ 10 \end{smallmatrix}, \begin{smallmatrix} 01 \\ 01 \end{smallmatrix}, \begin{smallmatrix} 11 \\ 00 \end{smallmatrix} \}$, $[4] = \{ \begin{smallmatrix} 01 \\ 10 \end{smallmatrix}, \begin{smallmatrix} 10 \\ 01 \end{smallmatrix} \}$. And

$$A = \begin{pmatrix} 1+y^4 & y^3+y^5 & 2y^4 & 2y^6 \\ 4y+4y^3 & y^2+6y^4+y^6 & 4y^3+4y^5 & 4y^5+4y^7 \\ 4y^2 & 2y^3+2y^5 & y^2+2y^4+y^6 & 4y^6 \\ 2y^2 & y^3+y^5 & 2y^4 & y^4+y^8 \end{pmatrix},$$

$v = (1, y^2, y^2, y^4)$, $v_1 = (2, 8, 4, 2)$, $v_2 = (1/8, 1/2, 1/4, 1/8)$. Therefore $F(z, y) = g/h$, with

$g = 2(y^4+6y^2+1)z - 16y^2(y^6+y^5+2y^4+2y^2+y+1)(y-1)^2z^2 + 2y^4(7y^6-10y^5+17y^4+4y^3+17y^2-10y+7)(y+1)^2(y-1)^4z^3 - 16y^8(y+1)^4(y-1)^8z^4$

and

$h = 1 + (-y^8-2y^6-10y^4-1-2y^2)z + 2y^2(y^8+5y^6+4y^4+5y^2+1)(y-1)^2(y+1)^2z^2 - y^4(y^8+2y^6+10y^4+1+2y^2)(y-1)^4(y+1)^4z^3 + y^8(y-1)^8(y+1)^8z^4$;

$$\frac{\partial F}{\partial y}\Big(\frac{z}{16}, 1\Big) = \frac{2z(1+z)}{(1-z)^2}$$

from which we get the expectation as

$$M(n) = 4n - 2. \tag{17}$$

Similarly,

$$\frac{\partial^2 F}{\partial y^2}\Big(\frac{z}{16}, 1\Big) = \frac{z(3+24z+5z^2)}{(1-z)^3},$$

from which we get

$$\begin{aligned} m_2 &= (8n-3)(2n-1), \\ \sigma^2(n) &= 2n-1. \end{aligned} \tag{18}$$

The next moments are given by

$$\begin{aligned} m_3 &= 2(8n-1)(2n-1)^2, \\ m_4 &= 256n^4 - 320n^3 + 108n^2 + 21n/2 - 21/2, \\ \mu_3 &= 0 \\ \mu_4 &= 12n^2 - 11n/2 - 5/2 \end{aligned}$$

We can, of course, also derive the Markov chain transition matrix P from A, only with more algebra and we can compute the number of changes PGF as given by

$$\kappa(y, n) = \psi(1)P^{n-1} \cdot 1.$$

6 An automatic procedure to generate H, A and P for $r \times n$ rectangular arrays

Let us first assume r even. We can classify our states into 4 classes. Set $s := r/2 - 1$. We can choose a canonical vector in each class that we represent in a compact way by $c := (d_1, a, b, d_2)$, where d_1 is the vector first bit, a is the integer representation of the s bits following d_1, b is the integer representation of the **reversed** next s bits and d_2 is the vector last bit. In each class $i, i = 1..4$, we have $size[i]$ states and, in each state of class i, we have $t[i]$ vectors (by reversing and complementing). Let $\bar{a} := 2^s - 1 - a$. We derive Table 1. Of course, $\sum_{i=1}^4 size[i]t[i] = 2^r$. Let us note that the gain in states number is given by $ST/2^r$, where $ST := \sum_{i=1}^4 size[i] = 2^s + 2^r/4$. For example, for $r = 4$, we obtain

$$\begin{aligned} class[1] &:= (0,0,0,0), (0,1,1,0), \\ class[2] &:= (0,0,1,0), \\ class[3] &:= (0,0,1,1), (0,1,0,1), \\ class[4] &:= (0,0,0,1). \end{aligned}$$

Canonical vector	$size[i]$	$t[i]$	range
$(0,a,a,0)$	2^s	2	$a = 0 \dots 2^s-1$
$(0,a,b,0), a<b$	$2^s(2^s-1)/2$	4	$a=0\dots b-1, b=1\dots 2^s-1$
$(0,a,\bar{a},1)$	2^s	2	$a = 0 \dots 2^s-1$
$(0,a,b,1), a<\bar{b}$	$2^s(2^s-1)/2$	4	$a=0\dots \bar{b}-1, b=0\dots 2^s-2$

Table 1: Canonical representation.

The generic 4×4 matrix H^* can be computed with the number of changes GF $\varphi(y)$ from a state c_1 of class i to a state c_2 of class j, which is given by
if $t[i]=2$, then $\varphi(y) = y^\gamma + y^{r-\gamma}$, where $\gamma = D(c_1,c_2)$,
if $t[i]=4$, then $\varphi(y) = y^{\gamma_1} + y^{r-\gamma_1} + y^{\gamma_2} + y^{r-\gamma_2}$, where $\gamma_1 = D(c_1,c_2), \gamma_2 = D(\mathrm{rev}(c_1),c_2)$,
and

$$D(c_1,c_2) := |c_1[1]-c_2[1]| + \delta(c_1[2],c_2[2]) + \delta(c_1[3],c_2[3]) + |c_1[4]-c_2[4]|,$$

$\delta(a,b)$ is the Hamming distance between a and b, i.e., the number of places where their bits are different and $\mathrm{rev}(c)$ is the reversal operator on c. The matrix H^* is given by

$$\begin{aligned}
H^*[1,1] &= y^{2\delta(a,\alpha)} + y^{r-2\delta(a,\alpha)},\\
H^*[1,2] &= y^{\delta(a,\alpha)+\delta(a,\beta)} + y^{r-\delta(a,\alpha)-\delta(a,\beta)},\\
H^*[1,3] &= y^{1+\delta(a,\alpha)+\delta(a,\bar{\alpha})} + y^{r-1-\delta(a,\alpha)-\delta(a,\bar{\alpha})},\\
H^*[1,4] &= y^{1+\delta(a,\alpha)+\delta(a,\beta)} + 2y^{r-1-\delta(a,\alpha)-\delta(a,\beta)},\\
H^*[2,1] &= 2y^{\delta(a,\alpha)+\delta(b,\alpha)} + 2y^{r-\delta(a,\alpha)-\delta(b,\alpha)},\\
H^*[2,2] &= y^{\delta(a,\alpha)+\delta(b,\beta)} + y^{r-\delta(a,\alpha)-\delta(b,\beta)} + y^{\delta(b,\alpha)+\delta(a,\beta)} + y^{r-\delta(b,\alpha)-\delta(a,\beta)},\\
H^*[2,3] &= y^{1+\delta(a,\alpha)+\delta(b,\bar{\alpha})} + y^{r-1-\delta(a,\alpha)-\delta(b,\bar{\alpha})} + y^{1+\delta(b,\alpha)+\delta(a,\bar{\alpha})} + y^{r-1-\delta(b,\alpha)-\delta(a,\bar{\alpha})},\\
H^*[2,4] &= y^{1+\delta(a,\alpha)+\delta(b,\beta)} + y^{r-1-\delta(a,\alpha)-\delta(b,\beta)} y^{1+\delta(b,\alpha)+\delta(a,\beta)} + y^{r-1-\delta(b,\alpha)-\delta(a,\beta)},\\
H^*[3,1] &= y^{1+\delta(a,\alpha)+\delta(\bar{a},\alpha)} + y^{r-1-\delta(a,\alpha)-\delta(\bar{a},\alpha)},\\
H^*[3,2] &= y^{1+\delta(a,\alpha)+\delta(\bar{a},\beta)} + y^{r-1-\delta(a,\alpha)-\delta(\bar{a},\beta)},\\
H^*[3,3] &= y^{\delta(a,\alpha)+\delta(\bar{a},\bar{\alpha})} + y^{r-\delta(a,\alpha)-\delta(\bar{a},\bar{\alpha})},\\
H^*[3,4] &= y^{\delta(a,\alpha)+\delta(\bar{a},\beta)} + y^{r-\delta(a,\alpha)-\delta(\bar{a},\beta)},\\
H^*[4,1] &= 2y^{1+\delta(a,\alpha)+\delta(b,\alpha)} + 2y^{r-1-\delta(a,\alpha)-\delta(b,\alpha)},\\
H^*[4,2] &= y^{1+\delta(a,\alpha)+\delta(b,\beta)} + y^{r-1-\delta(a,\alpha)-\delta(b,\beta)} + y^{1+\delta(b,\alpha)+\delta(a,\beta)} + y^{r-1-\delta(b,\alpha)-\delta(a,\beta)},\\
H^*[4,3] &= y^{\delta(a,\alpha)+\delta(b,\bar{\alpha})} + y^{r-\delta(a,\alpha)-\delta(b,\bar{\alpha})} + y^{2+\delta(b,\alpha)+\delta(a,\bar{\alpha})} + y^{r-2-\delta(b,\alpha)-\delta(a,\bar{\alpha})},\\
H^*[4,4] &= y^{\delta(a,\alpha)+\delta(b,\beta)} + y^{r-\delta(a,\alpha)-\delta(b,\beta)} + y^{2+\delta(b,\alpha)+\delta(a,\beta)} + y^{r-2-\delta(b,\alpha)-\delta(a,\beta)}.
\end{aligned}$$

To obtain the matrix $H(ST\times ST)$, it suffices to instantiate H^*, with the ranges given in Table 1. This is done by Maple in a matter of seconds. Each state is represented by its canonical vector. The vectors v_1 and v_2 are easily computed: $v_1[k] = t[class(k)]$ and $v_2[k] = t[class(k)]/2^r$. To obtain v, we compute, in each state, the number of changes γ of its canonical vector. The corresponding component of v is given by y^γ. The case r odd is left to the interested reader.

7 The general case

7.1 Mean and variance

Consider an $r \times n$ array. The number of changes $K(n)$ is given by

$$K = V_1 + \ldots + V_n + H_{12} + \ldots + H_{n-1,n}.$$

These are random variables, V_i is the number of vertical changes in column i, and $H_{i-1,i}$ the number of horizontal changes between columns $i-1$ and i.

The probability generating function for the vertical changes is given by

$$G_1 := \left(\frac{1+y}{2}\right)^{r-1}. \tag{19}$$

with expectation: $M_1 := (r-1)/2$.

For two adjacent columns, we obtain

$$G_2 := \left(\frac{1+y}{2}\right)^{r}. \tag{20}$$

with mean $M_2 = r/2$.

Altogether, the mean $M(n)$ is given by

$$m_1 = M(n) = nM_1 + (n-1)M_2 = n\frac{r-1}{2} + (n-1)\frac{r}{2} = nr - \frac{n+r}{2}.$$

It is symmetric in n and r, as it should. Also, as n gets large, we derive M^*:

$$M^* = r - 1/2 \tag{21}$$

Now we note that all $V_k, H_{i-1,i}$ are **two by two** independent. Of course, the triplets are not: for instance, if $V_1 = 0$ and $H_{1,2} = 0$, then $V_2 = 0$ too. But, for the variance computation, the doublet independence is enough.

So we consider $\mathbb{E}(K^2)$.

$$\mathbb{E}(K^2) = \sum_{i=1}^{n} \mathbb{E}(V_i^2) + \sum_{i=2}^{n} \mathbb{E}(H_{i-1,i}^2) + 2\sum_{i=1}^{n}\sum_{j=2}^{n} \mathbb{E}(V_i H_{j-1,j}) + 2\sum_{1\le i<j\le n} \mathbb{E}(V_i V_j) + 2\sum_{2\le i<j\le n} \mathbb{E}(H_{i-1,i} H_{j-1,j}).$$

Note that

$$\begin{aligned}
\mathbb{E}(V_i^2) &= \frac{r(r-1)}{4}, \quad \text{by } G_1, \\
\mathbb{E}(H_{i-1,i}^2) &= \frac{r(r+1)}{4}, \quad \text{by } G_2, \\
\mathbb{E}(V_i V_j) &= M_1^2 = \frac{(r-1)^2}{4}, \\
\mathbb{E}(H_{i-1,i} H_{j-1,j}) &= M_2^2 = \frac{r^2}{4}, \\
\mathbb{E}(V_i H_{j-1,j}) &= M_1 M_2 = \frac{r(r-1)}{4}.
\end{aligned}$$

This leads to

$$\mathbb{E}(K^2) = n\frac{r(r-1)}{4} + (n-1)\frac{r(r+1)}{4} + 2n(n-1)\frac{r(r-1)}{4} + n(n-1)\frac{(r-1)^2}{4} + (n-1)(n-2)\frac{r^2}{4} \tag{22}$$

or

$$m_2 = \mathbb{E}(K^2) = \frac{(2r-1)^2 n(n-1)}{4} + \frac{r(r-1)}{4}.$$

And the variance is given by

$$\sigma^2(n) = \frac{nr}{2} - \frac{n+r}{4}.$$

Also, as n gets large, we derive σ^{*2}:

$$\sigma^{*2} = r/2 - 1/4. \tag{23}$$

Of course, our analysis of Section 4.1 is still valid: we don't need the explicit form of A and P, only their existence is necessary. We can even get the stationary number of changes PGF. Indeed, again by local independence, this is given by the binomial

$$G_3 = G_1 G_2 = \left(\frac{1+y}{2}\right)^{r+r-1} = \left(\frac{1+y}{2}\right)^{2r-1}. \tag{24}$$

For instance, for $r = 3$, we derive

$$G_3 = \tfrac{1}{32}[1 + 5y + 10y^2 + 10y^3 + 5y^4 + y^5],$$

which immediately leads to v_5 as given by (13). The mean of (24) is given by $(2r-1)/2$ and the variance by $(2r-1)/4$ which correspond to M^*, σ^{*2}, as given by (21), (23).

The computation of mean and variance of K for the general $k-$dimensional array $d_1 \times d_2 \times \ldots \times d_k$ goes as follows.

Set $M_1[0] := 0$, $\mathbb{E}(V^2)[0] := 0$, $\sigma^2[0] := 0$, $d_0 := 1$, $D[0] := 1$, $D[i] := \prod_1^i d_u, i \geq 1$. Then we compute for all $i \geq 1$

$$\begin{aligned}
n &:= d_i, \\
M_2 &:= D[i-1]/2, \\
M_1[i] &:= nM_1[i-1] + (n-1)M_2, \\
\mathbb{E}(HH) &:= M_2^2, \\
\mathbb{E}(VV) &:= M_1[i-1]^2, \\
\mathbb{E}(VH) &:= M_1[i-1]M_2, \\
\mathbb{E}(H^2) &:= D[i-1](D[i-1]+1)/4, \\
m_2[i] := \mathbb{E}(V^2)[i] &:= n\mathbb{E}(V^2)[i-1] + (n-1)\mathbb{E}(H^2) + 2n(n-1)\mathbb{E}(VH) + n(n-1)\mathbb{E}(VV) \\
&\quad + (n-1)(n-2)\mathbb{E}(HH),
\end{aligned}$$

and finally,

$$\sigma^2[i] := \mathbb{E}(V^2)[i] - M_1[i]^2.$$

We recover, of course, $M(n) = M_1[2]$, $\sigma^2(n) = \sigma^2[2]$, and, for instance,

$$\begin{aligned}
M_1[3] &= (3d_1d_2d_3 - d_2d_3 - d_1d_3 - d_1d_2)/2, \\
m_2[3] &= (3d_1d_2d_3 - d_2d_3 - d_1d_3 - d_1d_2 + 1)(3d_1d_2d_3 - d_2d_3 - d_1d_3 - d_1d_2)/4, \\
\sigma^2[3] &= M_1[3]/2, \\
M_1[4] &= (4d_1d_2d_3d_4 - d_2d_3d_4 - d_1d_3d_4 - d_1d_2d_4 - d_1d_2d_3)/2, \\
m_2[4] &= (4d_1d_2d_3d_4 - d_2d_3d_4 - d_1d_3d_4 - d_1d_2d_4 - d_1d_2d_3 + 1)\cdot \\
&\quad \cdot(4d_1d_2d_3d_4 - d_2d_3d_4 - d_1d_3d_4 - d_1d_2d_4 - d_1d_2d_3)/4, \\
\sigma^2[4] &= M_1[4]/2.
\end{aligned}$$

Setting $d_1 = 2, d_2 = 2, d_3 = n$, we recover (17) and (18) . More generally,

$$M_1[i] = \frac{D[i]}{2}\Big[i - \sum_{k=1}^{i} \frac{1}{d_k}\Big], \tag{25}$$

$$\sigma^2[i] = M_1[i]/2, \tag{26}$$

which can be proved by induction.
Such a simple formula can be explained by the special structure of our lattice: actually, for any graph (without multiples edges and loops), let E be the number of edges. Mark each node with 0 or 1 (with probability $1/2$). Then the number of changes mean and variance are $E/2$ and $E/4$ again by independence. Equations (25) and (26) reflect the E counting in our lattice. Finally, we can get the stationary number of changes PGF $G_3(y)[i]$ for the $i-$dimensional array. Indeed, let $\kappa(y)[l]$ be the number of changes PGF related to the $l-$dimensional array. Then

$$G_3(y)[i] = \kappa(y)[i-1]G_2(y), \quad \text{with} \quad G_2 = \left(\frac{1+y}{2}\right)^{S[i-1]}.$$

7.2 The limiting processes

If the last dimension (or several of them) gets large, we obtain again a limiting Gaussian distribution: we don't need the explicit form of the matrix A, only its existence is necessary. The mean and variance are the only parameters we need for the Gaussian. More generally, we obtain a limiting Gaussian field. Indeed assume that among the i dimensions, j_2 stay fixed to $d_1 \dots d_{j_2}$ and that the remaining j_1 dimensions get large: $\delta_1 \dots \delta_{j_1}$. Set $D(j_2) := \prod_1^{j_2} d_k$ and $\Delta(j_1) := \prod_1^{j_1} \delta_u$. The mean is asymptotically given by $\Delta(j_1)D(j_2)(J^* + B)$, where

$$J^* = \Big[i - \sum_{k=1}^{j_2} \frac{1}{d_k}\Big]/2, \quad B = -\frac{1}{2}\sum_{u=1}^{j_1} \frac{1}{\delta_u}$$

Set

$$\sigma^* = \sqrt{J^* D(j_2)/2}.$$

We obtain the following result, where $B(t_1 \dots t_{j_1})$ is a Gaussian field.

Theorem 7.1

$$\frac{K(\lfloor \delta_1 t_1 \rfloor \dots \lfloor \delta_{j_1} t_{j_1} \rfloor) - D(j_2)\prod_1^{j_1} \delta_u t_u [J^* - \frac{1}{2}\sum_{u=1}^{j_1} \frac{1}{\delta_u t_u}]}{\sigma^* \sqrt{\Delta(j_1)}} \Rightarrow B(t_1 \dots t_{j_1}), \quad \delta_1 \dots \delta_{j_1} \to \infty,\ t_1 \dots t_{j_1} \in [0,1],$$

and

$$\mathbb{E}[B(t_1 \dots t_{j_1})B(\tau_1 \dots \tau_{j_1})] = \prod_1^{j_1} \min(t_u, \tau_u)$$

7.3 The next moments

We clearly need an automatic procedure. We first define the following indices and quantities:

- dim is the exponent of K in the moments. The case $dim = 1, 2$ has been computed already. We will deal with $dim = 3, 4$.
- l is the level of computation: $l = 0, 1, 2 \dots$
- s is the dimension of the matrix: $s = 2, 3$ here

- $mV[dim, l, s]$ is related to the moments of the number of changes in one slice:

$$\begin{aligned} mV[dim,0,2] &= \mathbb{E}(K(2,1)^{dim}), \\ mV[dim,1,2] &= \mathbb{E}(K(1,d_1)^{dim}), \end{aligned}$$

these quantities are immediately derived from G_1 as given by (19).

$$mV[dim,0,3] = \mathbb{E}(K(2,2)^{dim}),$$

this quantity is derived from the case $r = 2$ in Section 5.1

$$\begin{aligned} mV[dim,1,3] &= \mathbb{E}(K(2,d_1)^{dim}), \\ mV[dim,2,3] &= \mathbb{E}(K(d_1,d_2)^{dim}), \end{aligned}$$

- $mH[dim, l, s]$ is related to the moments of the number of changes $H(d_i, d_j)$ between two slices of dimension $d_i \times d_j$. These quantities are derived from G_2 as given by (20).

$$\begin{aligned} mH[dim,0,2] &= \mathbb{E}(H(2,1)^{dim}), \\ mH[dim,1,2] &= \mathbb{E}(H(1,d_1)^{dim}), \\ mH[dim,0,3] &= \mathbb{E}(H(2,2)^{dim}), \\ mH[dim,1,3] &= \mathbb{E}(H(2,d_1)^{dim}), \\ mH[dim,2,3] &= \mathbb{E}(H(d_1,d_2)^{dim}), \end{aligned}$$

- $mVC[dim, l, s]$ is related to particular matrices we need in the computation

$$\begin{aligned} mVC[dim,0,2] &= \mathbb{E}(K(2,2)^{dim}), \\ mVC[dim,1,2] &= \mathbb{E}(K(2,d_1)^{dim}), \end{aligned}$$

these quantities are derived from the case $r = 2$ in Section 5.1

$$mVC[dim,2,2] = \mathbb{E}(K(d_1,d_2)^{dim}).$$

We note that, for rectangular arrays,

$$m_{dim} = mVC[dim,2,2].$$

Also, we set

$$mVC[dim,0,3] = \mathbb{E}(K(2,2,2)^{dim}),$$

This quantity is derived from the case $2 \times 2 \times n$ of Section 5.2.

$$\begin{aligned} mVC[dim,1,3] &= \mathbb{E}(K(2,2,d_1)^{dim}), \\ mVC[dim,2,3] &= \mathbb{E}(K(2,d_1,d_2)^{dim}), \\ mVC[dim,3,3] &= \mathbb{E}(K(d_1,d_2,d_3)^{dim}). \end{aligned}$$

We note that, for 3−dimensional arrays,

$$m_{dim} = mVC[dim,3,3].$$

- $T4[dim, l, s]$ is related to the correlation $C(d_i, d_j)$ between two slices of dimension $d_i \times d_j$.

$$\begin{aligned} T4[dim,0,2] &= \mathbb{E}(C(2,1)^{dim}), \\ T4[dim,1,2] &= \mathbb{E}(C(1,d_1)^{dim}), \\ T4[dim,0,3] &= \mathbb{E}(C(2,2)^{dim}), \\ T4[dim,1,3] &= \mathbb{E}(C(2,d_1)^{dim}), \\ T4[dim,2,3] &= \mathbb{E}(C(d_1,d_2)^{dim}). \end{aligned}$$

Note that, from one level to another, a left shift is operated on the indices, with the addition of an index $d_\bullet$ on the right.

Consider now the number of changes $K(n)$ given by

$$K = V_1 + \ldots + V_n + H_{12} + \ldots + H_{n-1,n}.$$

These are random variables, V_i is the number of vertical changes in slice i (of any dimension) and $H_{i-1,i}$ the number of horizontal changes between slice $i-1$ and i.

7.3.1 Moments of order 3 and 4, 2−dimensional arrays($s = 2$)

How can we analyze $\mathcal{M}_3 = K^3$ at level l? By induction, we assume that we have already computed all necessary quantities for lower levels. We note that the only possibly correlated term must be of the form:

$$C = V_i V_{i+1} H_{i,i+1}$$

and there are $6(n-1)$ terms of this kind (here, n stands for n_l). However, there are $6n(n-1)/2\cdot(n-1)$ terms of the form

$$V_k V_j H_{i,i+1}.$$

So we must first compute $\mathbb{E}(C)$. This is done as follows (we drop the indices when unnecessary and use $\overline{X}$ for $\mathbb{E}(x)$). We have

$$\mathbb{E}((V_1 + V_2 + H_{1,2})^3) = 2\overline{V^3} + 6\overline{V}\,\overline{V^2} + 6\overline{V^2}\,\overline{H} + 6\overline{V}\,\overline{H^2} + \overline{H^3} + 6\overline{C}.$$

We know all quantities except $\overline{C}$. Indeed, $\overline{V^k}$ and $\overline{H^k}$ are given respectively by $mV[k,l,s]$ and $mH[k,l,s]$. So we obtain a value for $T4[3,l,s]$. Let us summarize our derivation by a procedure $M3(l,s,T4)$ which is easily written down in MAPLE

Now we return to $\mathcal{M}_3$. If we ignore the dependency, this is given by (an automatic procedure to generate such an expression is given later on)

$$\begin{aligned}
\widetilde{\mathcal{M}_3} &= n(n-1)(n-2)V[1]^3 + 3n(n-1)^2V[1]^2H[1] \\
&+ V[1](3n(n-1)(n-2)H[1]^2 + 3n(n-1)V[2] + 3n(n-1)(n-2)H[2]V[1]) \\
&+ 3n(n-1)V[2]H[1] + 3(n-1)(n-2)H[1]H[2] \\
&+ (n-1)(n-2)(n-3)H[3] + (n-1)H[3] + nV[3]
\end{aligned} \tag{27}$$

where $V[k]$ stands for independent copies of V^k and similarly for H. Actually, we generalyze here (22). Again n stands for d_l. To take care of the dependency, we must replace, in $\widetilde{\mathcal{M}_3}$, $6n(n-1)^2/2V[1]^2H[1]$ by

$$6(n-1)C + [6n(n-1)^2/2 - 6(n-1)]V[1]^2H[1]$$

i. e., we compute

$$\widehat{\mathcal{M}_3} = \widetilde{\mathcal{M}_3} + 6(n-1)[C - V[1]^2H[1]].$$

Finally, $\mathbb{E}(\mathcal{M}_3)$ is derived from $\widehat{\mathcal{M}_3}$ by simply replacing C and all V^k, H^k by their explicit mean values.

We now want to generate automatically expressions like (27). We know that all coefficients $P(n,l,k)$ of H^lV^k are polynomials in n of degree dim. So we compute

$$\left(\sum_{i=1}^{d} V_i + \sum_{i=1}^{d-1} H_i\right)^{dim}$$

for $d = dim{+}1 \cdots 2dim{+}1$. This gives $dim{+}1$ values for each coefficient of $H^lV^k : c_{dim+1}(l,k) \ldots c_{2dim+1}(l,k)$. A classical Lagrange interpolation gives

$$P(n,l,k) = \sum_{d=dim+1}^{2dim+1} c_d(l,k)\frac{\prod_{i=dim+1, i\neq d}^{2dim+1}(n-i)}{\prod_{i=dim+1, i\neq d}^{2dim+1}(d-i)}$$

and expressions of type (27) are automatically available. All this analysis leads to a value for $mVC[3,l,s]$. We summarize again our computations in a procedure $TC3(l,s,mVC)$, easily written down in MAPLE.

How can we, for instance, compute $m_3 = \mathbb{E}(K(d_1,d_2)^3) = mVC[dim,2,2]$ for $s=2$? We could start from level 0 but we know already all quantities at level 1. So we call $M3(1,2,T4)$. This gives $T4(3,1,2) = d_1(d_1-1)^2/8$, *which is equivalent to the independent case.* Although the RV are *not* independent, the correlation of order 3 appears to be equal to 0. This is basically explained by the following simple observation: if we compute directly $T4(3,0,2)$, we obtain 1/4 which is exactly the independent case value. This is the reason why *all* μ_3 are equal to 0.

Finally, we call $TC3(2,2,mVC)$. This leads to

$$m_3 = mVC[3,2,2] = (3-d_1-d_2+2d_1d_2)(-d_1-d_2+2d_1d_2)^2/8.$$

Let us now turn to $\mathcal{M}_4 = K^4$. It is easy to check that the correlated terms are now given by

$$C = V_iV_{i+1}H_{i,i+1}^2 + V_i^2V_{i+1}H_{i,i+1}V_iV_{i+1}^2H_{i,i+1}$$

and there are $12(n-1)$ terms of this kind. There are $12n(n-1)/2\cdot(n-1)$ terms of the form

$$V_kV_jH_{i,i+1}^2 + V_k^2V_jH_{i,i+1}V_kV_j^2H_{i,i+1}.$$

Now we compute $\mathbb{E}(C)$ from

$$\mathbb{E}((V_1+V_2+H_{1,2})^4) = 12\overline{H^2V^2} + 8\overline{V}\overline{V^3} + 8\overline{H^3}\overline{V} + 8\overline{H}\overline{V^3} + 6\overline{V^2}^2 + 2\overline{V^4} + \overline{H^4} + 12\overline{C}.$$

This is summarized in the procedure $M4(l,s,T4)$ which gives $T4(4,l,s)$. Now we define a procedure $TC4(l,s,mVC)$ quite similar to $TC3(l,s,mVC)$. (We omit the details). This gives $mVC(4,l,s)$.

Let us apply our procedures to the computation of m_4. Calling $M4(1,2,T4)$, we derive

$$T4[4,1,2] = (d_1-1)(3d_1^3-2d_1^2-d_1+2)/16$$

which is *different* from the independent case. This can also be explained by directly computing $T4[4,0,2]=1$ different from the independent equivalent given by 7/8. Now we apply $TC4(2,2,mVC)$, leading to

$$\begin{aligned}
m_4 &= mVC[4,2,2] = (3d_1^2/16 - 3d_1^3/8 - 11d_1/8 + d_1^4/16 + 3/2)\\
&+ (-d_1^4/2 - 11/8 + 5d_1^3/2 + 13d_1/8 - 15d_1^2/8)d_2\\
&+ (-15d_1/8 + 3d_1^4/2 + 3/16 + 45d_1^2/8 - 6d_1^3)d_2^2\\
&+ (5d_1/2 - 2d_1^4 - 6d_1^2 + 6d_1^3 - 3/8)d_2^3 + (1/16 - 2d_1^3 + d_1^4 + 3d_1^2/2 - d_1/2)d_2^4.
\end{aligned}$$

7.3.2 Moments of order 3 and 4, 3−dimensional arrays($s=3$)

Now we start at level 0, with $mVC[3,0,3]=270, mVC[4,0,3]=3957/2$. Applying $M3(0,3,T4)$ leads to $T4[3,0,3]=8$. Applying $TC3(1,3,mVC)$ leads to $mVC[3,1,3]$. Now we need $mV[dim,1,3]$ which is available from the already computed value $mVC[dim,1,2]$. Using $M3(1,3,T4)$ leads to $T4[3,1,3]$. Now $TC3(2,3,mVC)$ gives $mVC[3,2,3]$. The value $mV[dim,2,3]$ is available from $mVC[dim,2,2]$ already computed. Applying $M3(2,3,T4)$ leads to $T4[3,2,3]$. Finally applying $TC3(3,3,mVC)$ gives

$$\begin{aligned}
m_3 &= mVC[3,3,3] = \left[\left(27d_1^3/8 - 27d_1^2/8 + 9d_1/8 - 1/8\right)d_2^3 + \left(-27d_1^3/8 + 9d_1^2/4 - 3d_1/8\right)d_2^2\right.\\
&+ \left.\left(9d_1^3/8 - 3d_1^2/8\right)d_2 - d_1^3/8\right]d_3^3 + \left[\left(-27d_1^3/8 + 9d_1^2/4 - 3d_1/8\right)d_2^3\right.\\
&+ \left.\left(9d_1^3/4 + 21d_1^2/8 - 9d_1/4 + 3/8\right)d_2^2 + \left(-3d_1^3/8 - 9d_1^2/4 + 3d_1/4\right)d_2 + 3d_1^2/8\right]d_3^2\\
&+ \left[\left(9d_1^3/8 - 3d_1^2/8\right)d_2^3 + \left(-3d_1^3/8 - 9d_1^2/4 + 3d_1/4\right)d_2^2 + 3d_2d_1^2/4\right]d_3 - d_1^3d_2^3/8 + 3d_1^2d_2^2/8
\end{aligned}$$

We will skip the details of the computation of m_4, which are quite similar. We finally obtain $m_4 = mV[4,3,3]$ which is of course a quite long expression, but perfectly computed with MAPLE.

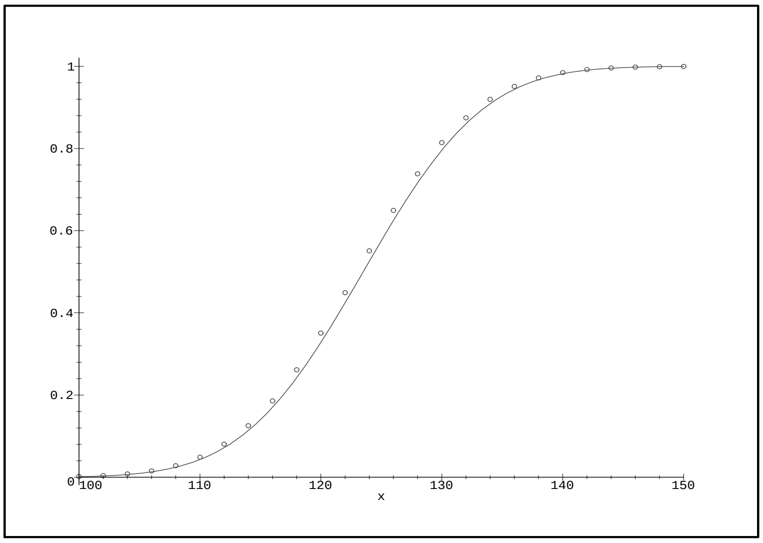

Figure 1: K distribution (exact = circle, asymptotic Gaussian = line)

8 Simulations

First of all, for $n = 50$, $r = 3$, we have compared the exact distribution, given by the PGF $\kappa(n)$ with a $\mathcal{N}(M(n), \sigma^2(n))$. This is given in Figure 1. Also, we have done a simulation of 1000 arrays. The comparison between the observed distribution of K and the Gaussian is given in Figure 2. A typical trajectory, for $n = 200$, normalized according to Thm 4.1, is given in Figure 3.

The stationary distribution of one column to column number of changes (as given by v_5 in (13)) is now compared to the observed one (based on $1000 \times (n-1)$ columns). This is given in Figure 4. As $r \uparrow \infty$, the binomial leads to a Gaussian.

Acknowledgements
The pertinent comments and suggestions of Y. Baryshnikov and of an anonymous referee led to substantial improvements in our analysis.

References

[1] R.J. Baxter. *Exactly Solved Models in Statistical Mechanics.* Academic Press, 1982.

[2] E.A. Bender. Central and local limit theorems applied to asymptotics enumeration. *Journal of Combinatorial Theory,* Series A, 15:91–111, 1973.

[3] P. Billingsley. *Convergence of Probability Measures.* Wiley, 1968.

[4] S. L. Campbell and C. D. Meyer. *Generalized Inverse of Linear Transformations.* Pitman, 1979.

[5] L. Carlitz. Rectangular arrays of zeros and ones. *Duke Mathematical Journal*, 39:153–164, 1972.

[6] J.M. Dufour, M. Hallin, and I. Mizera. Generalized runs tests for heteroscedastic time series. *Nonparametric Statistics*, 9:39–86, 1998.

[7] J. G. Kemeny, J. L. Snell, and A. W. Knapp. *Denumerable Markov Chains.* Van Nostrand, 1966.

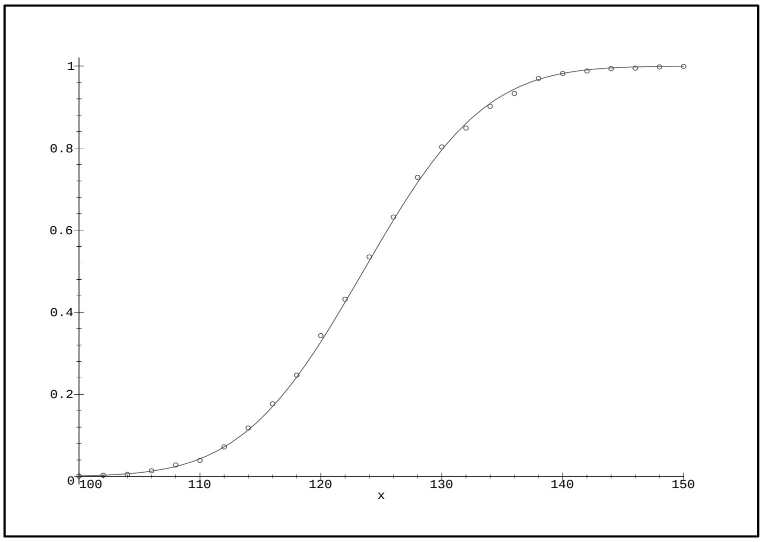

Figure 2: K distribution (observed = circle, asymptotic Gaussian = line)

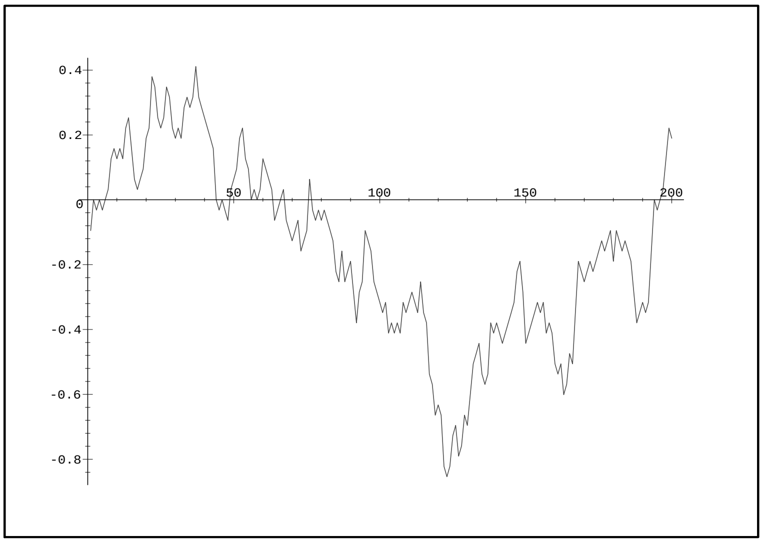

Figure 3: A typical normalized trajectory

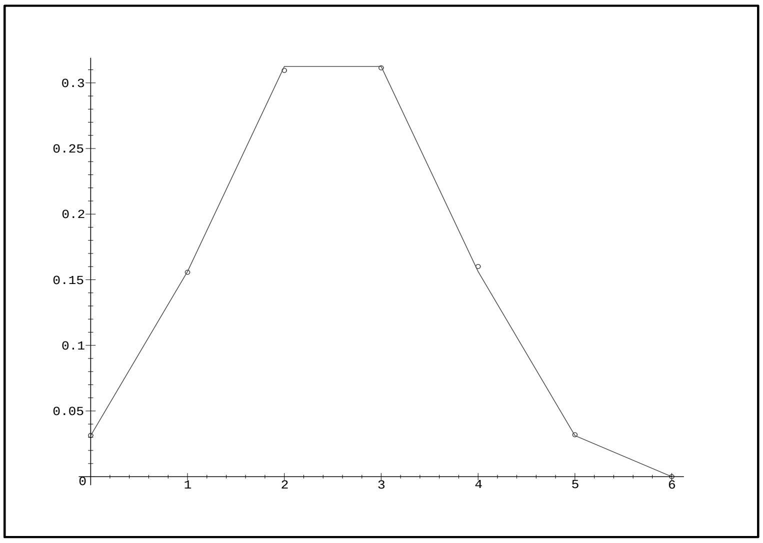

Figure 4: Stationary distribution (observed = circle, asymptotic = line)